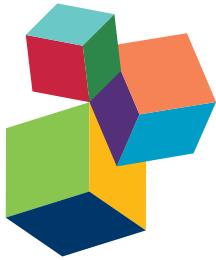


ARCHAEAL CELL ENVELOPE AND SURFACE STRUCTURES

EDITED BY: Sonja-Verena Albers and Mecky Pohlschroder

PUBLISHED IN: Frontiers in Microbiology



Frontiers Copyright Statement

© Copyright 2007-2016 Frontiers Media SA. All rights reserved.

All content included on this site, such as text, graphics, logos, button icons, images, video/audio clips, downloads, data compilations and software, is the property of or is licensed to Frontiers Media SA ("Frontiers") or its licensees and/or subcontractors. The copyright in the text of individual articles is the property of their respective authors, subject to a license granted to Frontiers.

The compilation of articles constituting this e-book, wherever published, as well as the compilation of all other content on this site, is the exclusive property of Frontiers. For the conditions for downloading and copying of e-books from Frontiers' website, please see the Terms for Website Use. If purchasing Frontiers e-books from other websites or sources, the conditions of the website concerned apply.

Images and graphics not forming part of user-contributed materials may not be downloaded or copied without permission.

Individual articles may be downloaded and reproduced in accordance with the principles of the CC-BY licence subject to any copyright or other notices. They may not be re-sold as an e-book.

As author or other contributor you grant a CC-BY licence to others to reproduce your articles, including any graphics and third-party materials supplied by you, in accordance with the Conditions for Website Use and subject to any copyright notices which you include in connection with your articles and materials.

All copyright, and all rights therein, are protected by national and international copyright laws.

The above represents a summary only. For the full conditions see the Conditions for Authors and the Conditions for Website Use.

ISSN 1664-8714

ISBN 978-2-88919-773-6

DOI 10.3389/978-2-88919-773-6

About Frontiers

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

Frontiers Journal Series

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

Dedication to Quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews.

Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view.

By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

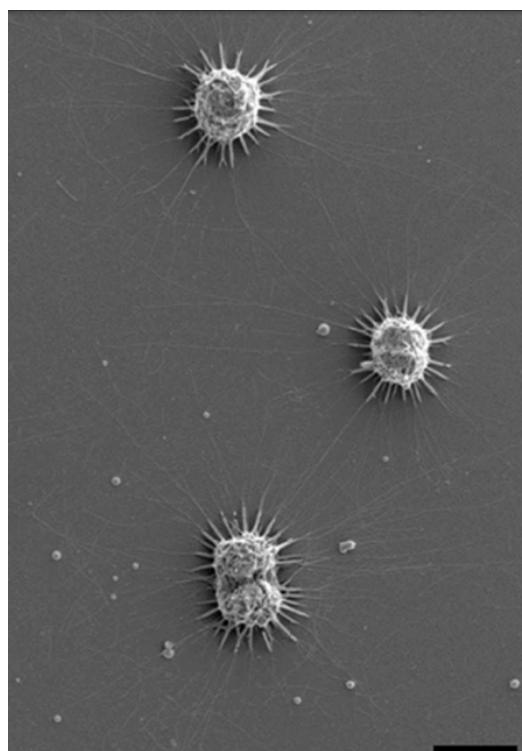
Frontiers Research Topics are very popular trademarks of the Frontiers Journals Series: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area! Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers Editorial Office: researchtopics@frontiersin.org

ARCHAEAL CELL ENVELOPE AND SURFACE STRUCTURES

Topic Editors:

Sonja-Verena Albers, University of Freiburg, Germany

Mecky Pohlschroder, University of Pennsylvania, USA



Single and dividing *Ca. A. hamiconexum* cells, connected by filamentous structures, the hamis. Image by G. Wanner, C. Moissl-Eichinger taken from Perras AK, Daum B, Ziegler C, Takahashi LK, Ahmed M, Wanner G, Klingl A, Leitinger G, Kolb-Lenz D, Gribaldo S, Auerbach A, Mora M, Probst AJ, Bellack A and Moissl-Eichinger C (2015) S-layers at second glance? Altarchaeal grappling hooks (hami) resemble archaeal S-layer proteins in structure and sequence. *Front. Microbiol.* 6:543. doi: 10.3389/fmicb.2015.00543.

Prokaryotes have a complex cell envelope which has several important functions, including providing a barrier that protects the cytoplasm from the environment. Along with its associated proteinaceous structures, it also ensures cell stability, facilitates motility, mediates adherence

to biotic and abiotic surfaces, and facilitates communication with the extracellular environment. Viruses have evolved to take advantage of cell envelope constituents to gain access to the cellular interior as well as for egress from the cell. While many aspects of the biosynthesis and structure of the cell envelope are similar across domains, archaeal cell envelopes have several unique characteristics including, among others, an isoprenoid lipid bilayer, a non-murein-based cell wall, and a unique motility structure, (important features that give archaeal cell envelopes characteristics that are significantly different from those of bacterial cell envelopes – possibly out). Recent analyses have revealed that the cell envelopes of distantly related archaea also display an immense diversity of characteristics. For instance, while many archaea have an S-layer, the subunits of S-layers of various archaeal species, as well as their posttranslational modifications, vary significantly. Moreover, like gram-negative bacteria, recent studies have shown that some archaeal species also have an outer membrane. In this collection of articles, we include contributions that focus on research that has expanded our understanding of the mechanisms underlying the biogenesis and functions of archaeal cell envelopes and their constituent surface structures.

Citation: Albers, S-V., Pohlschroder, M., eds. (2016). Archaeal Cell Envelope and Surface Structures. Lausanne: Frontiers Media. doi: 10.3389/978-2-88919-773-6

Table of Contents

- 06 Editorial: Archaeal Cell Envelope and Surface Structures**
Mechthild Pohlschroder and Sonja-Verena Albers
- 08 The archaeellum: How Archaea swim**
Sonja-Verena Albers and Ken F. Jarrell
- 20 Biofilm formation of mucosa-associated methanarchaeal strains**
Corinna Bang, Claudia Ehlers, Alvaro Orell, Daniela Prasse, Spinner Marlene, Stanislav N. Gorb, Sonja-Verena Albers and Ruth A. Schmitz
- 29 The universal tree of life: An update**
Patrick Forterre
- 47 Archaeal membrane-associated proteases: Insights on *Haloferax volcanii* and other haloarchaea**
María I. Giménez, Micaela Cerletti and Rosana E. De Castro
- 56 Biosynthesis of archaeal membrane ether lipids**
Samta Jain, Antonella Caforio and Arnold J. M. Driessen
- 72 Archaeal S-layer glycoproteins: Post-translational modification in the face of extremes**
Lina Kandiba and Jerry Eichler
- 77 Cytochromes c in Archaea: Distribution, maturation, cell architecture, and the special case of *Ignicoccus hospitalis***
Arnulf Kletzin, Thomas Heimerl, Jennifer Flechsler, Laura van Niftrik, Reinhard Rachel and Andreas Klingl
- 92 S-layer and cytoplasmic membrane – exceptions from the typical archaeal cell wall with a focus on double membranes**
Andreas Klingl
- 98 Mining proteomic data to expose protein modifications in *Methanosarcina mazei* strain Gö1**
Deborah R. Leon, A. Jimmy Ytterberg, Pinmanee Boontheung, Unmi Kim, Joseph A. Loo, Robert P. Gunsalus and Rachel R. Ogorzalek Loo
- 114 Novel pili-like surface structures of *Halobacterium salinarum* strain R1 are crucial for surface adhesion**
Gerald Losensky, Lucia Vidakovic, Andreas Klingl, Felicitas Pfeifer and Sabrina Fröls
- 125 Pyrococcus furiosus flagella: Biochemical and transcriptional analyses identify the newly detected flaB0 gene to encode the major flagellin**
Daniela J. Näther-Schindler, Simone Schopf, Annett Bellack, Reinhard Rachel and Reinhard Wirth

135 *S-layers at second glance? Altarchaeal grappling hooks (hami) resemble archaeal S-layer proteins in structure and sequence*

Alexandra K. Perras, Bertram Daum, Christine Ziegler, Lynelle K. Takahashi, Musahid Ahmed, Gerhard Wanner, Andreas Klingl, Gerd Leitinger, Dagmar Kolb-Lenz, Simonetta Gribaldo, Anna Auerbach, Maximilian Mora, Alexander J. Probst, Annett Bellack and Christine Moissl-Eichinger

147 *Archaeal type IV pili and their involvement in biofilm formation*

Mechthild Pohlschroder and Rianne N. Esquivel

163 *Archaeal viruses at the cell envelope: Entry and egress*

Emmanuelle R. J. Quemin1 and Tessa E. F. Quax

173 *Fluorescence microscopy visualization of halomucin, a secreted 927 kDa protein surrounding Haloquadratum walsbyi cells*

Ralf Zenke, Susanne von Gronau, Henk Bolhuis, Manuela Gruska, Friedhelm Pfeiffer and Dieter Oesterhelt



Editorial: Archaeal Cell Envelope and Surface Structures

Mechthild Pohlschroder^{1*} and Sonja-Verena Albers^{2*}

¹ Department of Biology, University of Pennsylvania, Philadelphia, PA, USA, ² Department of Microbiology, Institute of Biology, University of Freiburg, Freiburg, Germany

Keywords: archaea, membrane, S-layer, surface filaments

The Editorial on the research topic

Archaeal Cell Envelope and Surface Structures

Archaea and Bacteria have complex cell envelopes that play important roles in several vital cellular processes, including serving as a barrier that protects the cytoplasm from the environment. Along with associated proteinaceous structures, cell envelopes also ensure cell stability, promote motility, mediate adherence to biotic and abiotic surfaces, and facilitate communication with the extracellular environment. While some aspects of the biosynthesis and structure of the cell envelope are similar across the three domains of life, archaeal cell envelopes exhibit several unique characteristics. Moreover, recent analyses have revealed that many features of cell envelopes can vary greatly between distantly related archaea. The collection of reviews and original research papers in this focused issue describes research that has significantly expanded our understanding of the mechanisms underlying the biogenesis and functions of archaeal cell envelopes and their constituent surface structures.

Jain et al. provide a comprehensive review of our current knowledge of the unique archaeal cytoplasmic membrane, an isoprenoid lipid bilayer, as well as recently revealed aspects of cytoplasmic membrane biosynthesis that are conserved across the three domains of life. Complementing this review, Andreas Klingl summarizes the diverse structures and functions of archaeal cytoplasmic membranes (Klingl). While most archaeal cells have a single membrane, archaea having an outer membrane, which had been thought to be rare among archaea, have now been identified in a diverse variety of archaeal lineages. One particularly intriguing diderm is the hyperthermophilic archaeon *Ignicoccus hospitalis*, which has an outer cellular membrane that is energized and is able to use the electrochemical gradient across the membrane to synthesize ATP in the periplasmic space. Complementing this work, Kletzin provides an in-depth review of evolutionarily conserved and unique archaeal inner and outer membrane associated cytochromes (Kletzin et al.). The periplasmic space between the membranes of archaeal diderms does not contain a peptidoglycan layer. In fact, while the cytoplasmic membrane is superimposed by an S-layer in many monoderm archaea, it is unclear how diderms, and even some monoderm extremophiles that lack an S-layer, withstand osmotic stress. As noted by Klingl, glycocalyx, lipoglycans, or other protective cell-associated glycoproteins, may take on the functions of a cell wall in some archaea. One such secreted protein, as described by Zenke et al., is the halomucin of *Haloquadratum walsbyi* (Zenke et al.). While *H. walsbyi* does have a cell wall, halomucin, an unusually large protein (9159 aa), is thought to play an important role in protecting these extreme halophiles against desiccation.

Interestingly, *Candidatus Altarchaeum hamiconexum*, an uncultured diderm euryarchaeon, isolated from biofilms contain hami, cell surface proteins with the appearance of grappling hooks that connect cells to each other and to abiotic surfaces. Perra's stunning imaging suggests that

OPEN ACCESS

Edited and reviewed by:

Marc Strous,
University of Calgary, Canada

*Correspondence:

Mechthild Pohlschroder
pohlschr@sas.upenn.edu;
Sonja-Verena Albers
sonja.albers@biologie.uni-freiburg.de

Specialty section:

This article was submitted to
Microbial Physiology and Metabolism,
a section of the journal
Frontiers in Microbiology

Received: 27 October 2015

Accepted: 16 December 2015

Published: 06 January 2016

Citation:

Pohlschroder M and Albers S-V
(2016) Editorial: Archaeal Cell
Envelope and Surface Structures.
Front. Microbiol. 6:1515.
doi: 10.3389/fmicb.2015.01515

these hook-like filaments are connected to the inner membrane, and, surprisingly, are composed of subunits that share homology with S-layer glycoproteins, possibly suggesting a case of divergent evolution (Perras et al.).

Unlike hami, which appear to be limited to a subset of archaea, type IV pili, as pointed out by Pohlschroder and Esquivel as well as Losensky et al. are conserved across the prokaryotic domains, being found in the majority of sequenced archaea, where, as in bacteria, they play key roles in processes necessary for biofilm formation (Losensky et al.; Pohlschroder and Esquivel). Interestingly, as discussed by Albers and Jarrell, as well as Nather-Schindler et al., a type IV pilus-like structure is responsible for swimming motility in archaea.

Many secreted proteins, including the S-layer glycoprotein and pilin-like proteins, are heavily post-translationally modified. The known proteolytic modifications of the proteins of the model haloarchaeon *H. volcanii*, as reviewed here by Gimenez et al., highlight evolutionarily conserved characteristics, as well as the novel aspects, of these haloarchaeal proteases and their substrates. Using the results of proteomic studies, Leon et al. expand upon the existing experimental datasets of mature archaeal N-termini in the methanogen *Methanosarcina mazei* (Leon et al.), providing an invaluable resource for improving *in silico* prediction tools for the characterization of archaeal proteins, in general, but also specific phyla. Kandiba and Eichler review our current knowledge of N-glycosylation in archaea, including descriptions of the pathways the regulatory roles this post-translational modification plays in cellular processes (Kandiba and Eichler).

Considering the unique aspects of the archaeal cell envelope, including not only the protein structures, but their post-translational modifications as well, it is not surprising that archaeal viruses have evolved specific mechanisms to infect and egress from archaeal cells, which are reviewed in this issue by Quemin and Quax.

Understanding the roles that cell surfaces play in archaeal cellular processes can lead to important insights into the types of adaptations that allow some archaea to thrive in extreme environments, including the ability to form biofilms, which many archaea, including mucosa-associated methanogenic archaea, can establish, as described in this issue by Bang et al.. Archaeal cell membranes and S-layer glycoproteins have been used to make liposomes, and hami are also a potentially useful tool for nanobiological applications. Finally, a better understanding of the similarities and differences among the archaea as well as between the archaea and the other two domains will lead to the development of a more accurate phylogeny. In this issue, Forterre takes advantage of the recent profusion of genome studies, along with supporting *in vivo* work, to assemble an improved tree of life (Forterre).

AUTHOR CONTRIBUTIONS

MP and S-VA have both written the text.

ACKNOWLEDGMENTS

The support of the National Science Foundation MCB-1413158 to MP and the ERC starting grant 311523 (Archaeum) to SA are gratefully acknowledged.

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

Copyright © 2016 Pohlschroder and Albers. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



The archaeellum: how Archaea swim

Sonja-Verena Albers^{1,2*} and Ken F. Jarrell^{3*}

¹ Molecular Biology of Archaea, Institute of Biology II-Microbiology, University of Freiburg, Freiburg, Germany

² Molecular Biology of Archaea, Max Planck Institute for Terrestrial Microbiology, Marburg, Germany

³ Department of Biomedical and Molecular Sciences, Queen's University, Kingston, ON, Canada

Edited by:

Mechthild Pohlschroder, University of Pennsylvania, USA

Reviewed by:

Blanca Barquera, Rensselaer Polytechnic Institute, USA

Romé Voulhoux, Aix-Marseille University, France

Lori L. Burrows, McMaster University, Canada

***Correspondence:**

Sonja-Verena Albers, Molecular Biology of Archaea, Institute of Biology II-Microbiology, University of Freiburg, Schaenzlestrasse 1, 79104 Freiburg, Germany
e-mail: sonja.albers@biologie.uni-freiburg.de;

Ken F. Jarrell, Department of Biomedical and Molecular Sciences, Queen's University, 18 Stuart Street, Kingston, ON K7L 3N6, Canada
e-mail: jarrellk@queensu.ca

Recent studies on archaeal motility have shown that the archaeal motility structure is unique in several aspects. Although it fulfills the same swimming function as the bacterial flagellum, it is evolutionarily and structurally related to the type IV pilus. This was the basis for the recent proposal to term the archaeal motility structure the "archaeellum." This review illustrates the key findings that led to the realization that the archaeellum was a novel motility structure and presents the current knowledge about the structural composition, mechanism of assembly and regulation, and the posttranslational modifications of archaella.

Keywords: archaeal flagellum, archaeellum, motility, type IV pili, motor complex

THE ROAD FROM ARCHAEOAL FLAGELLUM TO THE ARCHAELLUM

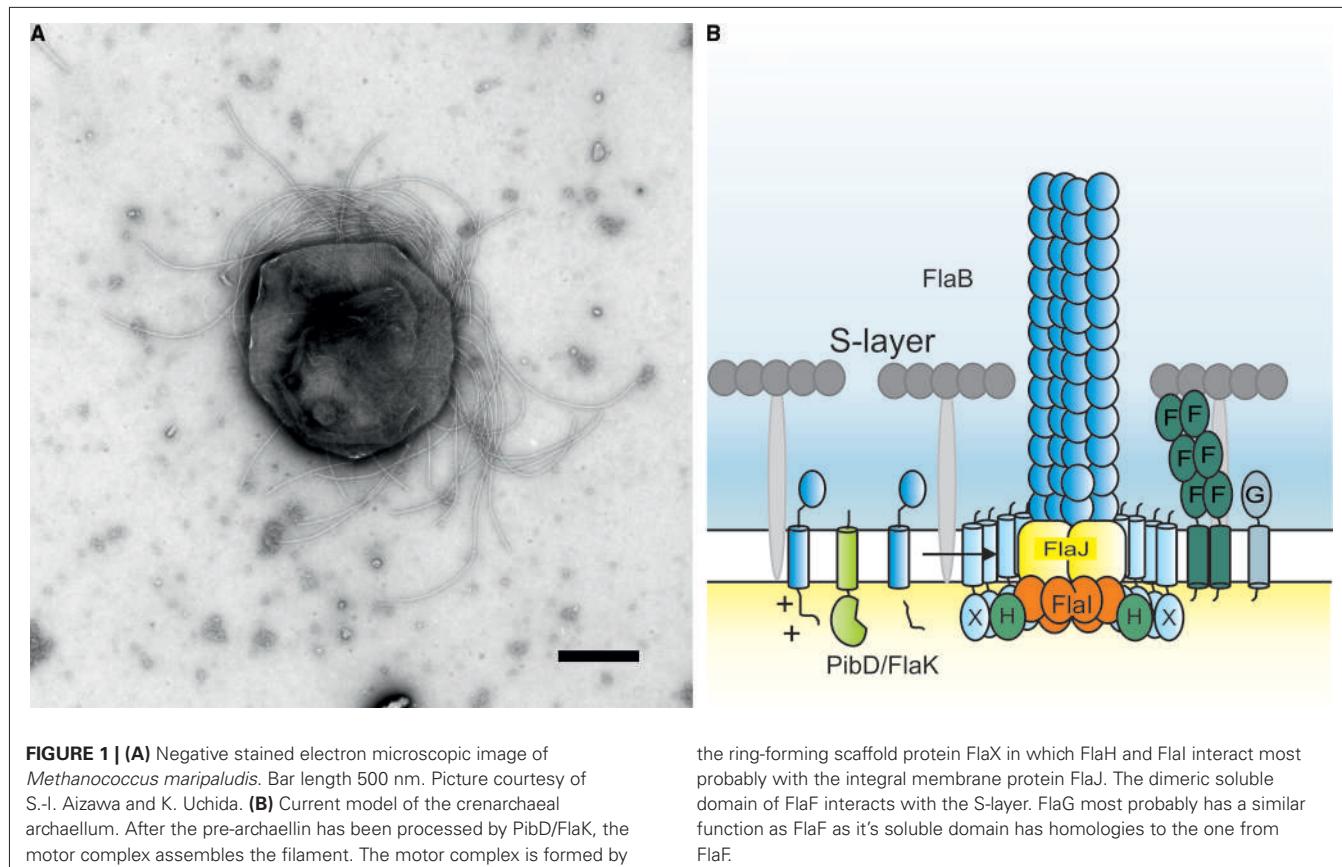
Motility is a trait that is widespread amongst all the different subgroupings of Archaea. While motile archaeal cells possess surface appendages involved in motility that superficially resemble bacterial flagella (**Figure 1A**), biochemical, genetic, and structural analyses of these archaeal appendages in several model organisms have demonstrated the uniqueness of the archaeal motility structure. This review provides an historical account of the investigations on the archaeal motility structure ending with current studies on the regulation of archaella flagella biosynthesis and determination of the roles of some of the specific components in assembly and function of the organelle.

EARLY WORK REVEALED UNUSUAL TRAITS OF ARCHAEOAL FLAGELLA

The first archaeon to have its flagella studied in detail was *Halobacterium salinarum* (*halobium*). Studies by Alam and Oesterhelt (1984) initially revealed several unusual features of the halobacterial flagella. Unlike most bacterial flagella, the flagella of *H. salinarum* form a right-handed helix. Using tethered cells, they showed that these flagella rotate and that the direction of rotation can change from clockwise to counter clockwise (Alam and Oesterhelt, 1984; Marwan et al., 1991). Cells swim forward

when the flagellar rotation is clockwise but backward when rotation is counter clockwise. Unlike peritrichously flagellated bacteria, the flagella bundle of *H. salinarum* did not fly apart when rotation direction changed. Flagella were isolated from a "super" flagella overproducer called strain M-175, a strain that shed large numbers of unattached flagella which aggregated into thick bundles containing 100s of individual flagellar filaments. Analysis of these flagella by SDS-PAGE revealed three bands with centers of intensity that corresponded to molecular masses of 26, 30, and 36 kDa, although each of these bands actually consisted of multiple bands in a ladder-like appearance indicating heterogeneity.

This striking pattern revealed by SDS-PAGE was recognized by Wieland et al. (1985) as almost identical to a pattern of heterogeneous sulfated proteins previously studied and thought to be related to bacteriopsin. Their work showed that the flagellin bands reported by Alam and Oesterhelt (1984) were indeed the same as the sulfated proteins. Further study revealed that the flagellins were modified with an N-linked oligosaccharide common to the S layer glycoprotein, the first prokaryotic glycoprotein identified. The N-linked glycan was determined to be Asn-Glc1-4GlcA1-4GlcA1-GlcA and Asn-Glc1-4GlcA1-4GlcA1-4Glc. They studied both the wildtype *H. salinarum* strain and also the superflagella producing M-175 strain and determined that while the pattern was similar in both cases, the entire set of bands was shifted



to lower apparent molecular masses in the M-175 strain. It was proposed that the M-175 strain had lost one or more glycosylation sites. Experimental investigation of this proposal was apparently never pursued but subsequent work identifying five flagellin genes (Gerl and Sumper, 1988) makes this explanation unlikely since a loss of a glycosylation site would presumably have to occur in all five flagellins to recreate the observed pattern. It seems more likely that the M-175 strain had a mutation in one of the N-glycan assembly or biosynthesis steps that rendered all the N-glycan-modified proteins migrate as smaller protein on SDS-PAGE. This type of effect was subsequently observed in other archaea like *Methanococcus* species (Chaban et al., 2006; VanDyke et al., 2009), *Haloferax volcanii* (*Hfx. volcanii*; Tripepi et al., 2012), and *Sulfolobus acidocaldarius* (Meyer et al., 2011). Nonetheless, in a prescient hypothesis, Wieland et al. (1985) thought that the overproduction of superflagella by the M-175 mutant could occur if correct glycosylation of the flagellins is necessary for proper incorporation of the flagella into the cell envelope. These were the first prokaryotic flagellins shown to be glycoproteins.

A further key finding was that N-glycosylation in *H. salinarum* occurred on the external surface of the cytoplasmic membrane (Sumper, 1987). This was shown by the addition of ethylenediaminetetraacetic acid (EDTA) which caused a shift in the flagellin molecular masses to the same values as occurs if the flagellins were chemically deglycosylated. In addition, it was shown that an

exogenously added peptide carrying an N-glycosylation sequon could be glycosylated even though it could not cross the cytoplasmic membrane. This extracellular site of glycosylation of the flagellins led Gerl and Sumper (1988) to state that “aggregation to a functional flagellum is likely to occur by a mechanism different from that proposed for the assembly of eubacterial flagella.”

Sumper’s group followed up the glycobiology aspect of the halobacterial flagella with genetic studies. Remarkably, they discovered that *H. salinarum* had five flagellin genes located at two distinct loci in the genome: two genes (*flgA1* and *flgA2*) were located in tandem at one locus while three others (*flgB1*, *flgB2*, and *flgB3*) were found tandemly at a second locus (Figure 2; Gerl and Sumper, 1988). All five flagellin proteins were 193–196 amino acids in length and were remarkably similar in amino acid sequence with large stretches being identical, although there were three short regions of hypervariability that were unique to each flagellin. The calculated molecular masses for all five flagellins were about 20.5 kDa, much smaller than the masses calculated by SDS-PAGE. However, three potential N-linked glycosylation sites were present in each protein. Since the flagellins were already known to be sulfated glycoproteins (Wieland et al., 1985), the heterogeneity seen on SDS-PAGE was explained by the presence of five different proteins which perhaps had different degrees of glycosylation. At the time, a search of protein databanks revealed no significant similarity to other sequences. Critically, the N-terminus of the 26 kDa band was resistant to Edman degradation.

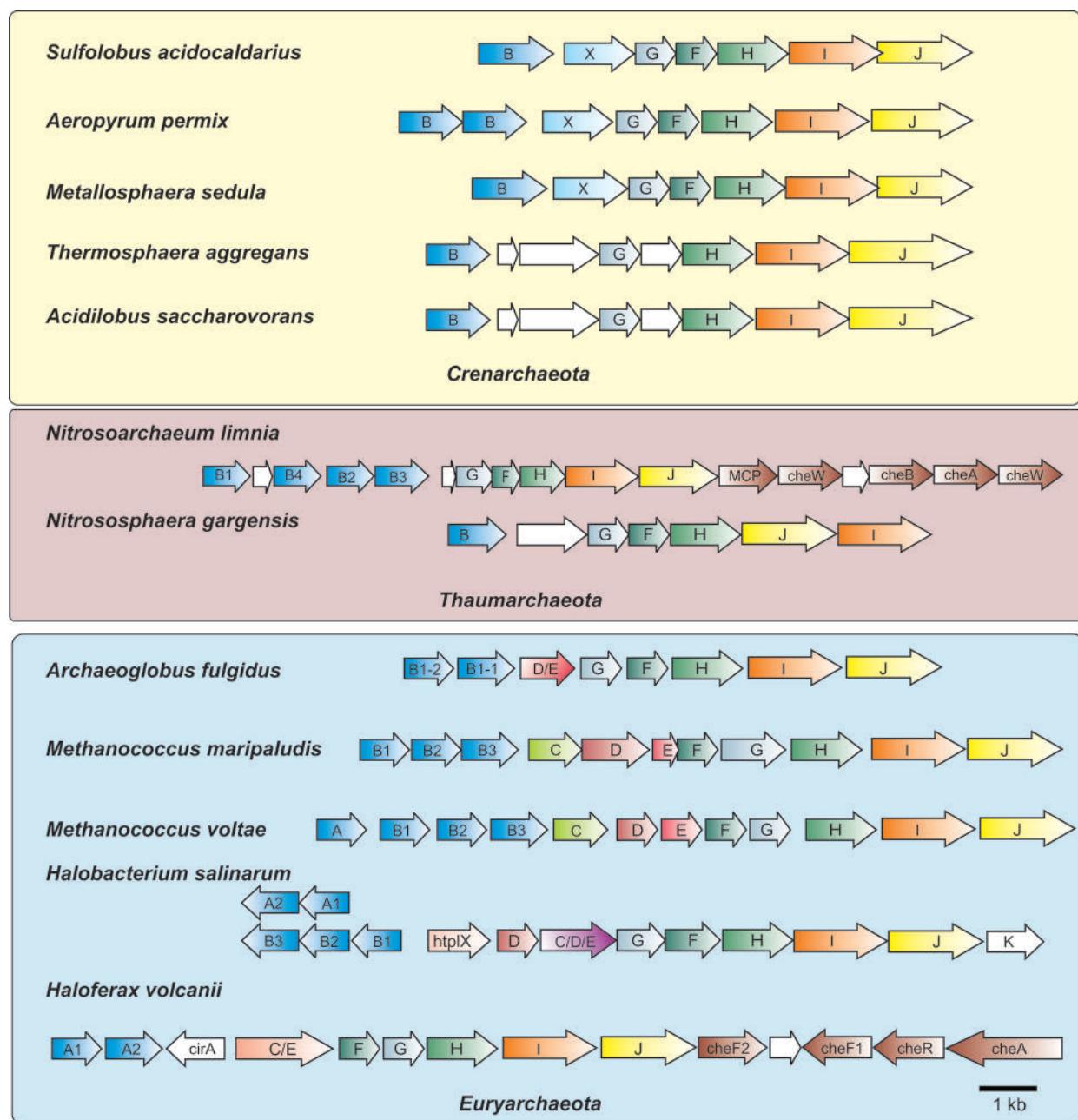


FIGURE 2 | Organization of archaella operons. Archaella operons of three of the archaeal kingdoms Crenarchaeota, Thaumarchaeota and Euryarchaeota are depicted. The *fla* genes are abbreviated using the respective letter of the *fla* gene. Homologous genes are shown in the same color. Genes of unknown

function are depicted in white. In the strain where chemotaxis genes are adjacent to the archaellum operon they are partially depicted. MCP, methyl accepting chemotaxis protein; *che* genes, genes encoding parts of the chemosensory system; *htpIX*, methyl accepting transducer.

A follow-up study (Gerl et al., 1989) demonstrated that all five of the flagellin proteins could be identified in purified flagella due to the unique amino acid sequences in the variable regions. Such methodology revealed that the flagellins in the 26 kDa band were FlgA2, FlgB1, and FlgB3 while only FlgA1 was found in the 30 kDa band and FlgB2 was the sole flagellin found in the 36 kDa band. Western blotting with specific antibody raised to amino acid sequences unique to the different flagellins also

revealed that FlgA1 antisera only reacted to the 30 kDa band and the FlgA2-specific antibodies only reacted to the 26 kDa band.

DISCOVERY OF SIGNAL PEPTIDES ON ARCHAEOAL FLAGELLINS

Flagella were subsequently purified from a number of archaea and the N-terminal amino acid sequence was obtained for a

number of these proteins, including one flagellin band from *Methanococcus voltae* (Kalmokoff et al., 1990; Faguy et al., 1994a, 1996). Remarkably, these N-terminal sequences showed no similarity to any bacterial flagellins but all the archaeal sequences showed high amino acid sequence similarity among themselves. Intriguing, the N-terminal sequences obtained aligned with the sequence predicted for the *H. salinarum* flagellin gene sequences but beginning at amino acid position 13, suggesting that the archaeal flagellins were made as preproteins with a signal peptide (Kalmokoff et al., 1990). Shortly thereafter, the flagellin genes of *M. voltae* were cloned and sequence analysis revealed that, indeed, all four flagellin genes of this organism encoded proteins with predicted short signal peptides (Kalmokoff and Jarrell, 1991). This was an unexpected finding since flagellins in bacteria are not made as preproteins and reach their final destination via a flagellum-specific type III secretion system located at the base of the flagellum (Macnab, 2004; Chevance and Hughes, 2008). The flagellins pass through the hollow organelle to the distal tip before incorporation under the flagellar cap protein. Thus, in addition to the unusual structural features reported by Alam and Oesterhelt (1984), archaeal flagella possessed two unique characteristics not found in bacterial flagella: its component subunits were made initially with signal peptides and they were modified with N-linked glycans (Wieland et al., 1985; Kalmokoff and Jarrell, 1991). These two properties suggested a completely novel assembly model was used in archaea for flagella biosynthesis.

SEQUENCE SIMILARITY OF ARCHAEL FLAGELLINS TO TYPE IV PILINS AND A NEW MODEL FOR FLAGELLA ASSEMBLY

While initial attempts did not find any relatives of archaeal flagellins in gene databases, Faguy et al. (1994b) reported that the N-terminal region of archaeal flagellins shared sequence similarity to the same highly conserved region in type IV pilins, which themselves formed a different type of appendage on the bacterial cell surface distinct from flagella (Pelicic, 2008; Burrows, 2012). Type IV pilins are known to be made initially as preproteins with unusual signal peptides. The signal peptide is cleaved at a conserved site by a dedicated signal peptidase, termed a prepilin peptidase or signal peptidase III, that is distinct from both signal peptidase I and II (Strom et al., 1994; Lory and Strom, 1997; Giltner et al., 2012). This noted similarity to type IV pilins led to the hypothesis that archaeal flagella could be assembled in a completely novel way compared to bacterial flagella, with insertion of new subunits at the base (Faguy et al., 1994b; Jarrell et al., 1996a). Following the development of the first genetic and transformation systems in *M. voltae* (Gernhardt et al., 1990; Patel et al., 1994), the flagellin genes of this methanogen were targeted and interrupted (Jarrell et al., 1996b). Mutants in the flagellin *flaB2* so generated were non-flagellated, thus linking these genes with the appearance of the flagella on the cell surface for the first time.

SIMILARITIES OF ARCHAEL FLAGELLA AND TYPE IV PILI: FURTHER STRUCTURAL AND GENETIC EVIDENCE

Evidence from several avenues of research supporting the notion that the archaeal flagella were distinct from bacterial flagella continued to appear. Electron microscopic examination of purified archaeal flagella revealed a knob at the cell proximal end but no

distinct ring structure as seen in flagella of both Gram negative and Gram positive flagella (Kalmokoff et al., 1988; Kupper et al., 1994). Curved hooks regions were observed in some archaeal flagella and specific flagellins were shown to be responsible for this region in both *Methanococcus* and *Halobacterium* (Bardy et al., 2002; Beznosov et al., 2007; Chaban et al., 2007), but this finding was not universal. For example, no hook region has been observed in *Sulfolobus solfataricus*, an archaeon possessing a single flagellin gene (Szabo et al., 2007b). Since most sequenced crenarchaeota genomes only possess a single flagellin gene, the flagella of these organisms would also be expected to lack a hook. Rotation of flagella in *H. salinarum* was shown to be ATP-dependent and not proton motive force (or sodium motive force) driven as it is in bacterial flagella (Streif et al., 2008). Structural studies by the Trachtenberg group revealed further crucial findings. The reconstructed 3D structure of flagella from distantly related archaea (*H. salinarum* and *Sulfolobus shibatae*) was shown to share common features with type IV pili and be distinct from known bacterial flagella structures (Cohen-Krausz and Trachtenberg, 2002, 2008; Trachtenberg and Cohen-Krausz, 2006). Critically, and in support of the type IV pili assembly model proposed earlier by Jarrell et al. (1996a), was the absence of a lumen in the interior of the archaeal flagella that could allow passage of subunits to the distal tip as occurs in bacterial flagella. This seemingly eliminated any potential chance for distal growth of archaeal flagella.

Meanwhile, further genetic evidence emerged that supported the evolutionary relationship of archaeal flagella to type IV pili. Sequencing of genes located downstream of the flagellin genes revealed the presence of two genes that encoded homologues to key components of the type IV pili assembly system, namely a PilB-like polymerizing ATPase (termed FlaI) and the conserved membrane/platform protein (FlaJ; Bayley and Jarrell, 1998; Peabody et al., 2003). Deletion of these genes in various archaea confirmed their involvement in the archaeal flagella system, since these mutants were consistently non-flagellated (Patenge et al., 2001; Thomas et al., 2001b; Chaban et al., 2007; Lassak et al., 2012b). With the advent of the genomic age, many sequenced archaeal genomes were examined and no genes encoding proteins involved in bacterial flagella structure (i.e., rod, hook, rings, etc) were identified (Faguy and Jarrell, 1999; Nutsch et al., 2005; Pyatibratov et al., 2008). Such analyses, as well as directed genetic studies in several archaea, revealed that a conserved group of so-called *fla* accessory genes, often *flaC–flaJ* in euryarchaeotes, was found usually directly downstream of, and co-transcribed with, flagellin genes (in some cases *fla* accessory genes are located in the immediate vicinity but in an opposite orientation to the flagellin genes; see Figure 2; Nagahisa et al., 1999; Patenge et al., 2001; Thomas and Jarrell, 2001; Ng et al., 2006). A typically smaller subset of these genes was observed in the genomes of crenarchaeotes (Ng et al., 2006; Lassak et al., 2012a).

PROPOSAL TO RENAME THE ARCHAEL FLAGELLUM AS THE ARCHAELLUM

By 2012, the evidence was overwhelming that there were two distinct flagella structures in the prokaryotic world: the bacterial one and the archaeal one. They were not evolutionarily related

and the Archaea domain structure was, in fact, closely related to type IV pili and the homologous type II secretion system which involves a piston-like pseudopilus comprised of pseudopilins and used to push exported proteins through the outer membrane of Gram negative bacteria (Peabody et al., 2003; Korotkov et al., 2012). The sole similarity of the bacterial and archaeal flagella was seemingly in their function as a rotating swimming organelle. With the realization that archaeal flagella were in fact a rotating variant of type IV pili with no evolutionary relationship to bacterial flagella, we proposed that this prokaryotic motility structure be designated the archaeellum (Jarrell and Albers, 2012), a distinct name that nevertheless fuses the concept of Archaea and flagellum and thus readily allows for similar terms common in the bacterial flagella field to be used in archaea (i.e., archaella/flagella, archaeellins/flagellins, archaeallated cells/flagellated cells). This proposal has met with both criticism and support and its acceptance is still under debate in the scientific community (Eichler, 2012; Wirth, 2012), but its use is becoming more common both within the archaeal research community (Stieglmeier et al., 2014; Syutkin et al., 2014) as well as outside the archaeal field (Giltner et al., 2012; Campos et al., 2013). What is undeniable is that each of the three domains of life, Eukarya, Bacteria, and Archaea has entirely distinct “flagella.”

KEY ENZYME IN ARCHAELLIN PROCESSING: THE PREPILIN PEPTIDASE-LIKE FlaK/PibD

Study of the archaeallin signal peptide processing led to the implementation of an assay based on type IV pilin processing to show *in vitro* processing of archaeallins that had been heterologously expressed in *Escherichia coli* (Bayley and Jarrell, 1999; Correia and Jarrell, 2000). Shortly thereafter, the gene encoding the prepilin peptidase-like enzyme (FlaK), responsible for processing of the prearchaellins, was identified in both *M. maripaludis* and *M. voltae* and its critical role demonstrated in archaella biosynthesis when deletion of the gene resulted in non-archaellated cells (Bardy and Jarrell, 2002, 2003). Shortly thereafter, a prepilin peptidase-like enzyme, designated PibD, was identified first in *S. solfataricus* and then other archaea that was much broader in its substrate specificity and capable of processing all type IV prepilin-like proteins including archaeallins, pilins, and sugar binding proteins (Albers et al., 2003; Tripepi et al., 2010; Henche et al., 2014).

The archaeal prepilin peptidases FlaK/PibD have both been demonstrated by site-directed mutagenesis studies to belong to the unusual family of aspartic acid proteases that also includes the prepilin peptidases of type IV pili systems in bacteria and presenilin, a protease involved in processing amyloid precursor proteins in humans (LaPointe and Taylor, 2000; Bardy and Jarrell, 2003; Szabo et al., 2006; Ng et al., 2007; Hu et al., 2011; Henche et al., 2014). Unlike the case with prepilin peptidases which methylate the N-terminal amino acid of the processed mature pilins (typically, but not always, a phenylalanine; Strom et al., 1993), the archaeal enzymes have not been shown to possess methyltransferase activity. In these polytopic membrane enzymes, two aspartic acid residues, one located within a conserved classic GxGD motif or a new variant GxHyD [Hy represents a hydrophobic amino acid, most commonly alanine, found in about 60% of

archaeal sequenced genomes (Hence et al., 2014)], are critical for the peptidase activity (LaPointe and Taylor, 2000; Bardy and Jarrell, 2003; Szabo et al., 2006; Hu et al., 2011). Recently, the crystal structure of the *M. maripaludis* FlaK was obtained (see Figure 3A; Hu et al., 2011). Analysis of the structure confirmed the presence of six transmembrane helices and demonstrated that FlaK must undergo a conformational change in order to bring the two critical aspartic acid residues, located in transmembrane helix 1 and 4 (the GXGD motif), into close proximity for catalysis.

The typical length of the processed part of the signal peptide on archaeallins is 6–12 amino acids (Ng et al., 2006), the short length typical of type IVa prepilins of bacteria (Giltner et al., 2012). In conjunction with studies that investigated the important amino acids in the signal peptidases necessary for catalysis, site-directed mutagenesis studies were also conducted to investigate the importance of various amino acid positions in the signal peptide of archaeallins themselves. In the archaeallins of *M. voltae*, the highly conserved glycine at the –1 position (position is relative to the cleavage site) was shown to be critical for peptidase cleavage, while the basic amino acids usually found at positions –2 and –3 as well as the conserved +3 glycine also were found to play important roles (Thomas et al., 2001a). Similar studies conducted on the glucose binding protein precursor, used as a model substrate for PibD activity in *S. solfataricus*, indicated PibD was more flexible in accepting amino acid substitutions around the cleavage site than was FlaK, as expected from its broader substrate range (Albers et al., 2003). In *M. maripaludis*, FlaK specifically processes pre-archaellins while the type IV pre-pilins are processed by another type IV prepilin-like peptidase, EppA (Szabo et al., 2007a). *S. solfataricus* PibD can also process the archaeallins of *M. voltae* (Ng et al., 2009). In that report, PibD was shown to cleave archaeallins engineered with signal peptides as short as 3 and 4 amino acids while for FlaK a minimal signal peptide length of five amino acids was needed for cleavage. This further supports the more flexible nature of the PibD enzyme. Recently, the prepilin peptidase in *Hfx. volcanii*, also designated PibD, was found to be responsible for the processing of both archaeallin FlgA2 and other type IV pilin proteins (Tripepi et al., 2010; Esquivel et al., 2013).

A PERL program termed FlaFind, using abundant archaeallin sequences available from complete genome sequencing projects as a training set, was developed to predict type IV pilin-like proteins in Archaea based on identification of signal peptides that were similar to those found in archaeallins that were known to be processed by archaeal prepilin peptidase-like enzymes (Szabo et al., 2007a). As more experimental evidence accumulated on the actual sequences processed by archaeal signal peptidase III enzymes, a newer version of FlaFind, FlaFind 1.2 (<http://signalfind.org/flafind.html>), was introduced that allowed for the presence of glutamate and aspartate at the –2 position. The program searches for the conserved signal peptide motif [KRDE][GA][ALIFQMVED][ILMVTAS] (Figure 3B; Esquivel et al., 2013).

BIOCHEMICAL AND STRUCTURAL ANALYSES OF ARCHAELLUM SUBUNITS

In all archaella operons, the genes *flaF,G,H,I*, and *J* are conserved and considered to encode the proteins that form the general

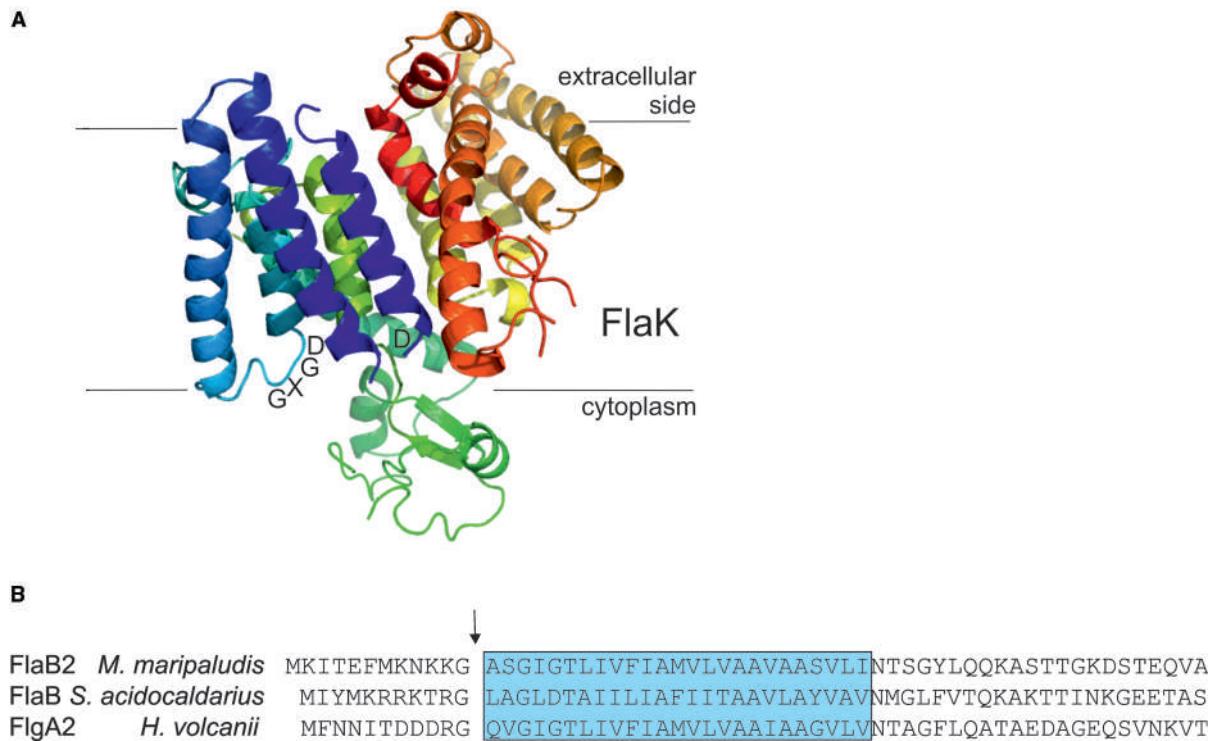


FIGURE 3 | The archaeal type IV prepilin peptidases. (A) Crystal structure of FlaK from *M. maripaludis*. Model was constructed with Pymol (www.pymol.org) using the pdb file 3SOX. The GXDG motif is indicated as well as the localization of the membrane. It is clear that the active site is directed toward the cytoplasmic side of the membrane where cleavage of the class III signal peptide occurs. **(B)** The N-terminal archaellin sequences

of *M. maripaludis*, *S. acidocaldarius*, and *Hfx. volcanii* are shown. The arrow indicates the cleavage site and the blue box delineates the hydrophobic, membrane inserted part of the mature N-terminus of the proteins. FlaK is specific for the archaellins in *M. maripaludis*, whereas PibD processes a variety of class III signal peptide containing substrates in various archaea.

assembly machinery and motor complex of this structure (see **Figures 1B and 2**). All of these genes are essential for archaella assembly and rotation (Patenge et al., 2001; Thomas et al., 2001b; Chaban et al., 2007; Lassak et al., 2012b; Reindl et al., 2013). FlaI was demonstrated to have ATP hydrolyzing activity, which was greatly stimulated by the addition of archaeal lipids (Albers and Driessens, 2005; Ghosh et al., 2011). FlaI forms an ATP-dependent hexamer and was crystallized in different nucleotide-bound states (Ghosh et al., 2011; Reindl et al., 2013). The C-terminal domain (CTD) of FlaI, which contains the Walker A and B motif for ATP-binding and hydrolysis, interacts more strongly with the N-terminal domain (NTD) of the neighboring monomer than with its own NTD. It is hypothesized that this strong interaction is essential for the function of FlaI in the rotation of the archaeellum filament.

In the FlaI hexamer, the N-termini of each monomer form the tips of the crown-like complex. In contrast to the nucleotide-free FlaI hexamer, the tips of the crown were rotated in a perpendicular fashion inside the hexamer in the nucleotide bound state. It is proposed that the tips of FlaI lock into the cytoplasmic loops of FlaJ, the only polytopic membrane protein of the archaeellum machinery, and thereby form a rigid motor complex to drive rotation of the archaeellum filament (Reindl et al., 2013).

Another subunit of the archaeellum, which was biochemically and structurally analyzed, is FlaX. While FlaX is essential for archaeellation in *S. acidocaldarius* (Lassak et al., 2012b), it is not found in euryarchaeotes. FlaX is a monotopic membrane protein and its soluble domain was shown to form large oligomeric ring structures of around 30 nm diameter (Banerjee et al., 2012). It was shown that the coiled-coil region that is present in the middle of its soluble domain is essential for FlaX ring formation. Both parts of FlaI, the N- and the C-terminus, were shown to interact with the soluble part of FlaX (Banerjee et al., 2013).

In addition to FlaI, FlaH is the only other predicted cytoplasmic component of the archaeellum assembly machinery. Although FlaH exhibits a Walker A motif, its non-canonical Walker B motif suggests that FlaH is not an active ATPase. It is proposed that it might modulate the activity of FlaI. A structure of *Pyrococcus horikoshii* FlaH (PH0284) is present in the Protein Data Bank, but has not yet been described. It shows high similarity to RecA folds, but no nucleotide was present in the structure. Using different biochemical assays, it was demonstrated that FlaX, FlaI, and FlaH indeed form a stable complex (Banerjee et al., 2013), which is thought to anchor the cytoplasmic part of the motor complex of the archaeellum. The binding affinities of the single subunits to each other were all in the nanomolar range.

FlaF and FlaG are also conserved components of the archaellum assembly machine. Their order in archaella operons, however, clearly sets euryarchaea apart from crenarchaea (Desmond et al., 2007). Both FlaF and FlaG are monotopic membrane proteins. FlaF contains a partial predicted archaellin domain implying that its soluble domain might be located in the pseudo-periplasm. Very recently, the crystal structure of FlaF from *S. acidocaldarius* was solved (Banerjee et al., submitted). It revealed a β -sheet-dominated structure with homologies to immunoglobulin folds and the recently solved structure of SbsB, the S-layer protein of *Geobacillus stearothermophilus* (Baranova et al., 2012). Binding assays with isolated *S. acidocaldarius* S-layer showed that FlaF bound to the S-layer, implying that it might be involved in anchoring the archaellum in the archaeal cell envelope (Banerjee et al., submitted). It was shown that dimerization is important for FlaF's function and therefore it is proposed that FlaF forms a channel between the cytoplasmic membrane and the S-layer in which the archaellum filament can cross the pseudo-periplasmic space and the S-layer. A current model of the crenarchaeal archaellum is depicted in **Figure 1B**.

THE ARCHAELLUM IS A ROTATING TYPE IV PILUS

When Alam and Oesterhelt (1984) showed that the archaella of *H. salinarum* were rotating, this was, at first sight, not surprising as they were being compared to bacterial flagella which were known to behave similarly. Later, *Sulfolobus* cells were also observed to rotate when tethered to a surface (Grogan, 1989). However, in light of the archaeal motility structure subsequently being identified as a type IV pilus structure, this rotation feature became exciting again. Type IVa pili are known to be extended and retracted by the action of two ATPases, PilB, and PilT, respectively (Merz et al., 2000). This feature enables bacteria to move across surfaces in a process termed twitching. The bacterium is pulled over a surface when extended pili adhere to the surface and subsequently retract (Burrows, 2005). However, type IV pili have not been reported to rotate, although a model was recently proposed in which the pseudopilus of a type II secretion system rotates during its assembly (Nivaskumar et al., 2014). While *H. salinarum* was shown to be able to switch the rotation direction of its archaella depending on different light pulses (Alam and Oesterhelt, 1984), the switching of the *Sulfolobus* archaellum seems to be a stochastic event (Shahapure et al., 2014). 72% of tethered *S. acidocaldarius* cells were found to be rotating counterclockwise, whereas 10% were switching spontaneously, and 18% of the cells were spinning clockwise. The archaellum switching events in *H. salinarum* are governed by the action of the chemotaxis/phototaxis system which has been studied in detail in this organism (Marwan et al., 1990; Rudolph et al., 1995, 1996; Rudolph and Oesterhelt, 1996; Schlesner et al., 2012). Many facets of the chemotaxis systems of bacteria and archaea seem to be conserved (Szurmant and Ordal, 2004) but it remains to be elucidated how the chemotaxis system can enact switching events in two absolutely different motility structures, the flagellum and the archaellum. For the archaellum, data on this topic is extremely limited. Schlesner et al. (2009) identified three proteins in *H. salinarum* that interacted with both chemotaxis proteins and the archaella proteins FlaCE and FlaD. Two of these

proteins belong to protein family DUF439 while the third is a HEAT_PBS family protein. Deletion of one of the DUF439 proteins or the HEAT_PBS family protein led to cells that could not switch the direction of archaella rotation. These proteins provide a link between the signal transduction of the chemotaxis system and the archaella.

KEY ROLE FOR N-LINKED GLYCOSYLATION IN ARCHAEL FLAGELLA ASSEMBLY AND FUNCTION

With the availability of complete genome sequences for many archaeallated archaea and the development of genetic techniques for generating targeted gene deletions, advances were made in the analysis and importance of the N-linked glycosylation found on the archaellins in model archaea (Gernhardt et al., 1990; Tumbula et al., 1994; Moore and Leigh, 2005; Leigh et al., 2011; Wagner et al., 2012; Jarrell et al., 2014). This work was originally performed on *Methanococcus* species where either a trisaccharide (*M. voltae*; Voisin et al., 2005) or tetrasaccharide (*M. maripaludis*; Kelly et al., 2009) glycan was found linked to each of the multiple archaellins that comprise the archaellum filament. It was quickly observed that deletion of *aglB* (the oligosaccharyltransferase responsible for transfer of the completed glycan from its dolichol lipid carrier onto the target protein) resulted in non-archaeallated cells, suggesting that the archaellins must undergo the N-glycosylation modification to be properly incorporated into a filament on the cell surface [Chaban et al., 2006; VanDyke et al., 2009; as considered earlier for *Halobacterium* M175 (Wieland et al., 1985)]. Further studies demonstrated that mutants carrying deletions in other *agl* (archaeal glycosylation) genes involved in either biosynthesis of the individual sugars of the glycan or its assembly on the lipid carrier (various glycosyltransferases) also led to defects in either archaellum assembly or motility (VanDyke et al., 2008, 2009; Chaban et al., 2009; Jones et al., 2012). In the case of both *Methanococcus* species, synthesis of a glycan of at least two sugars was necessary in order for cells to be archaeallated. In the case of *M. maripaludis*, motility was correlated directly with the size of the glycan with wildtype cells carrying the tetrasaccharide glycan being more motile than cells carrying archaellins with a trisaccharide glycan which in turn were more motile than cells carrying archaellins modified with a disaccharide (VanDyke et al., 2009). Similar observations were also reported in both *S. acidocaldarius* and *Hfx. volcanii* where studies on N-linked glycosylation, initially focused on effects on the S-layer protein (Eichler, 2013; Kaminski et al., 2013; Meyer and Albers, 2013), turned also to an examination of this posttranslational modification on surface appendages. Again interference in the N-glycosylation pathway had major effects on archaeallation and motility. In *Hfx. volcanii*, where archaellins are decorated with a pentasaccharide, mutants deleted for *aglB* were non-archaeallated (Tripepi et al., 2012). Investigations with strains deleted for other *agl* genes indicated that likely a minimum three sugar glycan was necessary for proper archaella formation and/or function. Site directed mutagenesis to remove each of the three N-glycosylation sites of archaellin FlgA indicated that modification at all sites was necessary for archaella formation. In *S. acidocaldarius*, recent evidence also showed that interference in the N-glycosylation system also led to non-archaeallated cells (Meyer et al., 2011, 2013).

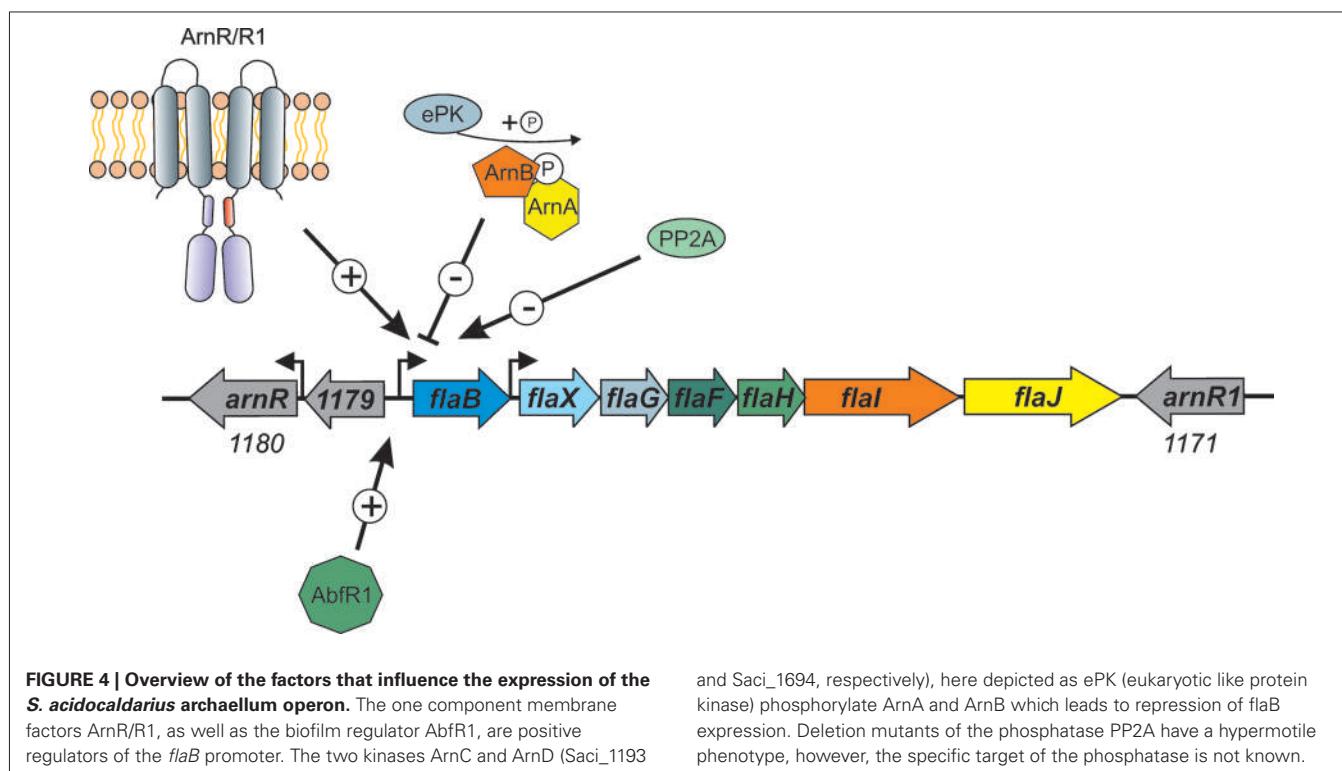
However, it could be demonstrated in this organism that it was not the glycosylation of the archaeallin itself that is important for archaella stability, but rather the N-glycosylation pathway is probably essential for archaella assembly. Deletion of five of the six N-glycosylation sites of the lone archaeallin led to no decrease in motility, whereas the deletion of genes of the N-glycosylation pathway did. Therefore, it was proposed that the correct N-glycosylation of cell wall components plays an important role in archaella assembly (Meyer et al., 2014). Interestingly, in *M. maripaludis*, elimination of the four N-glycosylation sites in all possible combinations in one of the major archaeallins, FlaB2, indicated that archaella could be assembled and function if FlaB2 was missing three of the four sites but not all of them (Ding et al., in press). Thus, it seems, that depending on which model organisms is being studied, N-glycosylation of the archaeallins may be necessary at all N-glycosylation sites (*Hfx. volcanii*), at none of the sites (*S. acidocaldarius*) or at some of the sites (*M. maripaludis*) for archaella assembly.

REGULATION OF ARCHAELLA COMPONENT EXPRESSION

The regulation of the archaeallum operon is, so far, restricted to a few examples. In studied methanogens, biosynthesis of archaella is not constitutive: it is known in both *Methanocaldooccus janaschii* and *M. maripaludis*, for example, that archaella synthesis is induced under H₂ limitation conditions (Mukhopadhyay et al., 2000; Hendrickson et al., 2008). Quantitative proteomics of nutrient-limited *M. maripaludis* further demonstrated that the expression of archaeallins was affected by multiple nutritional factors: decreased expression was observed under nitrogen limitation but increased expression when cells were phosphate limited (Xia

et al., 2009). To date, no transcriptional regulators involved in archaeallation have been identified in any euryarchaeon.

However, it is in the crenarchaea that most of the information concerning regulation of archaella is known. It was demonstrated in *S. solfataricus* that starvation induced the expression of the archaeallum operon (Szabo et al., 2007b). In *S. acidocaldarius*, a number of components of the archaeallum regulatory network (termed Arn proteins) were identified. ArnA, containing a fork head associated (FHA) domain and a zinc finger domain, was first shown in *S. tokodaii* (Wang et al., 2010) to be phosphorylated by kinase ST1565. A screen with *S. tokodaii* promoters identified the *flaX* promoter as a target, which was only bound when ArnA was in the phosphorylated state (Duan and He, 2011). ArnA is co-transcribed in an operon with ArnB, which contains a van Willebrand domain. These two proteins were demonstrated to strongly interact with each other both *in vitro* and *in vivo* in *S. acidocaldarius* (Reimann et al., 2012). As FHA domain containing proteins are known to bind to phosphorylated tyrosines, it is proposed that the ArnA and ArnB interaction relies on protein phosphorylation. Deletion of ArnA, ArnB or the zinc finger of ArnA led to the overexpression of archaella in *S. acidocaldarius* even without starvation conditions, indicating that both proteins act as repressors of the archaeallum operon (see Figure 4; Reimann et al., 2012). In the *fla* operons of Sulfolobales, three other conserved proteins were identified, Saci_1180 (ArnR), Saci_1171 (ArnR1) and Saci_1179. Saci_1179 is a small membrane protein; deletion of the corresponding gene did not lead to any deregulation of archaella in *S. acidocaldarius* (Lassak et al., 2013). On the contrary, deletion of Saci_1180 completely inhibited expression of FlaB (Lassak et al., 2013). Saci_1180 is a membrane bound one-component



regulator, termed ArnR, with an N-terminal helix turn helix (HTH) domain and two C-terminal transmembrane domains (**Figure 4**). In between these two domains a possible sensing domain is present which is believed to transmit a signal to the HTH domain. Interestingly, only in *S. acidocaldarius*, a gene duplication has occurred as downstream of *flaJ*, an *arnR* paralog is present, termed *arnR1* (see **Figure 4**). The HTH domains of ArnR and ArnR1 are nearly identical, whereas their sensing domains are quite different. Deletion of ArnR1 had a much less severe effect on *flaB* expression, indicating that it might be involved in fine tuning the expression of *flaB*. The archaellum operon in *S. acidocaldarius* has two transcriptional units of which one is *flaB* and the other locus is *flaX-J* (see **Figure 4**; Lassak et al., 2012b). Promoter fusion assays showed that ArnR and ArnR1 regulate the *flaB* promoter but not the *flaX-J* promoter. Moreover two inverted repeats, which are essential for the transcription of *flaB*, were identified in the promoter region of *flaB* (Lassak et al., 2012b).

The activity of members of the crenarchaeal archaellum regulatory network is regulated by protein phosphorylation. This was shown first for ArnA from *S. tokodaii* (Wang et al., 2010), then ArnA and ArnB were demonstrated to be phosphorylated by the protein kinase Saci_1193 (Reimann et al., 2012), now termed ArnC and only ArnB was phosphorylated by Saci-1694 in *S. acidocaldarius* (ArnD; Reimann et al., 2012). Moreover, in a phosphoproteomic study, the deletion of PP2A, the serine/threonine phosphatase of *S. acidocaldarius*, led to a strong overexpression of all archaealla genes, whereas the deletion of protein tyrosine phosphatase (PTP), the tyrosine phosphatase, had no effect on archaealla expression (**Figure 4**; Reimann et al., 2013).

Another regulator, AbfR1 (archaeal biofilm regulator 1) was also demonstrated to be involved in the archaellum regulatory network in *S. acidocaldarius* (Orell et al., 2013b). AbfR1 belongs to the Lrs14 regulator family of which two other members are also implicated in the regulation of biofilm growth. In the AbfR1 deletion mutant, the synthesis of archaealla components was impaired (**Figure 4**), leading to an increased production of EPS and biofilm (Orell et al., 2013a,b). In different archaea, the expression of other type IV pili also seems to influence the expression of archaealla. In *S. acidocaldarius*, the deletion of the gene encoding the membrane protein AapF from the archaeal adhesive pili operon unexpectedly led to a strong induction of archaealla, indicating that a switch exists that determines which of the surface structures is expressed (Henze et al., 2012). In *Hfx. volcanii*, it was recently observed that the deletion of *flgA2*, encoding the second archaeallin in this organism, led to hypermotile cells with an increased number of archaealla (Tripepi et al., 2013). Moreover, the presence of the H-domain of a set of type IV pilins (PilA1-A6) post-translationally influenced the assembly of archaealla in *Hfx. volcanii* (Esquivel and Pohlschroder, 2014). When the pilins were deleted, the cells were non-motile whereas the deletion of the pilus assembly machinery had no influence on archaealla assembly implying that the presence of the pilin subunits in the membrane is important for the regulation of archaeallum assembly in *Hfx. volcanii*.

In *Halocarcula marismortui*, the two different archaeallins FlaA2 and FlaB were produced under different growth conditions (Syutkin et al., 2014). Archaealla assembled from FlaA2 were

more stable than archaealla built from FlaB and, therefore, they were called ecoparalogs as they were produced under different environmental conditions.

CONCLUSION

During the last few years, an increasing amount of evidence has been collected proving that the archaeal motility structure is structurally and evolutionarily unrelated to the bacterial flagellum, leading directly to the proposal to rename the structure as the archaeallum. Although work on the regulation and the assembly of the archaeallum has been initiated, we still do not understand how this quite simple motor can achieve power comparable to that generated by the bacterial flagellum. Indeed, the recently measured swimming speeds of several hyperthermophilic archaea at 400–500 body lengths per second, clearly indicate that these organisms can be considered the fastest on Earth, all powered by archaealla (Herzog and Wirth, 2012). Future work will no doubt concentrate on this intriguing aspect of this unusual prokaryotic organelle as one major research focus even as efforts are also made to understand the regulation of the assembly of the structure and the critical role that the N-glycosylation pathway plays.

ACKNOWLEDGMENTS

Research in the authors' laboratory has been funded by Discovery Grants from the Natural Sciences and Engineering Research Council of Canada (to Ken F. Jarrell), the European Research Council (ERC starting grant 311523 ARCHAELLUM) and intramural funds of the Max Planck Society (to Sonja-Verena Albers). The authors wish to thank the various collaborators and students who have contributed so much to the studies generated in their laboratories.

REFERENCES

- Alam, M., and Oesterhelt, D. (1984). Morphology, function and isolation of halobacterial flagella. *J. Mol. Biol.* 176, 459–475. doi: 10.1016/0022-2836(84)90172-4
- Albers, S. V., and Driessens, A. J. M. (2005). Analysis of ATPases of putative secretion operons in the thermoacidophilic archaeon *Sulfolobus solfataricus*. *Microbiology* 151, 763–773. doi: 10.1099/mic.0.27699-0
- Albers, S. V., Szabo, Z., and Driessens, A. J. M. (2003). Archaeal homolog of bacterial type IV prepilin signal peptidases with broad substrate specificity. *J. Bacteriol.* 185, 3918–3925. doi: 10.1128/JB.185.13.3918-3925.2003
- Banerjee, A., Ghosh, A., Mills, D. J., Kahnt, J., Vonck, J., and Albers, S. V. (2012). FlaX, a unique component of the crenarchaeal archaeallum, forms oligomeric ring-shaped structures and interacts with the motor ATPase FlaI. *J. Biol. Chem.* 287, 43322–43330. doi: 10.1074/jbc.M112.414383
- Banerjee, A., Neiner, T., Tripp, P., and Albers, S. V. (2013). Insights into subunit interactions in the *Sulfolobus acidocaldarius* archaeallum cytoplasmic complex. *FEBS J.* 280, 6141–6149. doi: 10.1111/febs.12534
- Baranova, E., Fronzes, R., Garcia-Pino, A., Van Gerven, N., Papapostolou, D., Péhau-Arnaudet, G., et al. (2012). SbsB structure and lattice reconstruction unveil Ca^{2+} triggered S-layer assembly. *Nature* 487, 119–122. doi: 10.1038/nature11155
- Bardy, S. L., and Jarrell, K. F. (2003). Cleavage of preflagellins by an aspartic acid signal peptidase is essential for flagellation in the archaeon *Methanococcus voltae*. *Mol. Microbiol.* 50, 1339–1347. doi: 10.1046/j.1365-2958.2003.03758.x
- Bardy, S. L., and Jarrell, K. F. (2002). FlaK of the archaeon *Methanococcus maripaludis* possesses preflagellin peptidase activity. *FEMS Microbiol. Lett.* 208, 53–59. doi: 10.1111/j.1574-6968.2002.tb11060.x
- Bardy, S. L., Mori, T., Komoriya, K., Aizawa, S., and Jarrell, K. F. (2002). Identification and localization of flagellins FlaA and FlaB3 within flagella of *Methanococcus voltae*. *J. Bacteriol.* 184, 5223–5233. doi: 10.1128/JB.184.19.5223-5233.2002

- Bayley, D. P., and Jarrell, K. F. (1999). Overexpression of *Methanococcus voltae* flagellin subunits in *Escherichia coli* and *Pseudomonas aeruginosa*: a source of archaeal preflagellin. *J. Bacteriol.* 181, 4146–4153.
- Bayley, D. P., and Jarrell, K. F. (1998). Further evidence to suggest that archaeal flagella are related to bacterial type IV pili. *J. Mol. Evol.* 46, 370–373.
- Beznosov, S. N., Pyatibratov, M. G., and Fedorov, O. V. (2007). On the multicomponent nature of *Halobacterium salinarum* flagella. *Microbiology* 76, 435–441. doi: 10.1134/S002626170704008X
- Burrows, L. L. (2012). *Pseudomonas aeruginosa* twitching motility: type IV pili in action. *Annu. Rev. Microbiol.* 66, 493–520. doi: 10.1146/annurev-micro-092611-150055
- Burrows, L. L. (2005). Weapons of mass retraction. *Mol. Microbiol.* 57, 878–888. doi: 10.1111/j.1365-2958.2005.04703.x
- Campos, M., Cisneros, D. A., Nivaskumar, M., and Francetic, O. (2013). The type II secretion system—a dynamic fiber assembly nanomachine. *Res. Microbiol.* 164, 545–555. doi: 10.1016/j.resmic.2013.03.013
- Chaban, B., Logan, S. M., Kelly, J. F., and Jarrell, K. F. (2009). AglC and AglK are involved in biosynthesis and attachment of diacetylated glucuronic acid to the N-glycan in *Methanococcus voltae*. *J. Bacteriol.* 191, 187–195. doi: 10.1128/JB.00885-08
- Chaban, B., Ng, S. Y., Kanbe, M., Saltzman, I., Nimmo, G., Aizawa, S. I., et al. (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*. *Mol. Microbiol.* 66, 596–609. doi: 10.1111/j.1365-2958.2007.05913.x
- Chaban, B., Voisin, S., Kelly, J., Logan, S. M., and Jarrell, K. F. (2006). Identification of genes involved in the biosynthesis and attachment of *Methanococcus voltae* N-linked glycans: insight into N-linked glycosylation pathways in archaea. *Mol. Microbiol.* 61, 259–268. doi: 10.1111/j.1365-2958.2006.05226.x
- Chevance, F. F., and Hughes, K. T. (2008). Coordinating assembly of a bacterial macromolecular machine. *Nat. Rev. Microbiol.* 6, 455–465. doi: 10.1038/nrmicro1887
- Cohen-Krausz, S., and Trachtenberg, S. (2008). The flagellar filament structure of the extreme acidothermophile *Sulfolobus shibatae* B12 suggests that archaeabacterial flagella have a unique and common symmetry and design. *J. Mol. Biol.* 375, 1113–1124. doi: 10.1016/j.jmb.2007.10.048
- Cohen-Krausz, S., and Trachtenberg, S. (2002). The structure of the archeabacterial flagellar filament of the extreme halophile *Halobacterium salinarum* R1M1 and its relation to eubacterial flagellar filaments and type IV pili. *J. Mol. Biol.* 321, 383–395. doi: 10.1016/S0022-2836(02)00616-2
- Correia, J. D., and Jarrell, K. F. (2000). Posttranslational processing of *Methanococcus voltae* preflagellin by preflagellin peptidases of *M. voltae* and other methanogens. *J. Bacteriol.* 182, 855–858. doi: 10.1128/JB.182.3.855-858.2000
- Desmond, E., Brochier-Armanet, C., and Gribaldo, S. (2007). Phylogenomics of the archaeal flagellum: rare horizontal gene transfer in a unique motility structure. *BMC Evol. Biol.* 7:106. doi: 10.1186/1471-2148-7-106
- Ding, Y., Uchida, K., Aizawa, S. I., Murphy, K., Khursigara, C. M., Chong, J. P. J., et al. (in press). Effects of N-glycosylation site removal in archaellins on the assembly and function of archaella in *Methanococcus maripaludis*. *PLoS ONE* (accepted manuscript).
- Duan, X., and He, Z. G. (2011). Characterization of the specific interaction between archaeal FHA domain-containing protein and the promoter of a flagella-like gene-cluster and its regulation by phosphorylation. *Biochem. Biophys. Res. Commun.* 407, 242–247. doi: 10.1016/j.bbrc.2011.03.011
- Eichler, J. (2013). Extreme sweetness: protein glycosylation in archaea. *Nat. Rev. Microbiol.* 11, 151–156. doi: 10.1038/nrmicro2957
- Eichler, J. (2012). Response to Jarrell and Albers: the name says it all. *Trends Microbiol.* 20, 512–513. doi: 10.1016/j.tim.2012.08.007
- Esquivel, R. N., and Pohlschroder, M. (2014). A conserved type IV pilin signal peptide H-domain is critical for the post-translational regulation of flagella-dependent motility. *Mol. Microbiol.* 93, 494–504. doi: 10.1111/mmi.12673
- Esquivel, R. N., Xu, R., and Pohlschroder, M. (2013). Novel archaeal adhesion pilins with a conserved N terminus. *J. Bacteriol.* 17, 3808–3818. doi: 10.1128/JB.00572-13
- Faguy, D. M., Bayley, D. P., Kostyukova, A. S., Thomas, N. A., and Jarrell, K. F. (1996). Isolation and characterization of flagella and flagellin proteins from the thermoacidophilic archaea *Thermoplasma volcanium* and *Sulfolobus shibatae*. *J. Bacteriol.* 178, 902–905.
- Faguy, D. M., and Jarrell, K. F. (1999). A twisted tale: the origin and evolution of motility and chemotaxis in prokaryotes. *Microbiology* 145, 279–281. doi: 10.1099/13500872-145-2-279
- Faguy, D. M., Koval, S. F., and Jarrell, K. F. (1994a). Physical characterization of the flagella and flagellins from *Methanospirillum hungatei*. *J. Bacteriol.* 176, 7491–7498.
- Faguy, D. M., Jarrell, K. F., Kuzio, J., and Kalmokoff, M. L. (1994b). Molecular analysis of archaeal flagellins: similarity to the type IV pilin-transport superfamily widespread in bacteria. *Can. J. Microbiol.* 40, 67–71. doi: 10.1139/m94-011
- Gerl, L., Deutzmann, R., and Sumper, M. (1989). Halobacterial flagellins are encoded by a multigene family. Identification of all five gene products. *FEBS Lett.* 244, 137–140. doi: 10.1016/0014-5793(89)81179-2
- Gerl, L., and Sumper, M. (1988). Halobacterial flagellins are encoded by a multigene family. Characterization of five flagellin genes. *J. Biol. Chem.* 263, 13246–13251.
- Gernhardt, P., Possot, O., Foglino, M., Sibold, L., and Klein, A. (1990). Construction of an integration vector for use in the archaeabacterium *Methanococcus voltae* and expression of a eubacterial resistance gene. *Mol. Gen. Genet.* 221, 273–279. doi: 10.1007/BF00261731
- Ghosh, A., Hartung, S., van der Does, C., Tainer, J. A., and Albers, S. V. (2011). Archaeal flagellar ATPase motor shows ATP-dependent hexameric assembly and activity stimulation by specific lipid binding. *Biochem. J.* 437, 43–52. doi: 10.1042/BJ20110410
- Giltner, C. L., Nguyen, Y., and Burrows, L. L. (2012). Type IV pilin proteins: versatile molecular modules. *Microbiol. Mol. Biol. Rev.* 76, 740–772. doi: 10.1128/MMBR.00035-12
- Grogan, D. W. (1989). Phenotypic characterization of the archaeabacterial genus *Sulfolobus*: comparison of five wild-type strains. *J. Bacteriol.* 171, 6710–6719.
- Henché, A. L., Ghosh, A., Yu, X., Jeske, T., Egelman, E., and Albers, S. V. (2012). Structure and function of the adhesive type IV pilus of *Sulfolobus acidocaldarius*. *Environ. Microbiol.* 14, 3188–3202. doi: 10.1111/j.1462-2920.2012.02898.x
- Henché, A. L., van Wolferen, M., Ghosh, A., and Albers, S. V. (2014). Dissection of key determinants of cleavage activity in signal peptidase III (SPaseIII) PibD. *Extremophiles* 18, 905–913. doi: 10.1007/s00792-014-0675-4
- Hendrickson, E. L., Liu, Y., Rosas-Sandoval, G., Porat, I., Soll, D., Whitman, W. B., et al. (2008). Global responses of *Methanococcus maripaludis* to specific nutrient limitations and growth rate. *J. Bacteriol.* 190, 2198–2205. doi: 10.1128/JB.01805-07
- Herzog, B., and Wirth, R. (2012). Swimming behavior of selected species of archaea. *Appl. Environ. Microbiol.* 78, 1670–1674. doi: 10.1128/AEM.06723-11
- Hu, J., Xue, Y., Lee, S., and Ha, Y. (2011). The crystal structure of GxGD membrane protease FlaK. *Nature* 475, 528–531. doi: 10.1038/nature10218
- Jarrell, K. F., and Albers, S. V. (2012). The archaeallum: an old motility structure with a new name. *Trends Microbiol.* 20, 307–312. doi: 10.1016/j.tim.2012.04.007
- Jarrell, K. F., Bayley, D. P., and Kostyukova, A. S. (1996a). The archaeal flagellum: a unique motility structure. *J. Bacteriol.* 178, 5057–5064.
- Jarrell, K. F., Bayley, D. P., Florian, V., and Klein, A. (1996b). Isolation and characterization of insertional mutations in flagellin genes in the archaeon *Methanococcus voltae*. *Mol. Microbiol.* 20, 657–666. doi: 10.1046/j.1365-2958.1996.5371058.x
- Jarrell, K. F., Ding, Y., Meyer, B. H., Albers, S. V., Kaminski, L., and Eichler, J. (2014). N-linked glycosylation in archaea: a structural, functional, and genetic analysis. *Microbiol. Mol. Biol. Rev.* 78, 304–341. doi: 10.1128/MMBR.00052-13
- Jones, G. M., Wu, J., Ding, Y., Uchida, K., Aizawa, S., Robotham, A., et al. (2012). Identification of genes involved in the acetamidino group modification of the flagellin N-linked glycan of *Methanococcus maripaludis*. *J. Bacteriol.* 194, 2693–2702. doi: 10.1128/JB.00668-11
- Kalmokoff, M. L., and Jarrell, K. F. (1991). Cloning and sequencing of a multigene family encoding the flagellins of *Methanococcus voltae*. *J. Bacteriol.* 173, 7113–7125.
- Kalmokoff, M. L., Jarrell, K. F., and Koval, S. F. (1988). Isolation of flagella from the archaeabacterium *Methanococcus voltae* by phase separation with triton X-114. *J. Bacteriol.* 170, 1752–1758.
- Kalmokoff, M. L., Karnauchow, T. M., and Jarrell, K. F. (1990). Conserved N-terminal sequences in the flagellins of archaeabacteria. *Biochem. Biophys. Res. Commun.* 167, 154–160. doi: 10.1016/0006-291X(90)91744-D
- Kaminski, L., Naparstek, S., Kandiba, L., Cohen-Rosenzweig, C., Arbib, A., Konrad, Z., et al. (2013). Add salt, add sugar: N-glycosylation in *Haloflexax volcanii*. *Biochem. Soc. Trans.* 41, 432–435. doi: 10.1042/BST20120142

- Kelly, J., Logan, S. M., Jarrell, K. F., VanDyke, D. J., and Vinogradov, E. (2009). A novel N-linked flagellar glycan from *Methanococcus maripaludis*. *Carbohydr. Res.* 344, 648–653. doi: 10.1016/j.carres.2009.01.006
- Korotkov, K. V., Sandkvist, M., and Hol, W. G. (2012). The type II secretion system: biogenesis, molecular architecture and mechanism. *Nat. Rev. Microbiol.* 10, 336–351. doi: 10.1038/nrmicro2762
- Kupper, J., Marwan, W., Typke, D., Grunberg, H., Uwer, U., Gluch, M., et al. (1994). The flagellar bundle of *Halobacterium salinarium* is inserted into a distinct polar cap structure. *J. Bacteriol.* 176, 5184–5187.
- LaPointe, C. F., and Taylor, R. K. (2000). The type 4 prepilin peptidases comprise a novel family of aspartic acid proteases. *J. Biol. Chem.* 275, 1502–1510. doi: 10.1074/jbc.275.2.1502
- Lassak, K., Ghosh, A., and Albers, S. V. (2012a). Diversity, assembly and regulation of archaeal type IV pili-like and non-type-IV pili-like surface structures. *Res. Microbiol.* 163, 630–644. doi: 10.1016/j.resmic.2012.10.024
- Lassak, K., Neiner, T., Ghosh, A., Klingl, A., Wirth, R., and Albers, S. V. (2012b). Molecular analysis of the crenarchaeal flagellum. *Mol. Microbiol.* 83, 110–124. doi: 10.1111/j.1365-2958.2011.07916.x
- Lassak, K., Peeters, E., Wróbel, S., and Albers, S. V. (2013). The one-component system ArnR: a membrane-bound activator of the crenarchaeal archaeellum. *Mol. Microbiol.* 88, 125–139. doi: 10.1111/mmi.12173
- Leigh, J. A., Albers, S. V., Atomi, H., and Allers, T. (2011). Model organisms for genetics in the domain archaea: methanogens, halophiles, Thermococcales and Sulfolobales. *FEMS Microbiol. Rev.* 35, 577–608. doi: 10.1111/j.1574-6976.2011.00265.x
- Lory, S., and Strom, M. S. (1997). Structure-function relationship of type-IV prepilin peptidase of *Pseudomonas aeruginosa*—a review. *Gene* 192, 117–121. doi: 10.1016/S0378-1119(96)00830-X
- Macnab, R. M. (2004). Type III flagellar protein export and flagellar assembly. *Biochim. Biophys. Acta* 1694, 207–217. doi: 10.1016/j.bbamcr.2004.04.005
- Marwan, W., Alam, M., and Oesterhelt, D. (1991). Rotation and switching of the flagellar motor assembly in *Halobacterium halobium*. *J. Bacteriol.* 173, 1971–1977.
- Marwan, W., Schäfer, W., and Oesterhelt, D. (1990). Signal transduction in *Halobacterium* depends on fumarate. *EMBO J.* 9, 355–362.
- Merz, A. J., So, M., and Sheetz, M. P. (2000). Pilus retraction powers bacterial twitching motility. *Nature* 407, 98–102. doi: 10.1038/35024105
- Meyer, B. H., and Albers, S. V. (2013). Hot and sweet: protein glycosylation in crenarchaeota. *Biochem. Soc. Trans.* 41, 384–392. doi: 10.1042/BST20120296
- Meyer, B. H., Birich, A., and Albers, S. V. (2014). N-Glycosylation of the archaeal filament is not important for archaea assembly and motility, although N-Glycosylation is essential for motility in *Sulfolobus acidocaldarius*. *Biochimie* doi: 10.1016/j.biochi.2014.10.018 [Epub ahead of print].
- Meyer, B. H., Peyfoon, E., Dietrich, C., Hitchen, P., Dell, A., and Albers, S. V. (2013). Agl16, a thermophilic glycosyltransferase, mediating the last step of the N-glycan biosynthesis in the thermoacidophilic crenarchaeon *Sulfolobus acidocaldarius*. *J. Bacteriol.* 195, 2177–2186. doi: 10.1128/JB.00035-13
- Meyer, B. H., Zolghadr, B., Peyfoon, E., Pabst, M., Panico, M., Morris, H. R., et al. (2011). Sulfoquinovose synthase—an important enzyme in the N-glycosylation pathway of *Sulfolobus acidocaldarius*. *Mol. Microbiol.* 82, 1150–1163. doi: 10.1111/j.1365-2958.2011.07875.x
- Moore, B. C., and Leigh, J. A. (2005). Markerless mutagenesis in *Methanococcus maripaludis* demonstrates roles for alanine dehydrogenase, alanine racemase, and alanine permease. *J. Bacteriol.* 187, 972–979. doi: 10.1128/JB.187.3.972-979.2005
- Mukhopadhyay, B., Johnson, E. F., and Wolfe, R. S. (2000). A novel pH₂ control on the expression of flagella in the hyperthermophilic strictly hydrogenotrophic methanarchaeon *Methanococcus jannaschii*. *Proc. Natl. Acad. Sci. U.S.A.* 97, 11522–11527. doi: 10.1073/pnas.97.21.11522
- Nagahisa, K., Ezaki, S., Fujiwara, S., Imanaka, T., and Takagi, M. (1999). Sequence and transcriptional studies of five clustered flagellin genes from hyperthermophilic archaeon *Pyrococcus kodakarensis* KOD1. *FEMS Microbiol. Lett.* 178, 183–190. doi: 10.1111/j.1574-6968.1999.tb13776.x
- Ng, S. Y., Chaban, B., and Jarrell, K. F. (2006). Archaeal flagella, bacterial flagella and type IV pili: a comparison of genes and posttranslational modifications. *J. Mol. Microbiol. Biotechnol.* 11, 167–191. doi: 10.1159/000094053
- Ng, S. Y., Chaban, B., VanDyke, D. J., and Jarrell, K. F. (2007). Archaeal signal peptidases. *Microbiology* 153, 305–314. doi: 10.1099/mic.0.2006/003087-0
- Ng, S. Y., VanDyke, D. J., Chaban, B., Wu, J., Nosaka, Y., Aizawa, S., et al. (2009). Different minimal signal peptide lengths recognized by the archaeal prepilin-like peptidases FlaK and PibD. *J. Bacteriol.* 191, 6732–6740. doi: 10.1128/JB.00673-09
- Nivaskumar, M., Bouvier, G., Campos, M., Nadeau, N., Yu, X., Egelman, E. H., et al. (2014). Distinct docking and stabilization steps of the pseudopilus conformational transition path suggest rotational assembly of type IV pilus-like fibers. *Structure* 22, 685–696. doi: 10.1016/j.str.2014.03.001
- Nutsch, T., Oesterhelt, D., Gilles, E. D., and Marwan, W. (2005). A quantitative model of the switch cycle of an archaeal flagellar motor and its sensory control. *Biophys. J.* 89, 2307–2323. doi: 10.1529/biophysj.104.057570
- Orell, A., Fröls, S., and Albers, S. V. (2013a). Archaeal biofilms: the great unexplored. *Annu. Rev. Microbiol.* 67, 337–354. doi: 10.1146/annurev-micro-092412-155616
- Orell, A., Peeters, E., Vassen, V., Jachlewski, S., Schalles, S., Siebers, B., et al. (2013b). Lrs14 transcriptional regulators influence biofilm formation and cell motility of crenarchaea. *ISME J.* 7, 1886–1898. doi: 10.1038/ismej.2013.68
- Patel, G. B., Nash, J. H., Agnew, B. J., and Sprott, G. D. (1994). Natural and electroporation-mediated transformation of *Methanococcus voltae* protoplasts. *Appl. Environ. Microbiol.* 60, 903–907.
- Patenge, N., Berendes, A., Engelhardt, H., Schuster, S. C., and Oesterhelt, D. (2001). The fla gene cluster is involved in the biogenesis of flagella in *Halobacterium salinarum*. *Mol. Microbiol.* 41, 653–663. doi: 10.1046/j.1365-2958.2001.02542.x
- Peabody, C. R., Chung, Y. J., Yen, M. R., Vidal-Inigliardi, D., Pugsley, A. P., and Saier, M. H. Jr. (2003). Type II protein secretion and its relationship to bacterial type IV pili and archaeal flagella. *Microbiology* 149, 3051–3072. doi: 10.1099/mic.0.26364-0
- Pelicic, V. (2008). Type IV pili: *E pluribus unum?* *Mol. Microbiol.* 68, 827–837. doi: 10.1111/j.1365-2958.2008.06197.x
- Pyatibratov, M. G., Beznosov, S. N., Rachel, R., Tiktulpolo, E. I., Surin, A. K., Syutkin, A. S., et al. (2008). Alternative flagellar filament types in the haloarchaeon *Halococcus marismortui*. *Can. J. Microbiol.* 54, 835–844. doi: 10.1139/W08-076
- Reimann, J., Esser, D., Orell, A., Amman, F., Pham, T. K., Noirel, J., et al. (2013). Archaeal signal transduction: impact of protein phosphatase deletions on cell size, motility, and energy metabolism in *Sulfolobus acidocaldarius*. *Mol. Cell. Proteom.* 12, 3908–3923. doi: 10.1074/mcp.M113.027375
- Reimann, J., Lassak, K., Khadouma, S., Ettema, T. J., Yang, N., Driessen, A. J., et al. (2012). Regulation of archaea expression by the FHA and von Willebrand domain-containing proteins ArnA and ArnB in *Sulfolobus acidocaldarius*. *Mol. Microbiol.* 86, 24–36. doi: 10.1111/j.1365-2958.2012.08186.x
- Reindl, S., Ghosh, A., Williams, G. J., Lassak, K., Neiner, T., Henche, A. L., et al. (2013). Insights into Flal functions in archaeal motor assembly and motility from structures, conformations, and genetics. *Mol. Cell* 45, 1069–1082. doi: 10.1016/j.molcel.2013.01.014
- Rudolph, J., Nordmann, B., Storch, K. F., Gruenberg, H., Rodewald, K., and Oesterhelt, D. (1996). A family of halobacterial transducer proteins. *FEMS Microbiol. Lett.* 139, 161–168. doi: 10.1016/0378-1097(96)00136-X
- Rudolph, J., and Oesterhelt, D. (1996). Deletion analysis of the che operon in the archaeon *Halobacterium salinarum*. *J. Mol. Biol.* 258, 548–554. doi: 10.1006/jmbi.1996.0267
- Rudolph, J., Tolliday, N., Schmitt, C., Schuster, S. C., and Oesterhelt, D. (1995). Phosphorylation in halobacterial signal transduction. *EMBO J.* 14, 4249–4257.
- Schlesner, M., Miller, A., Besir, H., Aivaliotis, M., Streif, J., Scheffer, B., et al. (2012). The protein interaction network of a taxis signal transduction system in a halophilic archaeon. *BMC Microbiol.* 12:272. doi: 10.1186/1471-2180-12-272
- Schlesner, M., Miller, A., Streif, S., Staudinger, W. F., Muller, J., Scheffer, B., et al. (2009). Identification of archaea-specific chemotaxis proteins which interact with the flagellar apparatus. *BMC Microbiol.* 9:56. doi: 10.1186/1471-2180-9-56
- Shahapure, R., Driessen, R. P., Haurat, M. F., Albers, S. V., and Dame, R. T. (2014). The archaeellum: a rotating type IV pilus. *Mol. Microbiol.* 91, 716–723. doi: 10.1111/mmi.12486
- Stieglmeier, M., Klingl, A., Alves, R. J., Rittmann, S. K., Melcher, M., Leisch, N., et al. (2014). *Nitrososphaera viennensis* gen. nov., sp. nov., an aerobic and mesophilic, ammonia-oxidizing archaeon from soil and a member of the archaeal phylum Thaumarchaeota. *Int. J. Syst. Evol. Microbiol.* 64, 2738–2752. doi: 10.1099/ijss.0.063172-0

- Streif, S., Staudinger, W. F., Marwan, W., and Oesterhelt, D. (2008). Flagellar rotation in the archaeon *Halobacterium salinarum* depends on ATP. *J. Mol. Biol.* 384, 1–8. doi: 10.1016/j.jmb.2008.08.057
- Strom, M. S., Nunn, D. N., and Lory, S. (1993). A single bifunctional enzyme, PilD, catalyzes cleavage and N-methylation of proteins belonging to the type IV pilin family. *Proc. Natl. Acad. Sci. U.S.A.* 90, 2404–2408. doi: 10.1073/pnas.90.6.2404
- Strom, M. S., Nunn, D. N., and Lory, S. (1994). Posttranslational processing of type IV prepilin and homologs by PilD of *Pseudomonas aeruginosa*. *Methods Enzymol.* 235, 527–540. doi: 10.1016/0076-6879(94)35168-6
- Sumper, M. (1987). Halobacterial glycoprotein biosynthesis. *Biochim. Biophys. Acta* 906, 69–79. doi: 10.1016/0304-4157(87)90005-0
- Syutkin, A. S., Pyatibratov, M. G., Galitskaya, O. V., Rodríguez-Valera, F., and Fedorov, O. V. (2014). *Haloarcula marismortui* archaeallin genes as ecoparalogs. *Extremophiles* 18, 341–349. doi: 10.1007/s00792-013-0619-4
- Szabo, Z., Albers, S. V., and Driessen, A. J. M. (2006). Active-site residues in the type IV prepilin peptidase homologue PibD from the archaeon *Sulfolobus solfataricus*. *J. Bacteriol.* 188, 1437–1443. doi: 10.1128/JB.188.4.1437-1443.2006
- Szabo, Z., Stahl, A. O., Albers, S. V., Kissinger, J. C., Driessen, A. J. M., and Pohlschröder, M. (2007a). Identification of diverse archaeal proteins with class III signal peptides cleaved by distinct archaeal prepilin peptidases. *J. Bacteriol.* 189, 772–778. doi: 10.1128/JB.01547-06
- Szabo, Z., Sani, M., Groeneveld, M., Zolghadr, B., Schelert, J., Albers, S. V., et al. (2007b). Flagellar motility and structure in the hyperthermoacidophilic archaeon *Sulfolobus solfataricus*. *J. Bacteriol.* 189, 4305–4309. doi: 10.1128/JB.00042-07
- Szurmant, H., and Ordal, G. W. (2004). Diversity in chemotaxis mechanisms among the bacteria and archaea. *Microbiol. Mol. Biol. Rev.* 68, 301–319. doi: 10.1128/MMBR.68.2.301-319.2004
- Thomas, N. A., Chao, E. D., and Jarrell, K. F. (2001a). Identification of amino acids in the leader peptide of *Methanococcus voltae* preflagellin that are important in posttranslational processing. *Arch. Microbiol.* 175, 263–269. doi: 10.1007/s002030100254
- Thomas, N. A., Pawson, C. T., and Jarrell, K. F. (2001b). Insertional inactivation of the flaH gene in the archaeon *Methanococcus voltae* results in non-flagellated cells. *Mol. Genet. Genomics* 265, 596–603. doi: 10.1007/s004380100451
- Thomas, N. A., and Jarrell, K. F. (2001). Characterization of flagellum gene families of methanogenic archaea and localization of novel flagellum accessory proteins. *J. Bacteriol.* 183, 7154–7164. doi: 10.1128/JB.183.24.7154-7164.2001
- Trachtenberg, S., and Cohen-Krausz, S. (2006). The archaeabacterial flagellar filament: a bacterial propeller with a pilus-like structure. *J. Mol. Microbiol. Biotechnol.* 11, 208–220. doi: 10.1159/000094055
- Tripepi, M., Esquivel, R. N., Wirth, R., and Pohlschröder, M. (2013). *Halofexax volcanii* cells lacking the flagellin FlgA2 are hypermotile. *Microbiology* 159, 2249–2258. doi: 10.1099/mic.0.069617-0
- Tripepi, M., Imam, S., and Pohlschröder, M. (2010). *Halofexax volcanii* flagella are required for motility but are not involved in PibD-dependent surface adhesion. *J. Bacteriol.* 192, 3093–3102. doi: 10.1128/JB.00133-10
- Tripepi, M., You, J., Temel, S., Önder, Ö., Brisson, D., and Pohlschröder, M. (2012). N-glycosylation of *Halofexax volcanii* flagellins requires known agl proteins and is essential for biosynthesis of stable flagella. *J. Bacteriol.* 194, 4876–4887. doi: 10.1128/JB.00731-12
- Tumbula, D. L., Makula, R. A., and Whitman, W. B. (1994). Transformation of *Methanococcus maripaludis* and identification of a pstI-like restriction system. *FEMS Microbiol. Lett.* 121, 309–314. doi: 10.1111/j.1574-6968.1994.tb07118.x
- VanDyke, D. J., Wu, J., Logan, S. M., Kelly, J. F., Mizuno, S., Aizawa, S. I., et al. (2009). Identification of genes involved in the assembly and attachment of a novel flagellin N-linked tetrasaccharide important for motility in the archaeon *Methanococcus maripaludis*. *Mol. Microbiol.* 72, 633–644. doi: 10.1111/j.1365-2958.2009.06671.x
- VanDyke, D. J., Wu, J., Ng, S. Y., Kanbe, M., Chaban, B., Aizawa, S. I., et al. (2008). Identification of putative acetyltransferase gene, MMP0350, which affects proper assembly of both flagella and pili in the archaeon *Methanococcus maripaludis*. *J. Bacteriol.* 190, 5300–5307. doi: 10.1128/JB.00474-08
- Voisin, S., Houlston, R. S., Kelly, J., Brisson, J. R., Watson, D., Bardy, S. L., et al. (2005). Identification and characterization of the unique N-linked glycan common to the flagellins and S-layer glycoprotein of *Methanococcus voltae*. *J. Biol. Chem.* 280, 16586–16593. doi: 10.1074/jbc.M500329200
- Wagner, M., van Wolferen, M., Wagner, A., Lassak, K., Meyer, B., Reimann, J., et al. (2012). Versatile genetic toolbox for the crenarchaeote *Sulfolobus acidocaldarius*. *Front. Microbiol.* 3:214. doi: 10.3389/fmicb.2012.00214
- Wang, B., Yang, S., Zhang, L., and He, Z. G. (2010). Archaeal eukaryote-like serine/threonine protein kinase interacts with and phosphorylates a forkhead-associated-domain-containing protein. *J. Bacteriol.* 192, 1956–1964. doi: 10.1128/JB.01471-09
- Wieland, F., Paul, G., and Sumper, M. (1985). Halobacterial flagellins are sulfated glycoproteins. *J. Biol. Chem.* 260, 15180–15185.
- Wirth, R. (2012). Response to Jarrell and Albers: seven letters less does not say more. *Trends Microbiol.* 20, 511–512. doi: 10.1016/j.tim.2012.07.007
- Xia, Q., Wang, T., Hendrickson, E. L., Lie, T. J., Hackett, M., and Leigh, J. A. (2009). Quantitative proteomics of nutrient limitation in the hydrogenotrophic methanogen *Methanococcus maripaludis*. *BMC Microbiol.* 9:149. doi: 10.1186/1471-2180-9-149
- Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 11 November 2014; accepted: 08 January 2015; published online: 27 January 2015.

Citation: Albers SV and Jarrell KF (2015) The archaeellum: how Archaea swim. *Front. Microbiol.* 6:23. doi: 10.3389/fmicb.2015.00023

This article was submitted to Microbial Physiology and Metabolism, a section of the journal *Frontiers in Microbiology*.

Copyright © 2015 Albers and Jarrell. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Biofilm formation of mucosa-associated methanarchaeal strains

Corinna Bang^{1†}, Claudia Ehlers^{1†}, Alvaro Orell^{2,3†}, Daniela Prasse¹, Marlene Spinner⁴, Stanislav N. Gorb⁴, Sonja-Verena Albers² and Ruth A. Schmitz^{1*}

¹ Institute for General Microbiology, University of Kiel, Kiel, Germany

² Molecular Biology of Archaea, Max Planck Institute for Terrestrial Microbiology, Marburg, Germany

³ Molecular Microbiology of Extremophiles Research Group, Centre for Genomics and Bioinformatics, Faculty of Sciences, Universidad Mayor, Santiago, Chile

⁴ Functional Morphology and Biomechanics, Zoological Institute, University of Kiel, Kiel, Germany

Edited by:

Mechthild Pohlschroder, University of Pennsylvania, USA

Reviewed by:

Mike L. Dyall-Smith, Charles Sturt University, Australia

Wolfgang Buckel, Philipps-Universität Marburg, Germany

Sabrina Froels, Technische Universität Darmstadt, Germany

***Correspondence:**

Ruth A. Schmitz, Institute for General Microbiology, University of Kiel, Am Botanischen Garten 1-9, D-24118 Kiel, Germany

e-mail: rschmitz@ifam.uni-kiel.de

[†]Corinna Bang, Claudia Ehlers, and Alvaro Orell have contributed equally to this work.

Although in nature most microorganisms are known to occur predominantly in consortia or biofilms, data on archaeal biofilm formation are in general scarce. Here, the ability of three methanarchaeal strains, *Methanobrevibacter smithii* and *Methanospaera stadtmanae*, which form part of the human gut microbiota, and the *Methanosarcina mazei* strain Gö1 to grow on different surfaces and form biofilms was investigated. All three strains adhered to the substrate mica and grew predominantly as bilayers on its surface as demonstrated by confocal laser scanning microscopy analyses, though the formation of multi-layered biofilms of *Methanospaera stadtmanae* and *Methanobrevibacter smithii* was observed as well. Stable biofilm formation was further confirmed by scanning electron microscopy analysis. *Methanosarcina mazei* and *Methanobrevibacter smithii* also formed multi-layered biofilms in uncoated plastic µ-dishes™, which were very similar in morphology and reached a height of up to 40 µm. In contrast, biofilms formed by *Methanospaera stadtmanae* reached only a height of 2 µm. Staining with the two lectins ConA and IB4 indicated that all three strains produced relatively low amounts of extracellular polysaccharides most likely containing glucose, mannose, and galactose. Taken together, this study provides the first evidence that methanarchaea can develop and form biofilms on different substrates and thus, will contribute to our knowledge on the appearance and physiological role of *Methanobrevibacter smithii* and *Methanospaera stadtmanae* in the human intestine.

Keywords: biofilms, methanarchaea, human gut, microbiota

INTRODUCTION

Growth of microorganisms as complex microbial communities is the predominant lifestyle in nature and has been shown to occur on a wide variety of surfaces including living tissues (Donlan, 2002). Although the human gut harbors trillions of microorganisms forming a complex ecological community (Whitman et al., 1998; Hopkins et al., 2001; Macpherson and Harris, 2004; Abreu et al., 2005; Ley et al., 2006; O'Hara and Shanahan, 2006; Artis, 2008; Lozupone et al., 2012), the existence and significance of mucosa-associated biofilms was not considered for many years (Dongari-Bagtzoglou, 2008). However, during the last decade, the increasing numbers of studies dealing with the overall microbial diversity in the human gut have demonstrated bacterial biofilm formation on the mucus itself or the epithelial surface (Macfarlane and Dillon, 2007; Macfarlane et al., 2011). In this regard, the biofilm development on mucosal surfaces was shown to depend not only on environmental and nutritional factors but also on the host defense mechanisms (Macfarlane and Dillon, 2007). Particularly in patients suffering from inflammatory bowel diseases (IBD) the density and composition of mucosal biofilms has been shown to alter significantly when compared to healthy controls (Swidsinski et al., 2005). Biofilm formation on human mucosa surfaces are so-called "mucosal biofilms" involving microbial adhesion to the mucosa with subsequent cell-to-cell adhesion leading to

multicellular structure formation (Post et al., 2004; Dongari-Bagtzoglou, 2008). Structurally, members of those biofilms are embedded in a matrix of extracellular polymeric substances (EPS) that mediates protective functions as well as nutrient supply and enables communication between biofilm forming microorganisms (Flemming and Wingender, 2010). In addition, biofilm-associated microorganisms are phenotypically different from their planktonic counterpart, as indicated by the finding that large suites of genes are differentially transcribed (An and Parsek, 2007). Whereas environmental biofilms are mostly composed of various microbial species, medically relevant biofilms on epithelial tissues (such as the lung, the gut and the oral cavity) that are associated with infectious diseases are often composed of just a few species (Donlan, 2002). In this respect, diversity in mucosal biofilms was also found to be low, when compared to the overall microbial diversity in the human gut (Swidsinski et al., 2005; Dongari-Bagtzoglou, 2008). Studies of mucosal biofilms are mainly exclusively focused on bacterial species, though several members of the archaeal domain have been identified to be stable components of the complex microbial community in the human gut (Whitman et al., 1998; O'Hara and Shanahan, 2006; Hill and Artis, 2010). In particular, the methanarchaea *Methanobrevibacter smithii* and *Methanospaera stadtmanae* are known to be part of the human gut microbiota (Miller et al., 1982, 1984; Lovley et al., 1984; Miller and Wolin,

1985; Weaver et al., 1986; Backhed et al., 2005; Eckburg et al., 2005; Levitt et al., 2006; Dridi et al., 2009). Notably, *Methanobrevibacter smithii* has been shown to inhabit nearly every human individual gut ecosystem, whereas *Methanospaera stadtmanae* was found in 30% of individuals (Dridi et al., 2009; Dridi, 2012). Both strains, *Methanobrevibacter smithii* and *Methanospaera stadtmanae*, have been shown to be involved in fermentation processes by converting bacterial fermentation products like hydrogen, organic acids (e.g., formate, acetate), and carbon dioxide, to methane (Miller et al., 1984; Samuel and Gordon, 2006; Samuel et al., 2007). Apart from that, the knowledge on further functions of *Methanobrevibacter smithii* and *Methanospaera stadtmanae* in the human intestinal ecosystem is still limited, though *Methanobrevibacter smithii*'s role in the development of adiposity was proposed in several studies (Samuel et al., 2008; Mathur et al., 2013). Very recently, an influence of those predominating methanarchaeal strains on the immunomodulation within the human intestine was obtained (Bang et al., 2014). In addition, *Methanobrevibacter oralis*, which is a close relative of *Methanobrevibacter smithii*, was anticipated to play a role in the manifestation of periodontal disease and meanwhile its prevalence was shown to be increased in patients suffering from chronic periodontitis (Kulik et al., 2001; Vianna et al., 2006; Ashok et al., 2013). In general, these findings argue that the impact of (methano)archaea on human's health and disease might have been underestimated until now.

With respect to the identified syntrophic interactions between methanarchaea and bacterial gut inhabitants (Samuel and Gordon, 2006; Samuel et al., 2007), it appears most likely that methanarchaeal strains occur as biofilms within the human intestine together with gut bacteria such as *Bacteroides* species (Swidsinski et al., 2005). However, information on archaeal biofilm formation is in general rare and only a few examples are reported, which are reviewed in Fröls (2013) and Orell et al. (2013). On the other hand, it is known that the methanarchaeal strain *Methanosarcina mazei* easily forms cellular aggregates in the presence of environmental stressors (Mayerhofer et al., 1992). Thus, understanding how methanarchaea interact with gut bacteria and the mucosa itself potentially by forming biofilms is crucial for upcoming studies dealing with the immunomodulatory role of those microorganisms. Consequently, the aim of this study was to evaluate the general ability of the methanarchaeal gut inhabitants *Methanobrevibacter smithii* and *Methanospaera stadtmanae* to form biofilms on two different substrates as well as to examine structural characteristics of these biofilms, in particular in comparison with a methanarchaeon originally isolated from sewage sludge, *Methanosarcina mazei* strain Gö1.

MATERIALS AND METHODS

STRAINS AND GROWTH CONDITIONS

Methanosarcina mazei strain Gö1 (DSM 3647), *Methanospaera stadtmanae* (DSM 3091) and *Methanobrevibacter smithii* (DSM 861) were obtained from the Deutsche Sammlung von Mikroorganismen und Zellkulturen (DSMZ, Braunschweig, Germany). *Methanosarcina mazei* strain Gö1 was grown in minimal medium under strict anaerobic conditions as described earlier (Ehlers et al.,

2002; Bang et al., 2012). *Methanospaera stadtmanae* was grown in medium 322 (according to the DSMZ, <http://www.dsmz.de>) and *Methanobrevibacter smithii* in medium 119 (according to the DSMZ, <http://www.dsmz.de>) both containing 10% rumen fluid. The reductants Na₂S (1.25 mM) and cysteine (2.5 mM) were added after autoclaving of media and 1.5 atm H₂/CO₂ (80/20 vol/vol) was used as a gas phase. Medium for *Methanospaera stadtmanae* was further complemented with 150 mM methanol prior inoculation. To prevent bacterial contamination, the medium for all strains was in general supplemented with 100 µg/ml ampicillin.

GROWTH ON MICA

For initial adherence experiments of the methanarchaeal strains, mica plates (Baltic Präparation, Niesgrau, Germany) with an edge length of 0.5 cm were used. Those mica pieces were transferred into hungate tubes, autoclaved, and placed into an anaerobic chamber with an atmosphere of N₂/CO₂/H₂ (78/20/2 vol/vol/vol), which was constantly circulated through a 0.3 µm filter system (Coy Laboratory Products Inc., MI, USA) to ensure anaerobic and semi-sterile conditions. At least 24 h later, 3 ml of reduced and complemented media were filled in the prepared hungate tubes and 1 × 10⁷ cells of the respective methanarchaeal preculture during its exponential growth phase were added. Those preparations were vertically incubated and samples were taken after 48, 72, and 96 h. Samples for microscopic analysis were fixed with 2% glutaraldehyde (Sigma-Aldrich Biochemie GmbH, Hamburg, Germany, Number G5882), which was directly added to the hungate tubes for at least 4 h at 4°C prior washing in minimal medium and microscopic examination at 1000× magnification using an Axio Lab microscope (Carl Zeiss MicroImaging GmbH, Jena, Germany) supplied with a digital camera (AxioCam Mr5, Carl Zeiss MicroImaging GmbH). Phase-contrast micrographs were captured using the digital image analysis software AxioVision Rel. 4.7.1 (Carl Zeiss MicroImaging GmbH). In addition, fixed samples after 48 h of growth were visualized by using the autofluorescence of glutaraldehyde in a TCS-SP5 confocal laser scanning microscope (Leica, Bensheim, Germany) at an excitation wavelength of 520 nm and an emission wavelength of 540 nm. Obtained image data were edited by using the IMARIS software package (Bitplane AG, Zürich, Switzerland).

SCANNING ELECTRON MICROSCOPY (SEM)

After growing periods of 48 h (*Methanosarcina mazei*, *Methanospaera stadtmanae*) or 72 h (*Methanobrevibacter smithii*) cultures were prepared as described above and mica plates were fixed on the aluminum stubs with double-sided carbon conductive tapes (Plano, Wetzlar, Germany). Subsequently, samples were air dried in a desiccator with silica gel (Merck KGaA, Darmstadt, Germany) for a period of 72 h. After coating with a 10 nm thick layer of gold-palladium in a sputter coater (Leica EM SCD500, Leica Microsystems GmbH, Wetzlar, Germany), samples were examined in SEM Hitachi S-4800 (Hitachi High-Technologies Corp., Tokyo, Japan) at an accelerating voltage of 3 kV.

CONFOCAL LASER SCANNING MICROSCOPY (CLSM)

For CLSM images, the cells were grown for 72 h in uncoated plastic dishesTM (µ-DishesTM, 35 mm high; Ibidi, Martinsried,

Germany). Prior to confocal microscopy, the liquid supernatant of the biofilm, with the planktonic cells, was removed and 2 ml fresh medium was added. Images were recorded on an inverted TCS-SP5 confocal microscope (Leica). DAPI (4,6-diamidino-2-phenylindole), dissolved in water to 300 µg/ml, was used to visualize the cells of the biofilm. For this reason, 7 µl of the DAPI stock solution in 2 ml fresh medium were added to the biofilm, incubated at room temperature for at least 10 min and subsequently washed twice with 2 ml fresh medium. Images were taken at an excitation wavelength of 345 nm and an emission wavelength of 455 nm. Fluorescently labeled lectins were employed to visualize the EPS (extracellular polymeric substances) of the biofilms. Prior addition of lectins to the biofilm, fluorescein-conjugated concanavalin A (ConA; 5 mg/ml; Life Technologies GmbH, Darmstadt, Germany), which binds to α -mannopyranosyl and α -glucopyranosyl residues, was dissolved in 20 mM sodium bicarbonate (pH 8.0) to a final concentration of 10 mg/ml. Fluorescein-conjugated ConA has an excitation wavelength of 494 nm and an emission wavelength of 518 nm. Alexa FluorH594-conjugated IB4, specific for α -D-galactosyl residues (isolectin GS-IB4 from *Griffonia simplicifolia* 1 mg/ml; Life Technologies GmbH) was dissolved in 100 mM Tris-HCl pH 7.4 and 0.5 mM CaCl₂ to a final concentration of 8 mg/ml. The Alexa Fluor-conjugated lectin, which has an excitation wavelength of 591 nm and an emission wavelength of 618 nm, was used in concert with ConA. The lectin–biofilm mixtures were incubated at room temperature for 20–30 min in the absence of light. After incubation, the biofilm was washed with fresh media to remove excess label and images were taken by CSLM. Image data were processed by using the IMARIS software package (Bitplane AG).

DETERMINATION OF SURFACE COVERAGE

To evaluate cell surface coverage of the biofilms, pictures of the bottom layer were taken using a differential interference contrast (DIC) objective. Twelve images at different microscopy fields were recorded. By using Adobe Photoshop CS2 software DIC pictures were converted into black/white in order to calculate number of pixels/area thus representing the percentage surface coverage. Cell surface coverage determinations were performed in three biological replicates.

RESULTS

The aim of this study was to examine the general ability of several methanarchaeal strains to form biofilms and to evaluate potential differences between the human gut inhabitants *Methanobrevibacter smithii* and *Methanospaera stadtmanae* as well as *Methanosarcina mazei* strain Gö1, a member of the Methanosaecales inhabiting various anoxic environments (Depenmeier et al., 2002; Chaban et al., 2006).

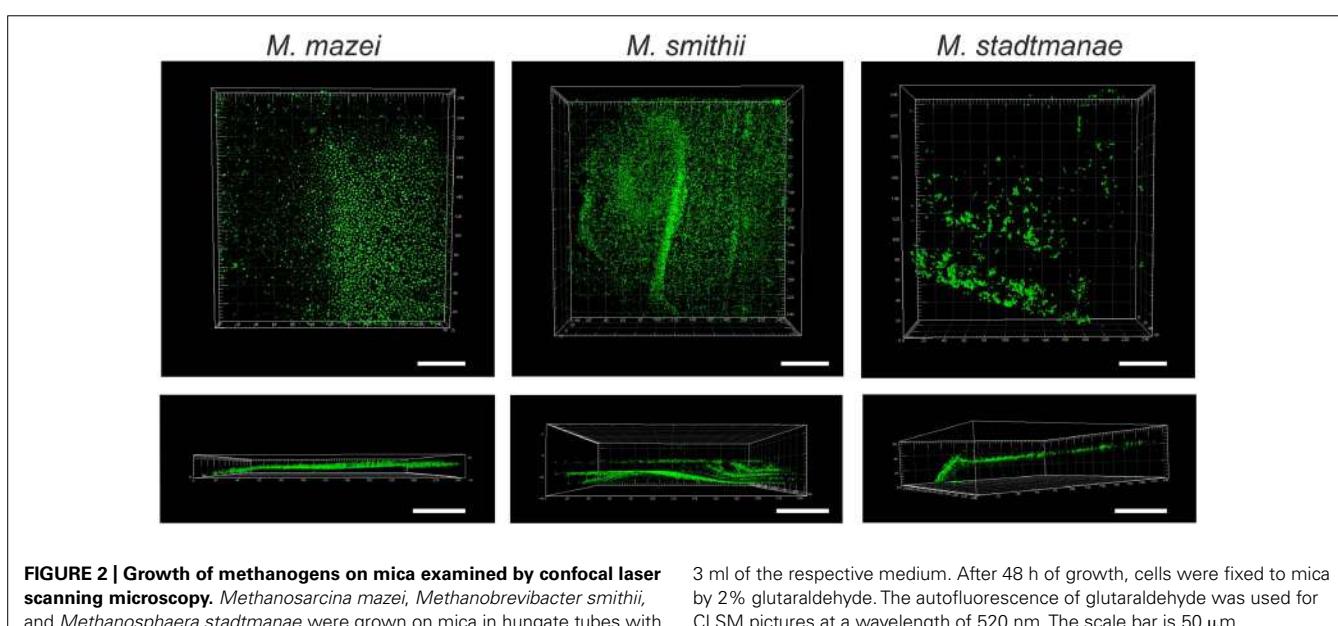
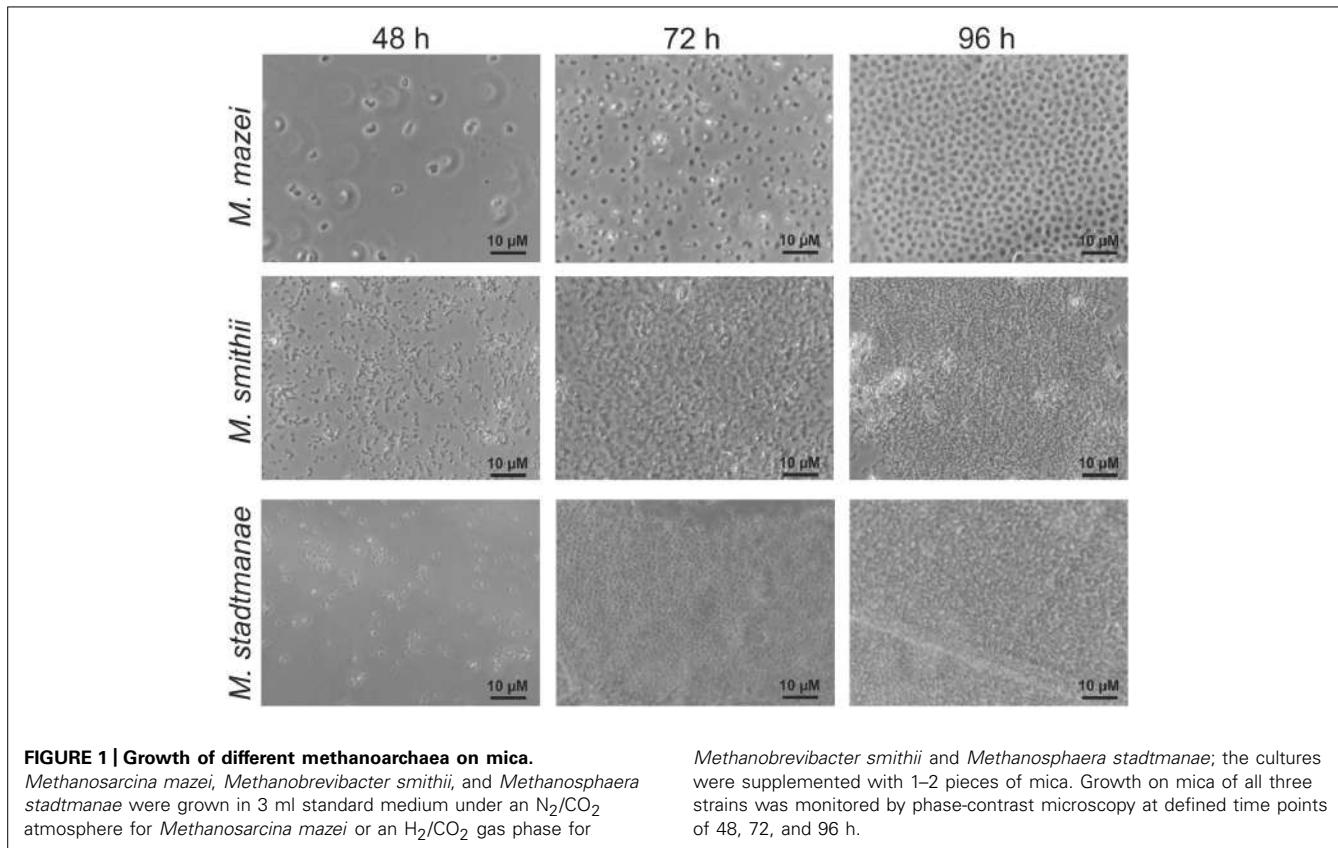
Since no information was available on biofilm formation of methanarchaeal gut inhabitants, initially static growth of *Methanobrevibacter smithii* and *Methanospaera stadtmanae* as well as of *Methanosarcina mazei* strain Gö1 on mica plates was investigated. For this purpose, methanarchaeal strains were grown for varying time periods in strain-specific media containing small pieces of mica plates. These preparations were fixed with 2% glutaraldehyde and washed prior to the subsequent

analysis. Phase-contrast microscopic examination of these mica plates after 48, 72, and 96 h revealed growth on mica for all three strains with increasing cell numbers during the time course (**Figure 1**). However, differences in the phenotype of the strains were observed during biofilm development. On the one hand, even after 96 h a precise space between the high numbers of attached *Methanosarcina mazei* cells resulting in no direct cell-to-cell contact was observed, which might potentially be coordinated by pili or EPS components. On the other hand, cells of *Methanobrevibacter smithii* and *Methanospaera stadtmanae* strongly formed aggregates attached to the surface with increasing cell numbers (**Figure 1**). In addition, all three strains appeared to form predominantly bilayer biofilms (**Figures 1–3**), although multi-layered growth was occasionally observed for *Methanobrevibacter smithii* and *Methanospaera stadtmanae* (**Figure 3**).

Confocal laser scanning microscopy was used to further visualize biofilm formation by the methanarchaeal strains after 48 h on the prepared mica plates. The autofluorescence of glutaraldehyde enabled visualization of methanarchaeal cell growth on the surface of mica plates by applying the respective wavelength (520 nm). This analysis revealed widespread adhesion of *Methanosarcina mazei* and *Methanobrevibacter smithii* cells over the surface of mica plates, whereas only small areas were shown to be overgrown by *Methanospaera stadtmanae* (**Figure 2**). Since comparable initial cell numbers of all strains were used as inoculum, these results demonstrated that *Methanobrevibacter smithii* and *Methanosarcina mazei* adhered better to the smooth surface of the mica when compared to *Methanospaera stadtmanae*.

Further morphological characteristics of the methanarchaeal biofilms were analyzed by using SEM. Cell-to-cell adhesion as well as adhesion to the mica surface could be demonstrated using this method for *Methanobrevibacter smithii* and *Methanospaera stadtmanae* (**Figure 3**). In addition, secretion of potential extracellular polymeric substances (EPS) by all tested strains was observed (**Figure 3**). The secretion of this potential EPS by *Methanobrevibacter smithii* and *Methanospaera stadtmanae* rose with increasing attached cell numbers; however, the highest production of potential EPS was detected for *Methanosarcina mazei*. Probably due to the air dry conditions, *Methanosarcina mazei* cells lost their integrity and thus, in SEM analysis of *Methanosarcina mazei* no single cells were found (**Figure 3** *Methanosarcina mazei*). Since difficulties during the preparation procedures of *Methanosarcina mazei* for electron microscopy analyses were already observed during an earlier study (Bang et al., 2012), SEM analysis only indicated the general ability of *Methanosarcina mazei* to form biofilms on mica.

For a more detailed analysis of the biofilm formation, *Methanosarcina mazei*, *Methanobrevibacter smithii*, and *Methanospaera stadtmanae* were incubated in strain-specific medium under static conditions in uncoated plastic μ -dishesTM for 72 h. Subsequently, the biofilms formed were analyzed by CLSM and DAPI was used to visualize the cells. Structurally, this method revealed that *Methanosarcina mazei* and *Methanobrevibacter smithii* formed multi-layered biofilms being very similar in respect to their morphology and height (up to 40 µm; **Figure 4**, DAPI). However, biofilms formed by *Methanobrevibacter smithii*



appeared to be denser and more compacted when compared to *Methanosarcina mazei*. In contrast to *Methanosarcina mazei* and *Methanobrevibacter smithii*, biofilms formed by *Methanospaera stadtmanae* developed only to a height of 2 µm, with occasional tower-like structures unevenly distributed on the surface (**Figure 4**, left panel, DAPI).

In order to confirm the observed production of potential EPS by the methanotrichaeal biofilms (visible in **Figure 3**), these sessile communities were additionally stained using two different fluorescently labeled lectins, ConA and IB4. A strong ConA signal was observed in biofilms formed by all three strains, indicating the presence of glucose and/or mannose residues. However, the

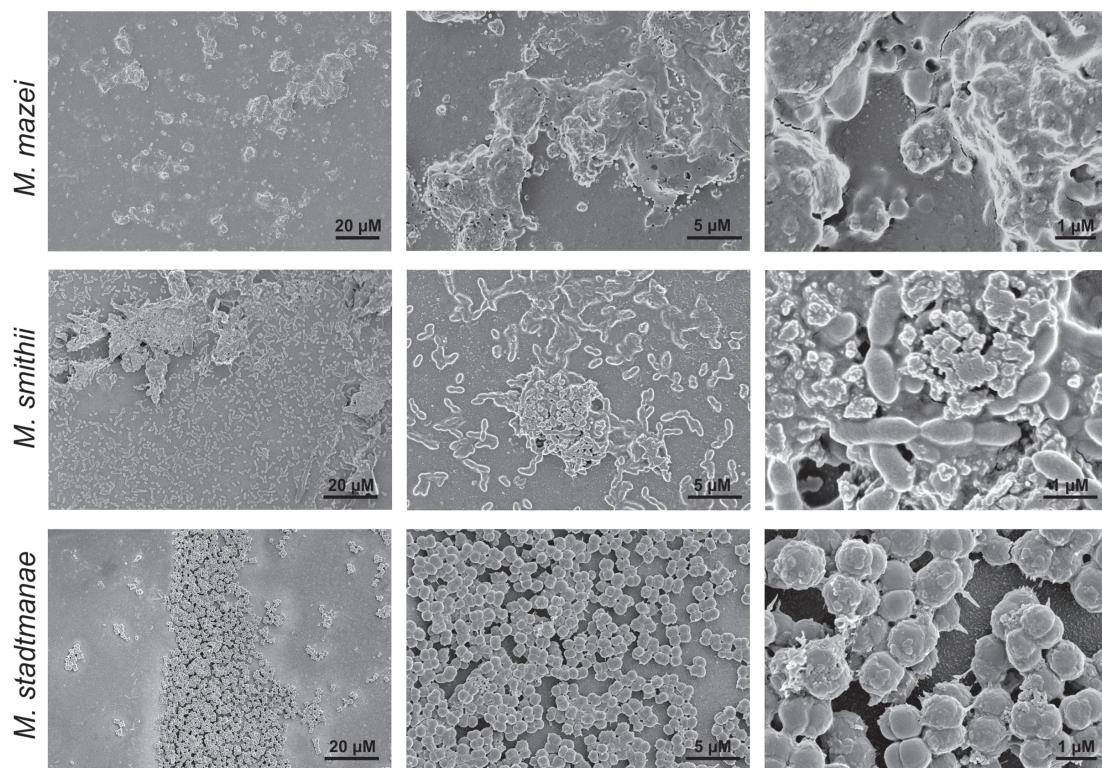


FIGURE 3 | Growth of methanogens on mica examined by SEM. *Methanosarcina mazei*, *Methanobrevibacter smithii*, and *Methanospaera stadtmanae* were grown on mica in挂管 tubes with 3 ml of the respective medium. After 48 h

(*Methanosarcina mazei* and *Methanospaera stadtmanae*) and 72 h (*Methanobrevibacter smithii*) of growth, cells were fixed to mica by 2% glutaraldehyde. Images are representative for the respective sample.

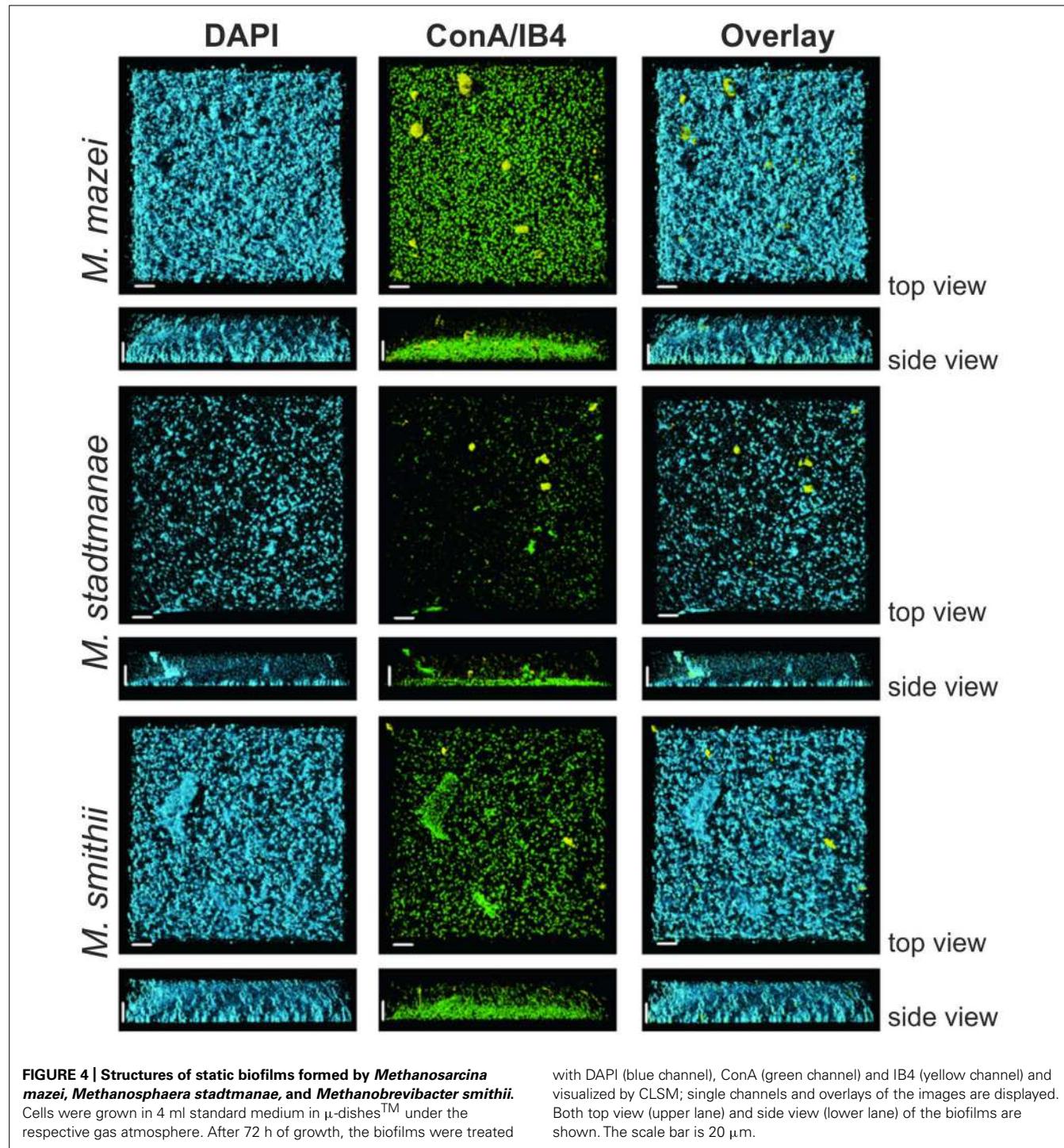
ConA signal (Figure 4, green signal) closely co-localized with the DAPI stained cells (Figure 4, blue signal). On the contrary, the IB4 signal (Figure 4, yellow signal), which is specific for α -galactosyl sugar residues, was only detected in very few clusters in all three biofilms and appeared not to be directly co-localized with cells.

The bottom layers of the static biofilms formed by *Methanosarcina mazei*, *Methanospaera stadtmanae*, and *Methanobrevibacter smithii* were imaged in order to calculate the respective surface coverage of the biofilms. This analysis revealed 50% higher coverage of the surface at the bottom of the μ -dish in the *Methanosarcina mazei* biofilm, when compared to the *Methanobrevibacter smithii* biofilm (Figure 5). Moreover, the surface coverage of the bottom layer of biofilms formed by *Methanospaera stadtmanae* was found to be only 30% of the one from *Methanosarcina mazei* and about 70% of the *Methanobrevibacter smithii* biofilm (Figure 5). However, it cannot completely ruled out that the surface coverage analysis of *Methanosarcina mazei* was affected by the potential EPS structures surrounding cells of *Methanosarcina mazei*, which were observed during SEM-analysis (Figure 3).

DISCUSSION

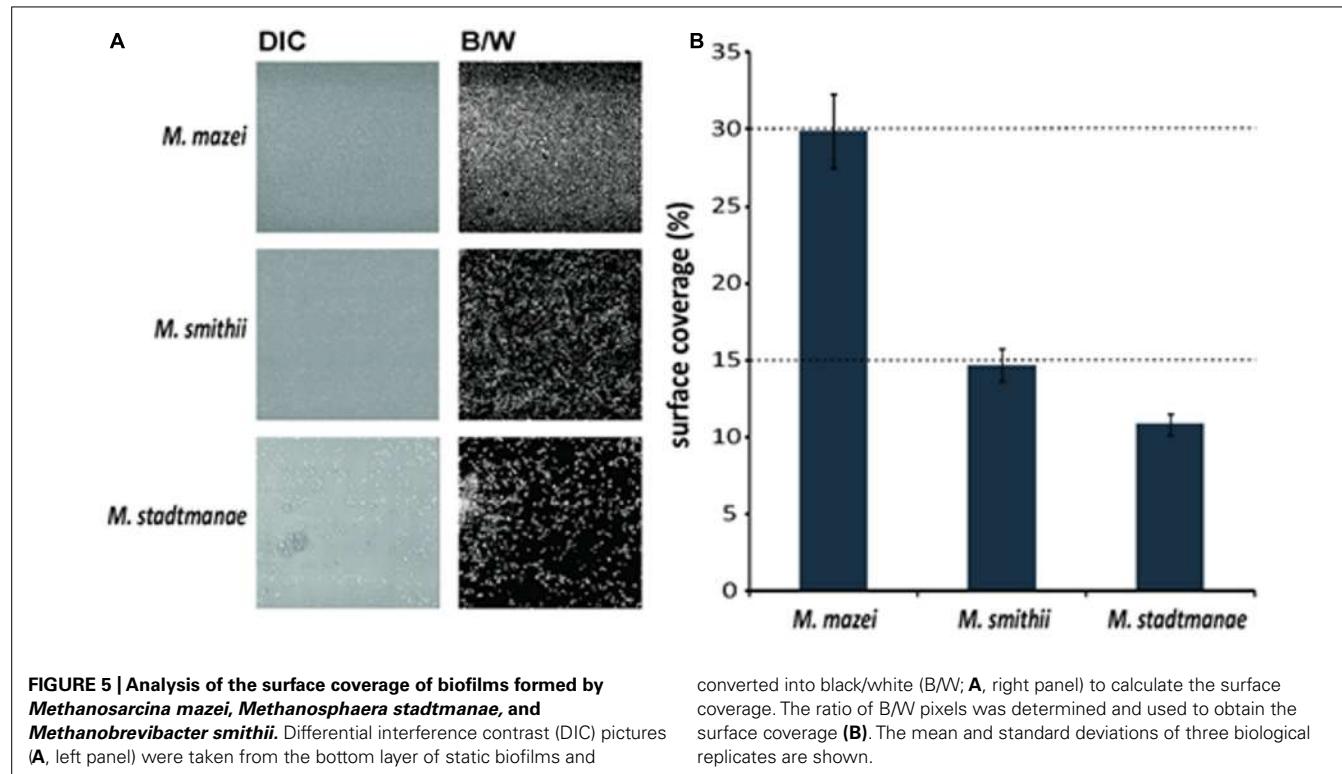
Although the knowledge on the functional importance of mucosal biofilms clearly increased in the last decade, the diversity

and characteristics of microbial communities associated with the human gut mucosa are still poorly understood (Dongari-Bagtzoglou, 2008). In addition, most studies dealing with the development and composition of human gut mucosal biofilms did only involve bacterial or fungal species (Swidsinski et al., 2005; Macfarlane and Dillon, 2007; Macfarlane et al., 2011). Thus, to our knowledge, this is the first report demonstrating biofilm formation of methanogenic archaea that frequently inhabit the human gut. By assessing static growth on different surfaces (mica and uncoated plastic μ -dishesTM) we showed that the studied methanoarchaeal strains, *Methanosarcina mazei*, *Methanobrevibacter smithii*, and *Methanospaera stadtmanae*, form biofilms with distinctive features. As it has been shown for other few archaeal species that form biofilms such as *Sulfolobus* spp. (Koerdt et al., 2010, 2011), the SM1 Euryarchaeon (Probst et al., 2013), several haloarchaeal strains (Fröls et al., 2012) and *Pyrococcus furiosus* as well as *Methanopyrus kandleri* (Schopf et al., 2008), each studied strain showed strain-specific characteristics during biofilm formation that were observed by using various microscopic techniques such as CLSM and SEM. In particular, significant differences in biofilm forming capabilities of the human gut inhabitants *Methanospaera stadtmanae* and *Methanobrevibacter smithii* were observed. In μ -dishesTM, *Methanobrevibacter smithii* biofilms reached heights up to 40 μ m, whereas *Methanospaera stadtmanae* biofilms grew only up to a



height of 2 μ m. However, surface coverage of *Methanospaera stadtmanae* (\sim 11%) was found to be almost similar to that obtained for *Methanobrevibacter smithii* (\sim 15%). Regarding to this, it has been shown in earlier studies that biofilm thickness and density increase with the number of participating microorganisms within the community (Costerton et al., 1995; Donlan, 2002). Thus, biofilm-forming communities consisting of both, bacteria and archaea, may reach significantly higher heights and

surface coverage as has been shown for various environmental biofilms (Orell et al., 2013). Furthermore, it has been demonstrated that bacterial human mucosal biofilm formation is favored in fluid flow or tissue motility such as the human gut (Tolker-Nielsen et al., 2000; Donlan, 2002; Dongari-Bagtzoglou, 2008). Hence, the determined static biofilm formation of methanarchaeal strains might underestimate their overall *in vivo* ability to form mucosal biofilms within the human gut. Interestingly, the



observed biofilm forming capabilities of the tested methanarchaeal strains differed within the two used systems. In particular, *Methanospaera stadtmanae*'s biofilm formation on mica plates appeared more pronounced when compared to the growth in μ -dishesTM. While the used mica plates are very smooth and hydrophilic, the surface of uncoated μ - dishesTM is more roughened and hydrophobic. Thus, surface properties are likely to influence the overall ability of methanarchaeal strains to form biofilms.

By using several lectins, only very low amounts of EPS were detected in these methanarchaeal biofilms (Figure 4). This observation might be due to the fact that the tested lectins did not exhibit the specificity needed to detect the secreted polysaccharides, since SEM analysis revealed high production of EPS for at least *Methanospaera stadtmanae* and *Methanobrevibacter smithii*. The tested lectin ConA mainly recognizes glucose and mannose residues, which form major components of EPS. However, the ConA signal was mainly co-localized with the DAPI stained cells; thus implying that the stained compound did not correspond to secreted exopolysaccharides, but most likely to the N-glycans that cover the outmost sheath of proteins or heteropolysaccharides surrounding the methanarchaeal cell surface (König, 1988, 2010; Kandler and König, 1998). In addition, the lectin IB4, specific for α -galactosyl residues, was rarely observed in all three biofilms. In this respect, further analysis is required to determine carbohydrate moieties of secreted EPS by methanarchaeal strains. On the other hand, high amounts of extracellular DNA (eDNA) have been observed in archaeal biofilms during earlier studies, particularly located in regions of sessile cell aggregates (Fröls et al., 2008; Koerdt et al., 2010; Orell et al., 2013). Hence, future studies should also

include examination of eDNA with an membrane-impermeable DNA-intercalating dye as well as detection of secretion proteins.

SEM analysis in this study revealed not only adhesion of methanarchaeal strains to the smooth mica surface, but also strong cell-to-cell adhesion of at least *Methanospaera stadtmanae* and *Methanobrevibacter smithii* during biofilm formation. The functional role of bacterial type-IV-pili-like structures and non-type-IV-pili-like structures involved by various archaeal species in biofilm formation has been confirmed in earlier studies (Fröls et al., 2008; Henche et al., 2012). However, the genomes of *Methanospaera stadtmanae* and *Methanobrevibacter smithii* lack coding sequences for archaellar or pili-like structures as well as for peptidases involved in processing pre-archaellins or pre-pilins indicating they cannot assemble an archaellum (archaeal flagellum) or type-IV-pili (Fricke et al., 2006; Samuel et al., 2007). Thus, adhesion of cells to the smooth surface of mica plates might also occur via interactions of either the heteropolysaccharide layer surrounding the cells of these two strains or by attachment of unknown cell appendages. Besides, under various stress conditions such as the treatment with human-derived antimicrobial peptides, alterations of the cell wall structure and increased cell aggregation of *Methanospaera stadtmanae* were observed in an earlier study (Bang et al., 2012). Furthermore, investigations of *Methanobrevibacter smithii* fecal strains as well as of *Methanospaera stadtmanae* revealed genomic adaptations to the human gut ecosystem such as the production of surface glycans resembling those found in the gut mucosa and a regulated expression of adhesion-like proteins (ALPs) (Fricke et al., 2006; Samuel et al., 2007). The expression of *Methanobrevibacter smithii*'s ALPs was later shown to differ between studied strains and to depend on

the existing concentration of formate (Hansen et al., 2011). Since biofilm formation often occurs during strong variations in living conditions such as nutrient limitations (Donlan, 2002; Dongari-Bagtzoglou, 2008), it might also be possible that biofilm formation of methanarchaeal strains is induced under certain stress conditions involving differential gene expression of ALPs among others. In this respect, it has also been shown that *Methanosarcina mazei* strain S-6 establishes multicellular forms (lamina) under certain stress conditions, which is thought to occur in adaptation to environmental changes (Mayerhofer et al., 1992). Besides, in response to changing culture conditions *Methanosarcina mazei* is able to switch between growth as (sarcina)packages and single cells (Sowers and Gunsalus, 1988). Thus, for *Methanosarcina mazei* it is also likely that it diversifies its cellular growth under static growth conditions in order to form a biofilm.

In summary, the present study demonstrated for the first time that methanarchaeal strains inhabiting the human gut have the ability to build up biofilms under static conditions. Though focusing on the evaluation of biofilm formation on abiogenic substrates, strong evidence was obtained that *Methanospaera stadtmanae* and *Methanobrevibacter smithii* might occur as an additional microbial part of mucosal biofilms in the human gut. This is in agreement with previous studies that demonstrated the interaction of these methanarchaeal strains with bacterial gut commensals such as *Bacteroides* species (Samuel and Gordon, 2006; Samuel et al., 2007). Microbial communities that occur in biofilms on the mucosal surface are currently thought to be crucially involved in modulating the host's immune system, since they are closer to the epithelium compared to microorganisms in the lumen (Macfarlane and Dillon, 2007; Macfarlane et al., 2011). More importantly, mucosal biofilms have been shown to be associated with many human infectious diseases that are reviewed in (Dongari-Bagtzoglou, 2008). In particular, the composition and density of mucosa-associated biofilms have been shown to alter in individuals with IBD, hence revealing evidence for an impact of sessile communities to human's gut diseases (Swidsinski et al., 2005). Regarding to this, increased prevalence of *Methanospaera stadtmanae* was recently found in patients with IBD (Blais-Lecours et al., 2014). Moreover, we recently demonstrated severe activation of human innate immune responses after exposure to this methanarchaeal strain, which might implicate its contribution to pathological conditions in the human gut (Bang et al., 2014). Thus, the observation in the present study demonstrating biofilm formation of mucosa-associated methanarchaeal strains might be important for the influence of *Methanospaera stadtmanae* and *Methanobrevibacter smithii* on the immunomodulation within the human gut that needs to be further elucidated.

AUTHOR CONTRIBUTIONS

Corinna Bang, Claudia Ehlers, Alvaro Orell, Marlene Spinner, Sonja-Verena Albers and Ruth A. Schmitz designed the research, Corinna Bang, Claudia Ehlers, Alvaro Orell, Daniela Prasse, and Marlene Spinner performed the research, Corinna Bang, Claudia Ehlers, Alvaro Orell, Marlene Spinner, Stanislav N. Gorb, Sonja-Verena Albers and Ruth A. Schmitz analyzed the data, and Corinna Bang, Claudia Ehlers, Alvaro Orell, Marlene Spinner, Sonja-Verena Albers, and Ruth A. Schmitz wrote the paper.

ACKNOWLEDGMENTS

Corinna Bang was funded by the German research foundation (DFG,SCHM1051/11-1). Alvaro Orell and Sonja-Verena Albers received intramural funding from the Max Planck Society and the Collaborative Research Center 987 from the German research foundation (DFG).

REFERENCES

- Abreu, M. T., Fukata, M., and Arditi, M. (2005). TLR signaling in the gut in health and disease. *J. Immunol.* 174, 4453–4460. doi: 10.4049/jimmunol.174.8.4453
- An, D., and Parsek, M. R. (2007). The promise and peril of transcriptional profiling in biofilm communities. *Curr. Opin. Microbiol.* 10, 292–296. doi: 10.1016/j.mib.2007.05.011
- Artis, D. (2008). Epithelial-cell recognition of commensal bacteria and maintenance of immune homeostasis in the gut. *Nat. Rev. Immunol.* 8, 411–420. doi: 10.1038/nri2316
- Ashok, N., Warad, S., Singh, V. P., Chaudhari, H., Narayanan, A., and Rodrigues, J. (2013). Prevalence of archaea in chronic periodontitis patients in an Indian population. *Ind. J. Dent. Res.* 24, 289–293. doi: 10.4103/0970-9290.117988
- Backhed, F., Ley, R. E., Sonnenburg, J. L., Peterson, D. A., and Gordon, J. I. (2005). Host-bacterial mutualism in the human intestine. *Science* 307, 1915–1920. doi: 10.1126/science.1104816
- Bang, C., Schilhabel, A., Weidenbach, K., Kopp, A., Goldmann, T., Gutsmann, T., et al. (2012). Effects of antimicrobial peptides on methanogenic archaea. *Antimicrob. Agents Chemother.* 56, 4123–4130. doi: 10.1128/aac.00661-12
- Bang, C., Weidenbach, K., Gutsmann, T., Heine, H., and Schmitz, R. A. (2014). The intestinal archaea *Methanospaera stadtmanae* and *Methanobrevibacter smithii* activate human dendritic cells. *PLoS ONE* 9:e99411. doi: 10.1371/journal.pone.0099411
- Blais-Lecours, P., Marsolais, D., Cormier, Y., Berberi, M., Hache, C., Bourdages, R., et al. (2014). Increased prevalence of *Methanospaera stadtmanae* in inflammatory bowel diseases. *PLoS ONE* 9:e87734. doi: 10.1371/journal.pone.0087734
- Chaban, B., Ng, S. Y., and Jarrell, K. F. (2006). Archaeal habitats – from the extreme to the ordinary. *Can. J. Microbiol.* 52, 73–116. doi: 10.1139/w05-147
- Costerton, J. W., Lewandowski, Z., Caldwell, D. E., Korber, D. R., and Lappin-Scott, H. M. (1995). Microbial biofilms. *Annu. Rev. Microbiol.* 49, 711–745. doi: 10.1146/annurev.mi.49.100195.003431
- Deppenmeier, U., Johann, A., Hartsch, T., Merkl, R., Schmitz, R. A., Martinez-Arias, R., et al. (2002). The genome of *Methanosarcina mazei*: evidence for lateral gene transfer between bacteria and archaea. *J. Mol. Microbiol. Biotechnol.* 4, 453–461.
- Dongari-Bagtzoglou, A. (2008). Pathogenesis of mucosal biofilm infections: challenges and progress. *Expert Rev. Anti Infect. Ther.* 6, 201–208. doi: 10.1586/14787210.6.2.201
- Donlan, R. M. (2002). Biofilms: microbial life on surfaces. *Emerg. Infect. Dis.* 8, 881–890. doi: 10.3201/eid0809.020063
- Dridi, B. (2012). Laboratory tools for detection of archaea in humans. *Clin. Microbiol. Infect.* 18, 825–833. doi: 10.1111/j.1469-0691.2012.03952.x
- Dridi, B., Henry, M., El Khechine, A., Raoult, D., and Drancourt, M. (2009). High prevalence of *Methanobrevibacter smithii* and *Methanospaera stadtmanae* detected in the human gut using an improved DNA detection protocol. *PLoS ONE* 4:e7063. doi: 10.1371/journal.pone.0007063
- Eckburg, P. B., Bik, E. M., Bernstein, C. N., Purdom, E., Dethlefsen, L., Sargent, M., et al. (2005). Diversity of the human intestinal microbial flora. *Science* 308, 1635–1638. doi: 10.1126/science.1110591
- Ehlers, C., Grabbe, R., Veit, K., and Schmitz, R. A. (2002). Characterization of GlnK1 from *Methanosarcina mazei* strain Gö1: complementation of an *Escherichia coli* glnK mutant strain by GlnK1. *J. Bacteriol.* 184, 1028–1040. doi: 10.1128/jb.184.4.1028
- Flemming, H. C., and Wingender, J. (2010). The biofilm matrix. *Nat. Rev. Microbiol.* 8, 623–633. doi: 10.1038/nrmicro2415
- Fricke, W. F., Seedorf, H., Henne, A., Krüer, M., Liesegang, H., Hedderich, R., et al. (2006). The genome sequence of *Methanospaera stadtmanae* reveals why this human intestinal archaeon is restricted to methanol and H₂ for methane formation and ATP synthesis. *J. Bacteriol.* 188, 642–658. doi: 10.1128/jb.188.2.642-658.2006
- Fröls, S. (2013). Archaeal biofilms: widespread and complex. *Biochem. Soc. Trans.* 41, 393–398. doi: 10.1042/BST20120304

- Fröls, S., Ajon, M., Wagner, M., Teichmann, D., Zolghadr, B., Folea, M., et al. (2008). UV-inducible cellular aggregation of the hyperthermophilic archaeon Sulfolobus solfataricus is mediated by pili formation. *Mol. Microbiol.* 70, 938–952. doi: 10.1111/j.1365-2958.2008.06459.x
- Fröls, S., Dyall-Smith, M., and Pfeifer, F. (2012). Biofilm formation by haloarchaea. *Environ. Microbiol.* 14, 3159–3174. doi: 10.1111/j.1462-2920.2012.02895.x
- Hansen, E. E., Lozupone, C. A., Rey, F. E., Wu, M., Guruge, J. L., Narra, A., et al. (2011). Pan-genome of the dominant human gut-associated archaeon, *Methanobrevibacter smithii*, studied in twins. *Proc. Natl. Acad. Sci. U.S.A.* 108(Suppl. 1), 4599–4606. doi: 10.1073/pnas.1000071108
- Henche, A. L., Ghosh, A., Yu, X., Jeske, T., Egelman, E., and Albers, S. V. (2012). Structure and function of the adhesive type IV pilus of *Sulfolobus acidocaldarius*. *Environ. Microbiol.* 14, 3188–3202. doi: 10.1111/j.1462-2920.2012.02898.x
- Hill, D. A., and Artis, D. (2010). Intestinal bacteria and the regulation of immune cell homeostasis. *Annu. Rev. Immunol.* 28, 623–667. doi: 10.1146/annurev-immunol-030409-101330
- Hopkins, K. L., Patrick, L. E., and Ball, T. I. (2001). Imaging findings of perforative appendicitis: a pictorial review. *Pediatr. Radiol.* 31, 173–179. doi: 10.1007/s002470000386
- Kandler, O., and König, H. (1998). Cell wall polymers in Archaea (Archaeabacteria). *Cell. Mol. Life Sci.* 54, 305–308. doi: 10.1007/s000180050156
- Koerdt, A., Godeke, J., Berger, J., Thormann, K. M., and Albers, S. V. (2010). Crenarchaeal biofilm formation under extreme conditions. *PLoS ONE* 5:e14104. doi: 10.1371/journal.pone.0014104
- Koerdt, A., Orell, A., Pham, T. K., Mukherjee, J., Włodkowiak, A., Karunakaran, E., et al. (2011). Macromolecular fingerprinting of sulfolobus species in biofilm: a transcriptomic and proteomic approach combined with spectroscopic analysis. *J. Proteome Res.* 10, 4105–4119. doi: 10.1021/pr2003006
- König, H. (1988). Archaeobacterial cell envelopes. *Can. J. Microbiol.* 35:395–406. doi: 10.1139/m88-071
- König, H. (2010). *Prokaryotic Cell Wall Compounds*. Berlin: Springer. doi: 10.1007/978-3-642-05062-6
- Kulik, E. M., Sandmeier, H., Hinni, K., and Meyer, J. (2001). Identification of archaeal rDNA from subgingival dental plaque by PCR amplification and sequence analysis. *FEMS Microbiol. Lett.* 196, 129–133. doi: 10.1111/j.1574-6968.2001.tb10553.x
- Levitt, M. D., Furne, J. K., Kuskowski, M., and Ruddy, J. (2006). Stability of human methanogenic flora over 35 years and a review of insights obtained from breath methane measurements. *Clin. Gastroenterol. Hepatol.* 4, 123–129. doi: 10.1016/j.cgh.2005.11.006
- Ley, R. E., Peterson, D. A., and Gordon, J. I. (2006). Ecological and evolutionary forces shaping microbial diversity in the human intestine. *Cell* 124, 837–848. doi: 10.1016/j.cell.2006.02.017
- Lovley, D. R., Greening, R. C., and Ferry, J. G. (1984). Rapidly growing rumen methanogenic organism that synthesizes coenzyme M and has a high affinity for formate. *Appl. Environ. Microbiol.* 48, 81–87.
- Lozupone, C. A., Stombaugh, J. I., Gordon, J. I., Jansson, J. K., and Knight, R. (2012). Diversity, stability and resilience of the human gut microbiota. *Nature* 489, 220–230. doi: 10.1038/nature11550
- Macfarlane, S., and Dillon, J. F. (2007). Microbial biofilms in the human gastrointestinal tract. *J. Appl. Microbiol.* 102, 1187–1196. doi: 10.1111/j.1365-2627.2007.03287.x
- Macfarlane, S., Bahrami, B., and Macfarlane, G. T. (2011). Mucosal biofilm communities in the human intestinal tract. *Adv. Appl. Microbiol.* 75, 111–143. doi: 10.1016/B978-0-12-387046-9.00005-0
- Macpherson, A. J., and Harris, N. L. (2004). Interactions between commensal intestinal bacteria and the immune system. *Nat. Rev. Immunol.* 4, 478–485. doi: 10.1038/nri1373
- Mathur, R., Amichai, M., Chua, K. S., Mirocha, J., Barlow, G. M., and Pimentel, M. (2013). Methane and hydrogen positivity on breath test is associated with greater body mass index and body fat. *J. Clin. Endocrinol. Metab.* 98, E698–E702. doi: 10.1210/jc.2012-3144
- Mayerhofer, L. E., Macario, A. J., and Conway De Macario, E. (1992). Lamina, a novel multicellular form of *Methanosarcina mazei* S-6. *J. Bacteriol.* 174, 309–314.
- Miller, T. L., Weaver, G. A., and Wolin, M. J. (1984). Methanogens and anaerobes in a colon segment isolated from the normal fecal stream. *Appl. Environ. Microbiol.* 48, 449–450.
- Miller, T. L., and Wolin, M. J. (1985). *Methanospaera stadtmaniae* gen. nov., sp. nov.: a species that forms methane by reducing methanol with hydrogen. *Arch. Microbiol.* 141, 116–122. doi: 10.1007/BF00423270
- Miller, T. L., Wolin, M. J., Conway De Macario, E., and Macario, A. J. (1982). Isolation of *Methanobrevibacter smithii* from human feces. *Appl. Environ. Microbiol.* 43, 227–232.
- O'Hara, A. M., and Shanahan, F. (2006). The gut flora as a forgotten organ. *EMBO Rep.* 7, 688–693. doi: 10.1038/sj.embo.7400731
- Orell, A., Fröls, S., and Albers, S. V. (2013). Archaeal biofilms: the great unexplored. *Annu. Rev. Microbiol.* 67, 337–354. doi: 10.1146/annurev-micro-092412-155616
- Post, J. C., Stoodley, P., Hall-Stoodley, L., and Ehrlich, G. D. (2004). The role of biofilms in otolaryngologic infections. *Curr. Opin. Otolaryngol. Head Neck Surg.* 12, 185–190. doi: 10.1097/01.moo.0000124936.46948.6a
- Probst, A. J., Holman, H. Y., Desantis, T. Z., Andersen, G. L., Birarda, G., Bechtel, H. A., et al. (2013). Tackling the minority: sulfate-reducing bacteria in an archaea-dominated subsurface biofilm. *ISMEJ* 7, 635–651. doi: 10.1038/ismej.2012.133
- Samuel, B. S., and Gordon, J. I. (2006). A humanized gnotobiotic mouse model of host-archaeal-bacterial mutualism. *Proc. Natl. Acad. Sci. U.S.A.* 103, 10011–10016. doi: 10.1073/pnas.0602187103
- Samuel, B. S., Hansen, E. E., Manchester, J. K., Coutinho, P. M., Henrissat, B., Fulton, R., et al. (2007). Genomic and metabolic adaptations of *Methanobrevibacter smithii* to the human gut. *Proc. Natl. Acad. Sci. U.S.A.* 104, 10643–10648. doi: 10.1073/pnas.0704189104
- Samuel, B. S., Shaito, A., Motoike, T., Rey, F. E., Backhed, F., Manchester, J. K., et al. (2008). Effects of the gut microbiota on host adiposity are modulated by the short-chain fatty-acid binding G protein-coupled receptor, Gpr41. *Proc. Natl. Acad. Sci. U.S.A.* 105, 16767–16772. doi: 10.1073/pnas.0808567105
- Schopf, S., Wanner, G., Rachel, R., and Wirth, R. (2008). An archaeal bi-species biofilm formed by *Pyrococcus furiosus* and *Methanopyrus kandleri*. *Arch. Microbiol.* 190, 371–377. doi: 10.1007/s00203-008-0371-9
- Sowers, K. R., and Gunsalus, R. P. (1988). Adaptation for growth at various saline concentrations by the archaeabacterium *Methanosarcina thermophila*. *J. Bacteriol.* 170, 998–1002.
- Swidsinski, A., Weber, J., Loening-Baucke, V., Hale, L. P., and Lochs, H. (2005). Spatial organization and composition of the mucosal flora in patients with inflammatory bowel disease. *J. Clin. Microbiol.* 43, 3380–3389. doi: 10.1128/JCM.43.7.3380-3389.2005
- Tolker-Nielsen, T., Brinch, U. C., Ragas, P. C., Andersen, J. B., Jacobsen, C. S., and Molin, S. (2000). Development and dynamics of *Pseudomonas* sp. biofilms. *J. Bacteriol.* 182, 6482–6489. doi: 10.1128/JB.182.22.6482-6489.2000
- Vianna, M. E., Conrads, G., Gomes, B. P. F. A., and Horz, H. P. (2006). Identification and quantification of archaea involved in primary endodontic infections. *J. Clin. Microbiol.* 44, 1274–1282. doi: 10.1128/jcm.44.4.1274-1282.2006
- Weaver, G. A., Krause, J. A., Miller, T. L., and Wolin, M. J. (1986). Incidence of methanogenic bacteria in a sigmoidoscopy population: an association of methanogenic bacteria and diverticulosis. *Gut* 27, 698–704. doi: 10.1136/gut.27.6.698
- Whitman, W. B., Coleman, D. C., and Wiebe, W. J. (1998). Prokaryotes: the unseen majority. *Proc. Natl. Acad. Sci. U.S.A.* 95, 6578–6583. doi: 10.1073/pnas.95.12.6578

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

Received: 25 April 2014; accepted: 24 June 2014; published online: 08 July 2014.

Citation: Bang C, Ehlers C, Orell A, Prasse D, Spinner M, Gorb SN, Albers S-V and Schmitz RA (2014) Biofilm formation of mucosa-associated methanarchaeal strains. *Front. Microbiol.* 5:353. doi: 10.3389/fmicb.2014.00353

This article was submitted to Microbial Physiology and Metabolism, a section of the journal *Frontiers in Microbiology*.

Copyright © 2014 Bang, Ehlers, Orell, Prasse, Spinner, Gorb, Albers and Schmitz. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

The universal tree of life: an update

Patrick Forterre^{1,2*}

¹ Unité de Biologie Moléculaire du Gène chez les Extrémophiles, Département de Microbiologie, Institut Pasteur, Paris, France, ² Institut de Biologie Intégrative de la cellule, Université Paris-Saclay, Paris, France

OPEN ACCESS

Edited by:

Mechthild Pohlschroder,
University of Pennsylvania, USA

Reviewed by:

Dong-Woo Lee,
Kyungpook National University, South
Korea

Reinhard Rachel,
University of Regensburg,
Germany

David Penny,
Massey University, New Zealand

***Correspondence:**

Patrick Forterre,
Unité de Biologie Moléculaire du
Gène chez les Extrémophiles,
Département de Microbiologie,
Institut Pasteur, 25 Rue du Docteur
Roux, Paris 75015, France
forterre@pasteur.fr

Specialty section:

This article was submitted to
Microbial Physiology and Metabolism,
a section of the journal
Frontiers in Microbiology

Received: 05 January 2015

Accepted: 29 June 2015

Published: 21 July 2015

Citation:

Forterre P (2015) The universal tree
of life: an update.
Front. Microbiol. 6:717.
doi: 10.3389/fmicb.2015.00717

Biologists used to draw schematic “universal” trees of life as metaphors illustrating the history of life. It is indeed a priori possible to construct an organismal tree connecting the three major domains of ribosome encoding organisms: Archaea, Bacteria and Eukarya, since they originated by cell division from LUCA. Several universal trees based on ribosomal RNA sequence comparisons proposed at the end of the last century are still widely used, although some of their main features have been challenged by subsequent analyses. Several authors have proposed to replace the traditional universal tree with a ring of life, whereas others have proposed more recently to include viruses as new domains. These proposals are misleading, suggesting that endosymbiosis can modify the shape of a tree or that viruses originated from the last universal common ancestor (LUCA). I propose here an updated version of Woese’s universal tree that includes several rootings for each domain and internal branching within domains that are supported by recent phylogenomic analyses of domain specific proteins. The tree is rooted between Bacteria and Arkarya, a new name proposed for the clade grouping Archaea and Eukarya. A consensus version, in which each of the three domains is unrooted, and a version in which eukaryotes emerged within archaea are also presented. This last scenario assumes the transformation of a modern domain into another, a controversial evolutionary pathway. Viruses are not indicated in these trees but are intrinsically present because they infect the tree from its roots to its leaves. Finally, I present a detailed tree of the domain Archaea, proposing the sub-phylum neo-Euryarchaeota for the monophyletic group of euryarchaeota containing DNA gyrase. These trees, that will be easily updated as new data become available, could be useful to discuss controversial scenarios regarding early life evolution.

Keywords: archaea, bacteria, eukarya, LUCA, universal tree, evolution

Introduction

The editors of research topic on “archaeal cell envelopes and surface structures” gave me the challenging task of drawing an updated version of the universal tree of life. This is a daunting task indeed, given that the concept of a universal “tree” is disputed by some scientists, who have suggested replacing trees with networks, and that major features of the tree are still controversial (Gribaldo et al., 2010; Forterre, 2012). Thus, it will be difficult to draw a consensus tree welcomed by all scientists in the field. In this paper, I thus try to propose updated versions of the universal tree that include as many features as possible validated by robust phylogenetic analyses and/or comparative molecular biology and biochemistry. I will draw a “universal tree” limited to ribosome-encoding organisms (Raoult and Forterre, 2008) that diverged from the last universal common ancestor (LUCA). Viruses (capsid encoding organisms) are polyphyletic, therefore their evolution can be neither illustrated by a single tree nor included in the universal tree as additional domains

(Forterre et al., 2014a). However, this should not be viewed as neglecting the role of viruses in biological evolution because “*the tree of life is infected by viruses from the root to the leaves*” (Forterre et al., 2014a). The universal tree of ribosome-encoding organisms contains cellular organisms that, unlike viruses, reproduce via cell cycles that imply the formation of new cells from the division of mother cells. This implies a fascinating continuity in the heredity of the cell membrane from LUCA to modern members of the three domains. A robust universal tree is critical to make sense of the evolution of the several types of lipids, cell envelopes and surface structures that originated and evolved on top of this continuity.

The Textbook Trees

Several popular drawings of the universal tree of life are widespread in the scientific literature and textbooks. These include the radial rRNA tree of Pace (1997), the tree of Stetter (1996) rooted in a hyperthermophilic LUCA and the most famous tree, which was published by Woese et al. (1990) to support their proposal to change the name “Archaeabacteria” to Archaea and to define the Domain as the highest taxonomic level (Woese et al., 1990; Sapp, 2009). All these trees are based on ribosomal DNA (rDNA) sequence comparisons and are rooted between Bacteria and a common ancestor of Archaea and Eukarya, a rooting that was first suggested by phylogenetic analyses of paralogous proteins (Iwabe et al., 1989; Gogarten et al., 1989).

The rDNA trees are still widely used despite their age (from 17 to 25 years old) because scientists need metaphors to represent the history of living organisms (despite all criticisms to the tree concept itself) and because few new trees have been proposed in the past two decades that are accepted by most biologists. Doolittle (2000) published a widely popularized tree in which the bases of the three domains are mixed by widespread lateral gene transfer (LGT). However, studies on archaeal phylogeny and more recently in Bacteria and Eukarya, as well as comparative genomic analyses, have shown that the history of organisms (not to be confused with the history of genes) can be inferred from a core of highly conserved genes (Brochier et al., 2005a; Gribaldo and Brochier, 2009; Puigbò et al., 2013; He et al., 2014; Ramulu et al., 2014; Raymann et al., 2014; Petitjean et al., 2015).

Comparative genomics has confirmed the existence of three versions (*sensu* Woese) of all universally conserved proteins, validating the three domains concept at the genomic level and opening the way to protein based universal trees (Olsen and Woese, 1997). A fairly popular radial tree based on a set of universal proteins was published in 2006 by Bork and colleagues (Ciccarelli et al., 2006). Unfortunately, this tree is biased by the over-representation of bacteria, because of the method used to create it, which requires complete genome sequences. Furthermore, several detailed internal branches within each domain are either controversial, such as the presence of Chlamydiae and Planctomycetes in different bacterial superphyla (Kamke et al., 2014). A well-thought-out universal tree based on rRNA was published in by López-García and Moreira (2008). This unrooted tree depicts each domain as a radial form with many phyla, without resolving the relationships between phyla within domains. However, like all previous trees, this tree does not show

some major lineages identified in the past decade, such as the Thaumarchaeota, which are now recognized as one of the three major archaeal phyla (Brochier-Armanet et al., 2008a; Spang et al., 2010, 2015).

Problems with Textbook Trees

The universal trees of the 1990s based on rDNA that are still widely used in textbooks, reviews and seminars provide a misleading view of the history of organisms. For instance, they all depict the division of Eukarya between a crown including Plants, Metazoa, Fungi, and several lineages of protists, and several basal long branches leading to various other unicellular eukaryotes, of which the most basal are protists lacking mitochondria (formerly called Archaezoa). This topology of the eukaryotic tree was very popular in the 1990s but is the result of a long branch attraction artifact. At the beginning of this century, it was acknowledged that all major eukaryotic divisions should be somewhere in the crown (Embley and Hirt, 1998; Keeling and McFadden, 1998; Philippe and Adoutte, 1998; Gribaldo and Philippe, 2002).

Another problem still present in many textbook trees is the position of hyperthermophiles. All rDNA trees of the 1990s were rooted within hyperthermophilic archaea and bacteria (Woese et al., 1990; Stetter, 1996; Pace, 1997). In particular, the hyperthermophilic bacteria of the genera *Thermotoga* and *Aquifex* were the two most basal bacterial lineages in all these trees. This explains why these bacteria are still often labeled as “deep-branching bacteria” (Braakman and Smith, 2014). However, the analysis of ribosomal RNA sequences at slowly evolving nucleotide positions (Brochier and Philippe, 2002) and phylogenetic analyses based on protein sequences do not support the deep branching of *Thermotoga* and *Aquifex* in the bacterial tree (Boussau et al., 2008b; Zhaxybayeva et al., 2009). The exact position of these hyperthermophilic bacteria currently remains controversial, because of the unusual extent of LGT that occurred between these bacteria and some other bacterial groups (Boussau et al., 2008b; Zhaxybayeva et al., 2009; Eveleigh et al., 2013).

The clustering of hyperthermophiles around the root and their “short branches” have been widely cited as support for the idea of a hot LUCA and a hot origin of life (Stetter, 1996), although the phenotype of an organism at the tip of a branch does not necessarily reflect that of its ancestor at the base. However, early reports noted that both features could be explained by the very high GC content of the ribosomal RNA of hyperthermophiles that limits the sequence space available for the evolution of these molecules (Forterre, 1996; Galtier et al., 1999; Boussau et al., 2008a). Indeed, the reconstruction of ribosomal RNA and protein sequences in LUCA shed serious doubt on its hyperthermophilic nature, and suggests instead that it was either a mesophilic or a moderate thermophilic organism (Galtier et al., 1999; Boussau et al., 2008a). This result is in agreement with the facts that specific thermostadaptation features of lipids in Archaea and Bacteria are not homologous and that reverse gyrase, a protein required for life at very high temperature, was probably not present in the common ancestor of Archaea and Bacteria (Forterre, 1996; Brochier-Armanet and Forterre, 2007; Glansdorff et al., 2008). These observations suggest that thermal adaptation from LUCA

to the ancestors of Archaea and Bacteria took place from cold to hot and not the other way around.

A Tree or a Ring?

Several authors during the past three decades have proposed to replace the universal tree with a “ring of life” (*sensu* Rivera and Lake, 2004) because they think that Eukarya originated from the association of a bacterium with an archaeon (for recent reviews, see Forterre, 2011; Martijn and Ettema, 2013). The most recent version of ring of life scenario is that eukaryogenesis was triggered by the engulfment of an alpha-proteobacterium by a wall-less giant archaeon capable of phagocytosis (Martijn and Ettema, 2013). Proponents of fusion (association) scenarios argue that such fusion is required to explain why eukaryotic genomes contain both archaeal and bacterial-like genes. However, this is not the case, since the presence of archaeal-like genes in Eukarya is a logical consequence of the sisterhood of Archaea and Eukarya, whereas the presence of bacterial-like genes is the expected result of mitochondrial endosymbiosis. Additional bacterial genes might have been introduced in proto-eukaryotes by LGT (Doolittle, 1998), which may have been partly mediated by large DNA viruses (Forterre, 2013a). Finally, ring of life scenarios do not easily explain the presence of many core eukaryotic genes (around 40%) that were already present in the last eukaryotic common ancestor (LECA) but have no detectable bacterial or archaeal homologs (Fritz-Laylin et al., 2010).

Ring of life scenarios, as well as scenarios in which Eukarya emerged from within Archaea (see below) assume the transformation of one and/or two of the modern domain into a third one. These scenarios have been criticized by several authors, as being biologically unsound (Woese, 2000; Kurland et al., 2006; De Duve, 2007; Cavalier-Smith, 2010; Forterre, 2011, 2013a). In particular, Woese (2000) argued that: “*modern cells are sufficiently integrated and “individualized” that further change in their designs does not appear possible.*” However, even if eukaryogenesis was actually triggered by the endosymbiotic event that produced mitochondria, this would not be a good reason to replace the tree of life with a ring. The universal tree should depict evolutionary relationships between domains defined according to the translation apparatus reflecting the history of cells (and their envelope; Woese et al., 1990) and not according to the global genomic composition that is influenced by LGT, virus integration and endosymbiosis, the history of which is incredibly complex.

This is well illustrated by the case of Plantae. The endosymbiosis of a *Cyanobacterium* that created this eukaryotic megagroup don’t prevent evolutionists to draw a tree of Eukarya in which Plantae are represented as one branch of the tree, and not as the product of a ring (He et al., 2014). The tree of any particular taxonomic unit is indeed not affected by the presence (or absence) of endosymbionts in some of its branches! Thus, a universal tree of life depicting the three domains as three separate entities does not contradict the fusion/endosymbiotic hypotheses at the origin of Eukarya, as long as this event had no effect on the eukaryotic ribosome itself. This is not the case, because the eukaryotic ribosome is not a mixture of archaeal and bacterial ribosomes; it shares 33 proteins with archaeal ribosomes that are

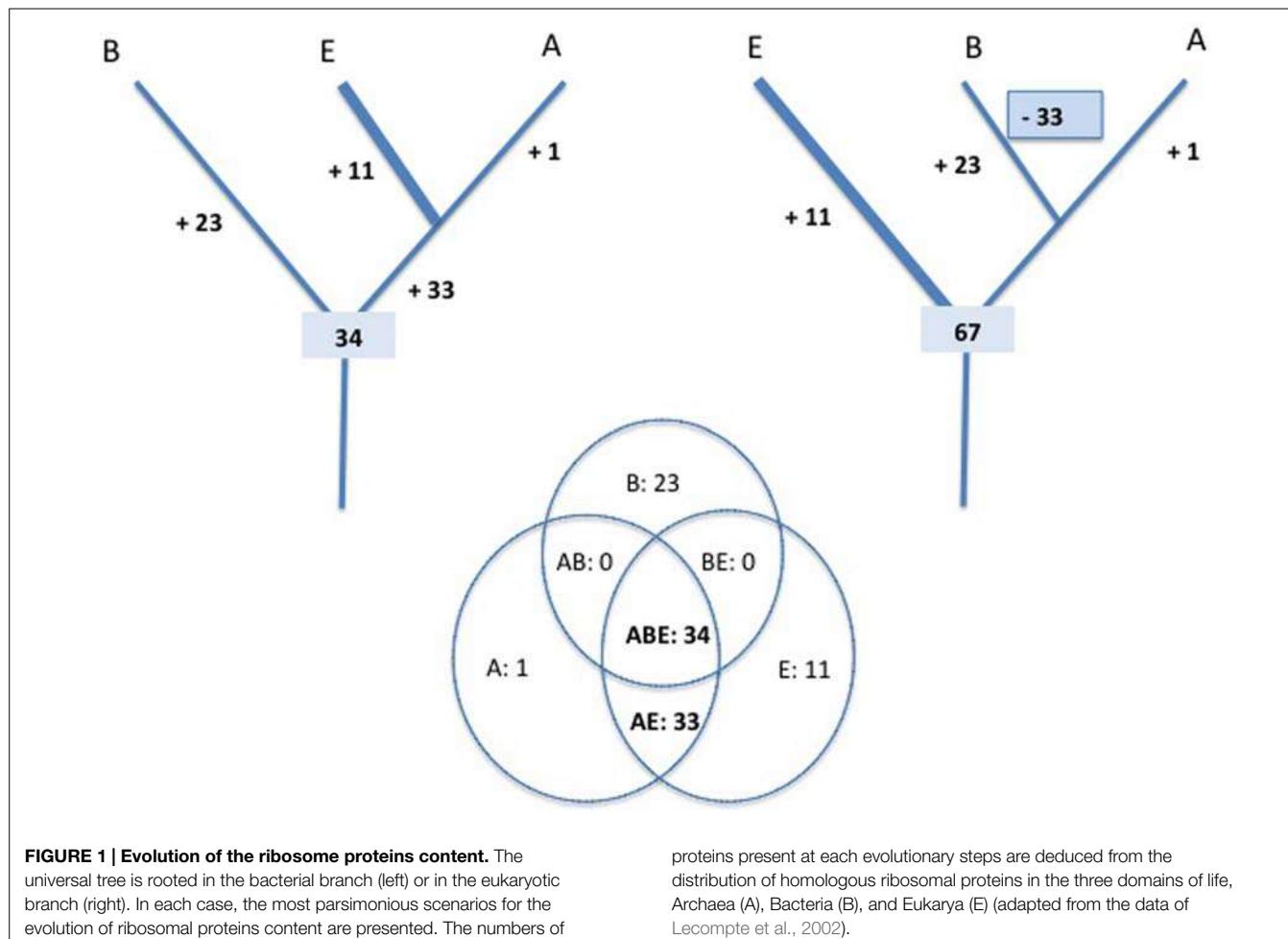
not present in Bacteria, but none with the bacterial ribosome that are not present in Archaea (Lecompte et al., 2002; **Figure 1**).

The “eocyte” Question

In the 1980s, James Lake proposed a universal tree in which Eukarya are sister group of a subgroup of Archaea (later recognized as Crenarchaeota) that he called Eocytes (Lake et al., 1984; Sapp, 2009). Most phylogeneticists now support a new version of the “Eocyte tree,” in which Eukarya emerge from within Archaea, and are a sister group of a superphylum encompassing Thaumarchaeota, Aigarchaeota, Crenarchaeota, and Korarchaeota (the TACK superphylum; Guy and Ettema, 2011; Williams et al., 2013; McInerney et al., 2014; Raymann et al., 2015), a sister group of Thaumarchaeota (Katz and Grant, 2014) or a sister group of Korarchaeota (Raymann et al., 2015). In a recent study, Eukarya are even proposed to have emerge from within a new phylum of the TACK superphylum (tentative phylum Lokiarchaeota; Spang et al., 2015).

The “eocyte” scenario is supported by phylogenetic analyses of universal proteins that use sophisticated methods for tree reconstruction, which are thought to be very efficient at identifying weak phylogenetic signals. However, these data are controversial, because most universal proteins are small (e.g., ribosomal proteins) and very divergent between Bacteria and Archaea/Eukarya, which makes archaeal/eukaryal relationships difficult to resolve (Gribaldo et al., 2010). For instance, the elongation factor datasets are saturated and unable to identify deep phylogenetic relationships between eukaryal phyla (Philippe and Forterre, 1999), and it is therefore challenging to use them as phylogenetic markers to resolve even deeper evolutionary relationships. In fact, in the single trees obtained for universal proteins by Cox et al. (2008), Eukarya branch within Euryarchaeota in about half of the trees and within Crenarchaeota in the other half, and they are characterized by poor node resolution and many aberrant groupings within Archaea (Cox et al., 2008, supporting information online). Similar unresolved and contradictory single-gene trees were again obtained by Williams and Embley in their more recent universal phylogeny (see supplementary Figure S1 in Williams and Embley, 2014).

One should be cautious in the interpretation of trees obtained from the concatenation of protein sequences that produce such contradictory individual trees. Indeed, the Microsporidia *Encephalitozoon*, a derived fungus, appears at the base of the eukaryotic tree published by Cox et al. (2008). This is reminiscent of the phylogenies of the 1990s that misplaced Microsporidia and other amitochondriate eukaryotes at the base of the tree of Eukarya (Hashimoto et al., 1995; Kamaishi et al., 1996). Similarly, several long basal eukaryotic branches (Fornicata, Archamoeba) emerged between Thaumarchaeota and other Eukarya in the tree of Katz and Grant (2014). In the tree supporting the emergence of Eukaryotes from Lokiarchaeota, the archaeal tree is rooted in the branch leading to *Methanopyrus kandleri* (Spang et al., 2015) a fast evolving archaeon (Brochier et al., 2004). Finally, in the analysis of Gribaldo and coworkers, the archaeal tree is rooted between Euryarchaeota and the putative TACK superphylum



using eukaryotic proteins as outgroup, but it is rooted in the branch leading to Korarchaeota when universal bacterial proteins are added to the dataset (Raymann et al., 2015).

The universal tree published by Gribaldo and co-workers reflects our best present knowledge of the internal branching order within Archaea and recovers the monophyly of most phyla in the three domains (Raymann et al., 2015). Notably, Archaea are rooted in this tree within Euryarchaeota when bacterial proteins are used as an outgroup, suggesting a new root for Archaea. However, Moreira and colleagues found instead that the root of the archaeal tree is located between Euryarchaeota and Crenarchaeota when using a bacterial outgroup (Petitjean et al., 2014).

I previously reported an observation that questions the validity of the methods used for tree reconstruction in some of these analyses. The three domains are each monophyletic, with well resolved evolutionary relationships within domains, in a tree of the universal protein Kae1/YgjD published in 2007 (see Figure S1 in Hecker et al., 2007). In contrast, Archaea are paraphyletic for the same protein in the analysis of Cox et al. (2008) (see Figure 2 in Forterre, 2013a) or in a more recent analysis by the same research group (Williams and Embley, 2014, supporting information online). In the tree of Cox et al. (2008),

Eukarya are a sister group of a clade containing crenarchaea and the euryarchaeon *Methanopyrus kandleri*, whereas in the tree of Williams and Embley, eukaryotic Kae1 emerges from a clade grouping Methanobacteriales and Methanococccales. This illustrates the importance to present in supplementary material the individual trees of universal proteins beside those obtained with concatenation of protein sequences.

The various sets of universal proteins used by different groups to investigate the relationships between Archaea and Eukarya show substantial overlap and it is probable that most protein data sets lack valid phylogenetic signal (Gribaldo et al., 2010). Two groups that analyzed similar sets of proteins with various methods came to a similar conclusion (Lasek-Nesselquist and Gogarten, 2013; Rochette et al., 2014). Lasek-Nesselquist and Gogarten (2013) noticed that “*the methods used*” to recover the eocyte tree “*generate trees with known defects...revealing that it is still error prone*,” whereas Rochette et al. (2014) concluded that “*the high frequency of paraphyletic-Archaea topologies for near-universal genes may be the consequence of stochastic effects*.”

Generally speaking, it is very difficult to resolve ancient relationships by molecular phylogenetic methods for both practical and theoretical reasons, essentially because the informative signal is completely erased at long evolutionary distances

(Forterre and Philippe, 1999; Mossel, 2003; Penny et al., 2014). One possibility to bypass this phylogenetic impasse is to focus on biological plausibility. Trees in which the three domains are each monophyletic are more plausible than trees in which Archaea are monophyletic because they explain more easily the existence of three versions of the ribosome discovered by Woese et al. (1990). In the classical Woese tree, these versions emerged from ancestors that differed from modern cells, at a time when the tempo of protein evolution was faster than today. By contrast, scenarios in which Eukarya emerged from within Archaea assume that members from a modern domain (Archaea) were transformed recently (after the diversification of this domain) into completely different organisms (Eukarya), something difficult to imagine from a biological point of view (Woese, 2000; Kurland et al., 2006; De Duve, 2007; Cavalier-Smith, 2010; Forterre, 2011, 2013a).

It has been proposed that this dramatic transformation (implying among others the replacement of archaeal type lipids by bacterial type lipids) was initiated in a particular archaeal lineage that, in contrast to all other lineages, already evolved toward more complex forms before eukaryogenesis (Martijn and Ettema, 2013). The recently proposed novel archaeal phylum, Lokiarchaeota, appears to be a good candidate in that case because its genome apparently encodes many genes potentially involved in the manipulation of membranes or in the formation of a cytoskeleton, including several homologs of eukaryal proteins that are observed for the first time in Archaea (Spang et al., 2015). Eukaryotic-like features present in Archaea (the archaeal eukaryome, *sensu* Koonin and Yutin, 2014) are indeed widely dispersed among the various archaeal phyla (Koonin and Yutin, 2014), suggesting that all these features were present in the last archaeal common ancestor (LACA), which was more complex than modern archaea (Forterre, 2013a). This observation is easily explained in the framework of the Woese tree by the selective loss of these features (present in the last common ancestor of Archaea and Eukarya) in different archaeal phyla during the streamlining process that led to the emergence of modern archaea (Forterre, 2013a). Comparative genomic analyses have indeed previously revealed a tendency toward reduction in the evolution of Archaea (Csurös and Miklos, 2009; Makarova et al., 2010; Koonin and Yutin, 2014).

In contrast, in the eocyte scenario, most eukaryotic features present in Archaea originated during the transition between Eurarchaeota and the TACK superphylum. This scenario further implies that all archaea “stopped evolving,” remaining archaea, whilst progressively and randomly losing some of these eukaryotic features, except for one particular lineage of Lokiarchaeota that experienced a dramatic burst of accelerated evolution and was transformed into eukaryotes.

It is traditionally suggested that the process that led to this transformation was triggered by the endosymbiosis event that created mitochondria (Lane and Martin, 2010). This seems to be a leap of faith, because there is no example of such a drastic transformation of the host molecular fabric at the more basic and fundamental levels (translation, transcription, replication) triggered by an endosymbiotic event (Forterre, 2013a). For instance, Plantae remain *bona fide* Eukarya (with typical eukaryotic version of all universal proteins) despite the

fact that about 20% of their genes originated from cyanobacteria (Martin et al., 2002).

An argument often used for scenarios in which Eukarya descended from modern lineages of prokaryotes is that Eukarya only appeared recently in the fossil record (McInerney et al., 2014). Most authors supporting this scenario systematically ignore the discovery 5 years ago of possible multicellular eukaryotes in sediments dating 2.1 billions years old (El Albani et al., 2010). This suggests that the last common ancestor of modern eukaryotes could have originated much earlier than previously thought (well before 2.1 Gyr ago) and that proto-eukaryotes could have been already present even earlier. In any case, this also confirms that the tempo of evolution of the central components of the molecular cell fabric has decreased around three Gyr ago, possibly after the emergence of the three domains, as first suggested by (Woese, 2000; Forterre, 2006). This also explains why it is so challenging to determine the precise topology of the universal tree of life, considering that we are dealing with six Gyr of evolution, encompassing periods with very different evolutionary tempo, when we are comparing two modern sequences of universal proteins.

The Elusive Root of the Tree

A major problem in drawing the universal tree of life is the position of the root. The tree is rooted between Bacteria and Archaea/Eukarya in the classical Woese tree (Woese et al., 1990). This rooting was initially supported by phylogenetic analyses of protein paralogs (elongation factors and V/F types ATPase subunits) that originated by duplication before LUCA (Gogarten et al., 1989; Iwabe et al., 1989). This rooting was criticized in the 1990s because the bacterial branches are much longer than the other two branches in the trees of these protein paralogs (Forterre and Philippe, 1999; Philippe and Forterre, 1999). Furthermore, the elongation factors and V/F type ATPase subunits data sets, as well as other groups of paralogs (e.g., signal recognition particles, SRP) that also used to place the root between Bacteria and a common ancestor of Archaea and Eukarya were saturated with mutations (Philippe and Forterre, 1999). Therefore, it was unclear if this rooting reflects the real history of life on our planet or if it is due to a long branch attraction artifact, e.g., the “bacterial branch” being attracted by the long branch of the outgroup sequences of the paralogs. Statistical analysis of slowly evolving positions in the two paralogous subunits of SRP confirmed that the bacterial rooting obtained by more classical phylogenetic analyses with SRP was due to a long branch attraction artifact and suggested that the root is located between Archaea/Bacteria and Eukarya (Brinkmann and Philippe, 1999); however, this analysis was not followed up.

As is the case for archaeal/eukaryal relationships, there is probably no valid phylogenetic signal left in the universal protein data set to resolve the rooting of the universal tree by molecular phylogeny. This was confirmed in the case of the elongation factors data set by a cladistic analysis of individual amino-acid alignments that discriminate between primitive and share derived characters (Forterre et al., 1992). Only 23 positions could be subjected to this analysis in the elongation factor data set, of

which 22 gave ambiguous results and only one supported bacterial rooting!

These past 20 years, the rooting problem has been neglected—with a few exceptions (see for instance Harish et al., 2013) that I have no space to discuss here. Indeed, “ring of life” scenarios or those in which Eukaryotes originated from Archaea automatically root the tree between Archaea and Bacteria (rejuvenating the pre-Woesien prokaryote/eukaryote paradigm). However, comparative molecular biology has now revealed several situations that can help us to root the universal tree and decide between alternative scenarios.

Rooting From Comparative Molecular Biology

Comparative genomic analyses have shown that most proteins central for cellular function (both informational and operational) show higher sequence similarity between Archaea and Eukarya than between Eukarya and Bacteria. Furthermore, Archaea and Eukarya also share many proteins that are either absent in Bacteria or replaced with non-homologous proteins with the same function. Surprisingly, comparative genomic analyses have also shown that critical components of the DNA replication machinery (replicase, primase, helicase) are non-homologous between Archaea/Eukarya and Bacteria (Leipe et al., 1999; Forterre, 2006). This is also the case for the proteins that allow the bacterial F-type ATPase and archaeal A-type ATPase to work as ATP synthases (Mulkidjanian et al., 2007). All these observations are difficult to interpret if the universal tree is rooted in the archaeal or eukaryotic branches and/or if the archaeal/eukaryal specific proteins were present in LUCA. Indeed, this would have required many non-orthologous replacement events that occurred specifically in the bacterial branch. **Figure 1** illustrates the case of the ribosome evolution. The eukaryotic (or archaeal) rooting is clearly less parsimonious than the bacterial ones since it implies the loss of 33 ancestral ribosomal proteins in the bacterial branch concomitant with the gain of 23 new proteins. Such non-orthologous replacement scenario cannot be completely ruled out since a similar gain and loss event occurred during the evolution of the mitochondrial ribosome from the bacterial one (Desmond et al., 2011). However, there is no evidence that the emergence of bacteria involved a dramatic evolutionary event similar to the drastic reductive evolution that occurred during the emergence of mitochondria.

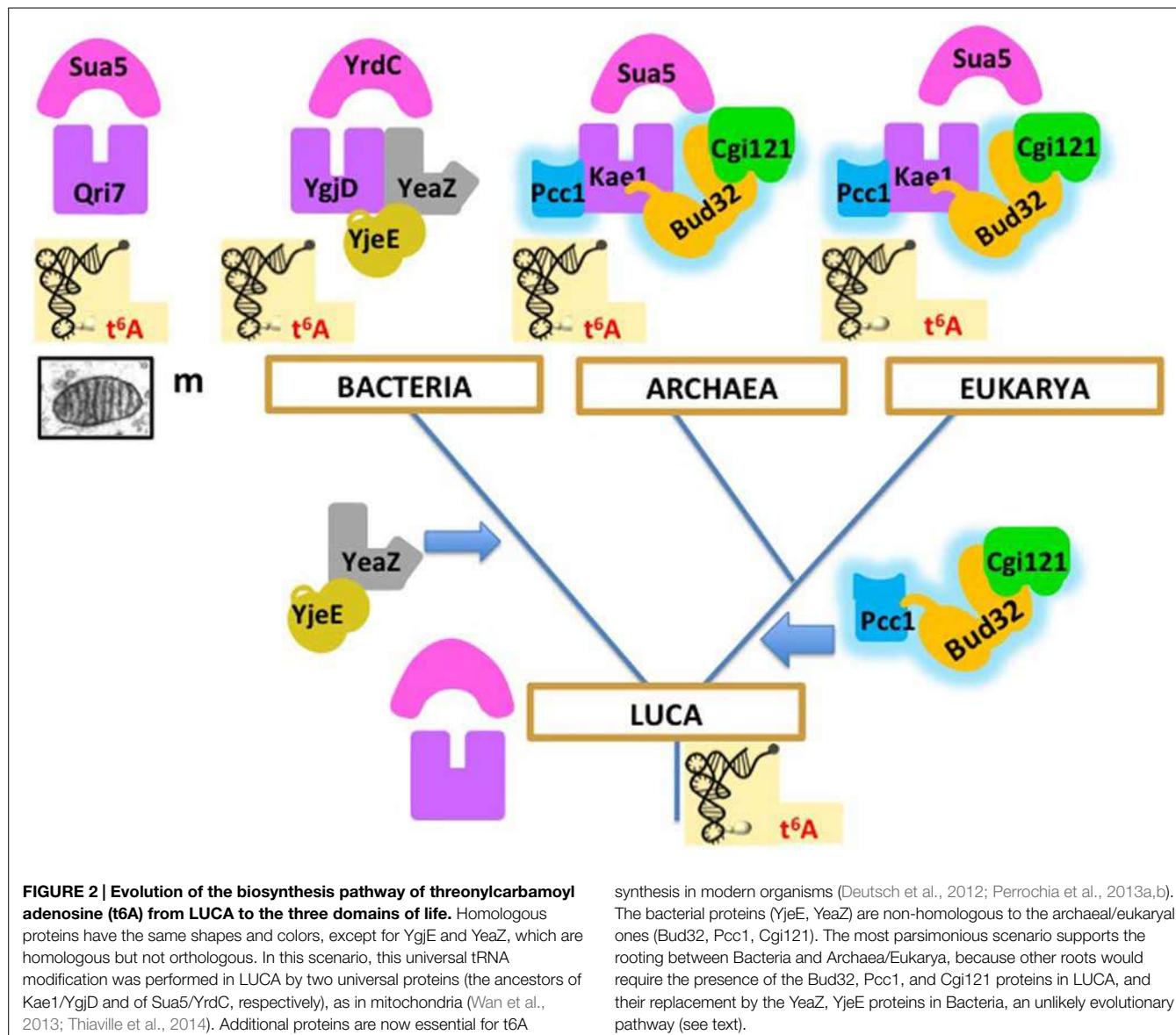
It was previously suggested that non-orthologous replacement had indeed occurred for the DNA replication machinery, with the ancestral DNA replication machinery in LUCA being replaced by non-homologous DNA replication proteins of viral origin either in Bacteria or in Archaea/Eukarya (Forterre, 1999). However, this type of explanation cannot be easily generalized. It seems unlikely that multiple non-orthologous replacements can explain all other major differences between archaeal/eukaryal and bacterial analogous but non-homologous systems! In the case of the DNA replication machineries, it is simpler to imagine that two versions present in modern cells were independently transferred from viruses to cells, once in the bacterial lineage and once in the archaeal/eukaryal lineage (Mushegian and Koonin, 1996;

Forterre, 2002, 2013b). For instance, our preliminary analyses of universal proteins sequence alignments indicate that the Lokiarchaeon is probably neither an early branching archaeon nor a missing link between Archaea and Eukarya (see also Nasir et al., 2015). Similarly, other analogous, but non-homologous, systems, such as the two distinct rotary proteins involved in ATP synthesis by F^o/F₁ and A/V ATPases, might have originated independently in the bacterial and in the archaeal/eukaryal lineages (Mulkidjanian et al., 2007).

Another explanation for the existence of non-homologous systems between Archaeal/Eukaryal and Bacteria is that LUCA contained two redundant systems and that one of them was later on lost at random in each domain (Edgell and Doolittle, 1997; Glansdorff et al., 2008). However, it is unlikely that both versions of all non-homologous systems between Archaeal/Eukaryal and Bacteria were present in LUCA. For instance, no modern cells have two non-homologous versions of DNA replication machineries or two versions of RNA polymerases (the bacterial and the archaeal ones). Some systems could have been randomly distributed between LUCA and other contemporary cellular (or viral) lineages, and redistributed thereafter by LGT, but this seems very unlikely in the case of the ribosome.

Recent biochemical work in our laboratory exemplifies why comparative biochemistry data support a universal tree in Archaea and Eukarya are indeed sister domains. Several research groups, as well as our team in Orsay, succeeded in reconstituting *in vitro* the protein complexes involved in the biosynthesis of the universal threonylcarbamoyl adenosine (t6A) tRNA modification in position 37 of tRNA in the three domains of life (Deutsch et al., 2012; Perrochia et al., 2013a,b) and in mitochondria (Wan et al., 2013; Thiaville et al., 2014). In Bacteria, Archaea and Eukarya, the reactions require the combination of two universal proteins and essential accessory proteins that exist in two versions, one present in Bacteria, the other present in Archaea and Eukarya. Interestingly, the same reaction can be performed in mitochondria by the two universal proteins alone, one (Qri7) that came from Bacteria via the endosymbiotic route and the other (Kae1) corresponding to the eukaryotic version (Wan et al., 2013; Thiaville et al., 2014). These results suggest that LUCA was able to perform this universally conserved reaction with the ancestors of the two universal proteins and that accessory proteins (now essential) were added independently in the bacterial and in the archaeal/eukaryal lineages. The most parsimonious scenario, illustrated in **Figure 2**, supports the rooting between Bacteria and Archaea/Eukarya, because other roots would require the presence of the archaeal/eukaryal set of accessory proteins in LUCA, and its replacement by the non-homologous bacterial set in Bacteria. This seems unlikely because biochemical analyses have shown that the bacterial and archaeal/eukaryal accessory proteins are not functionally equivalent (Deutsch et al., 2012; Perrochia et al., 2013b). It is therefore difficult to imagine intermediate steps in the replacement process. Furthermore, such replacement, even partial, never occurred during the diversification of the three domains.

Woese and Fox (1977) were thus possibly right when they proposed that the molecular fabric of LUCA was simpler than that of modern organisms, and that this organism still had an RNA



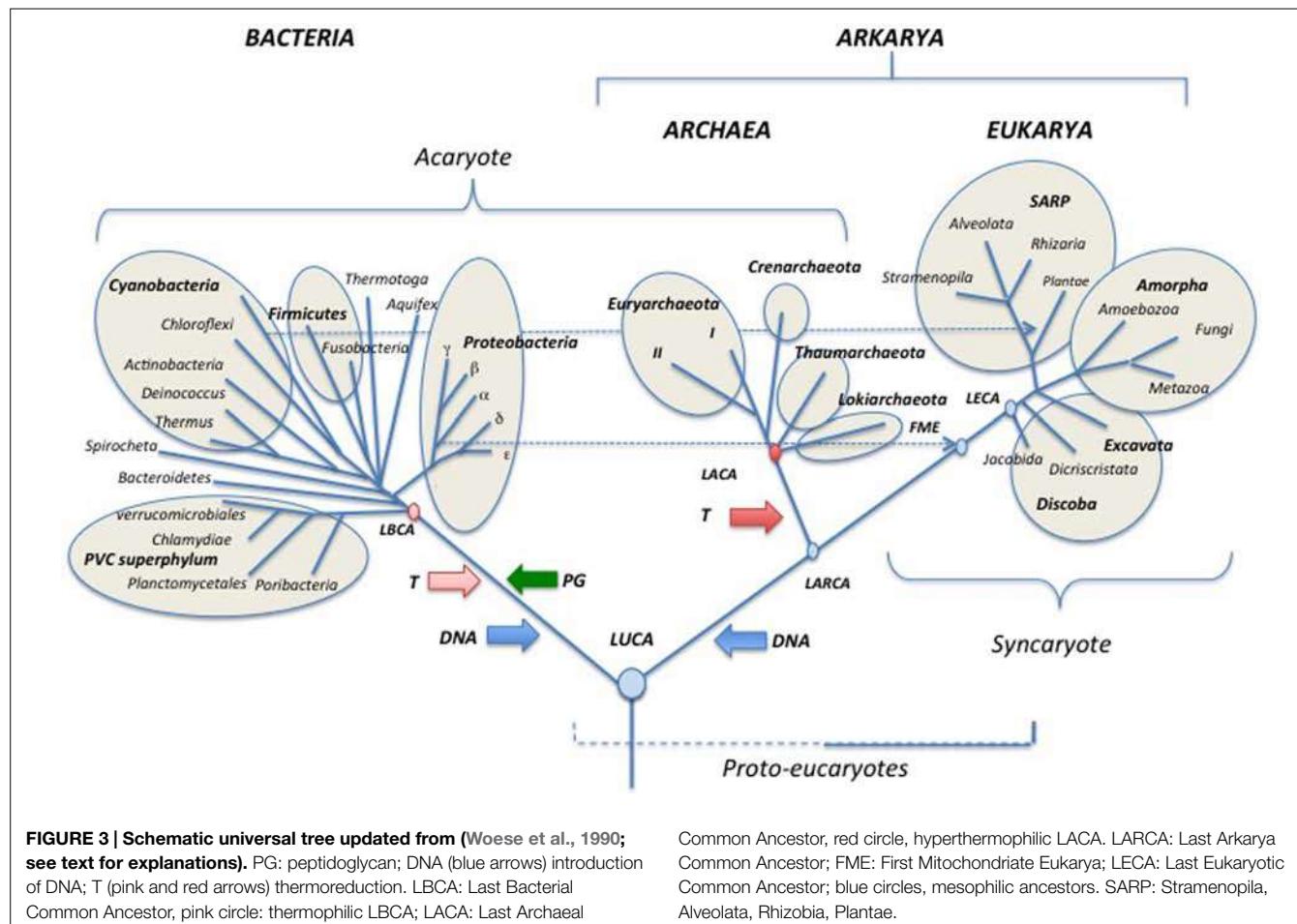
genome. In this scenario, major molecular machineries, such as the DNA replication machineries or the ATP synthases, emerged and/or became sophisticated independently in the branches leading to Bacteria on one side and to the common ancestor of Archaea and Eukarya on the other.

The rooting of the universal tree in the so-called “bacterial branch” (Figures 3–5) has been often interpreted as suggesting a “prokaryotic phenotype” for LUCA. This is a misleading interpretation that again confuses the phenotypes at the tip and base of a branch. The rooting between a lineage leading to Bacteria and a lineage leading to Archaea and Eukarya is compatible with diverse types of LUCA, including a LUCA with some “eukaryotic-like features” that were lost in Archaea and Bacteria (Forterre, 2013a).

Importantly, rooting of the universal tree in the “bacterial branch” formally requires giving a name to the clade grouping Archaea and Eukarya. Woese (2000) never proposed such a

name, adopting a “gradist” view of life evolution, with the three Domains emerging independently from a “communal LUCA” before the “Darwinian threshold” (Woese, 2000, 2002). In such view, the notion of clade itself cannot be used to group organisms that diverged at the time of LUCA when no real speciation occurred. I have criticized the Darwinian threshold concept, assuming—with many others—that Darwinian evolution started as soon as biological evolution take off (see Forterre, 2012, and references therein). In particular, extensive genes exchanges that possibly take place at the time of LUCA (but see Poole, 2009) cannot be opposed to Darwinian evolution occurring through variation and selection, since gene transfer only corresponds to a specific type of variation (Forterre, 2012).

I think that it’s time now to look back at the universal tree with a cladistics perspective and to propose a name for the clade grouping Archaea and Eukarya. It is challenging to find a common synapomorphy to Archaea and Eukarya that could provide a



good name for the clade corresponding to these two domains. David Prangishvili suggests Arkarya (personal communication), combining the names of the two domains belonging to this clade (**Figure 3**). Notably, universal trees in which Eukarya emerge from within Archaea are often viewed as “two domains trees” versus the “three domains tree” of Carl Woese (Gribaldo et al., 2010). However, the new nomenclature proposed here emphasizes that the classical Woese tree is also *stricto sensu*, a two domains tree (Bacteria and Arkarya)!

Updated Trees for Everybody

The backbone of the updated universal trees proposed here (**Figures 3–5**) was selected from the 1990 tree of Woese et al. (1990) as a tribute to Carl Woese and the historical work of the Urbana school (Sapp, 2009; Albers et al., 2013). The relative lengths of the branches linking the three domains together combine features of the rDNA and protein trees. It is indeed puzzling that Archaea and Eukaryotes are very close in trees based on universal protein sequence comparison (Rochette et al., 2014), but are more divergent in those based on rDNA (Pace et al., 1986). The reason for this discrepancy remains unclear and should be worth exploring further. The rather long branches between domains in the trees of **Figures 3–5** also reflect the “three

major transformation events” (*sensu* Forterre and Philippe, 1999) that occurred between LUCA and the formation of each domain (Woese, 1998; Forterre and Philippe, 1999).

Evidently, it is not possible to draw a tree including all presently recognized phyla, especially in the bacterial domain, so I made arbitrary choices, and tried to include most well studied bacterial and archaeal phyla, as well as major eukaryotic divisions and/or supergroups. In the case of Archaea, I only indicate the phyla Euryarchaeota, Crenarchaeota, Thaumarchaeota and the candidate phylum “Lokiarchaeota” because other proposed archaeal phyla are represented by a single species and/or their phylum status is controversial or has been refuted by robust phylogenetic and phylogenomic analyses (see below). Although still preliminary, the study of three partial lokiarchaeal genomes has shown that these archaea encode many eukaryotic-like genes absent in Thaumarchaeota and Crenarchaeota in phylogenetic analysis (Spang et al., 2015). Furthermore, Lokiarchaeota correspond to a large clade of abundant and diversified uncultivated archaea, previously named deep-sea archaeal group (DSAG), that are widely distributed in both marine and fresh water environments (Jørgensen et al., 2013).

I divide the phylum Euryarchaeota in sub-phylum I (I) and sub-phylum II (II), according to the presence/absence of

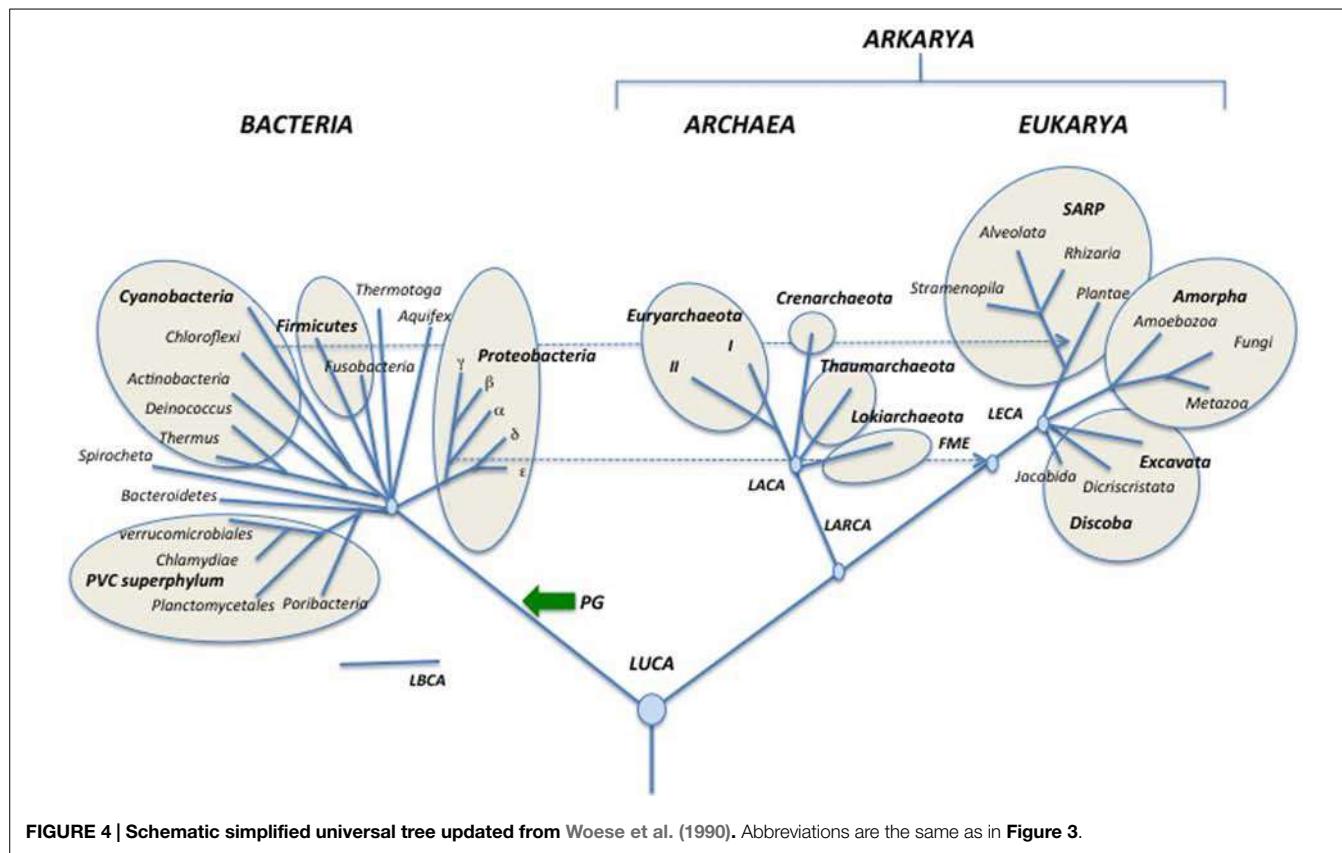


FIGURE 4 | Schematic simplified universal tree updated from Woese et al. (1990). Abbreviations are the same as in **Figure 3**.

DNA gyrase (see below; Forterre et al., 2014b). Dotted lines indicate the endosymbiosis events that had a major impact on the history of life by triggering the emergence of both modern eukaryotes (mitochondria) and Plantae (chloroplasts). In particular, this reminds us that the first mitochondrial eukaryote (FME) emerged after the diversification of alpha proteobacteria, indicating that “modern eukarya” are indeed much more recent than Archaea and Bacteria.

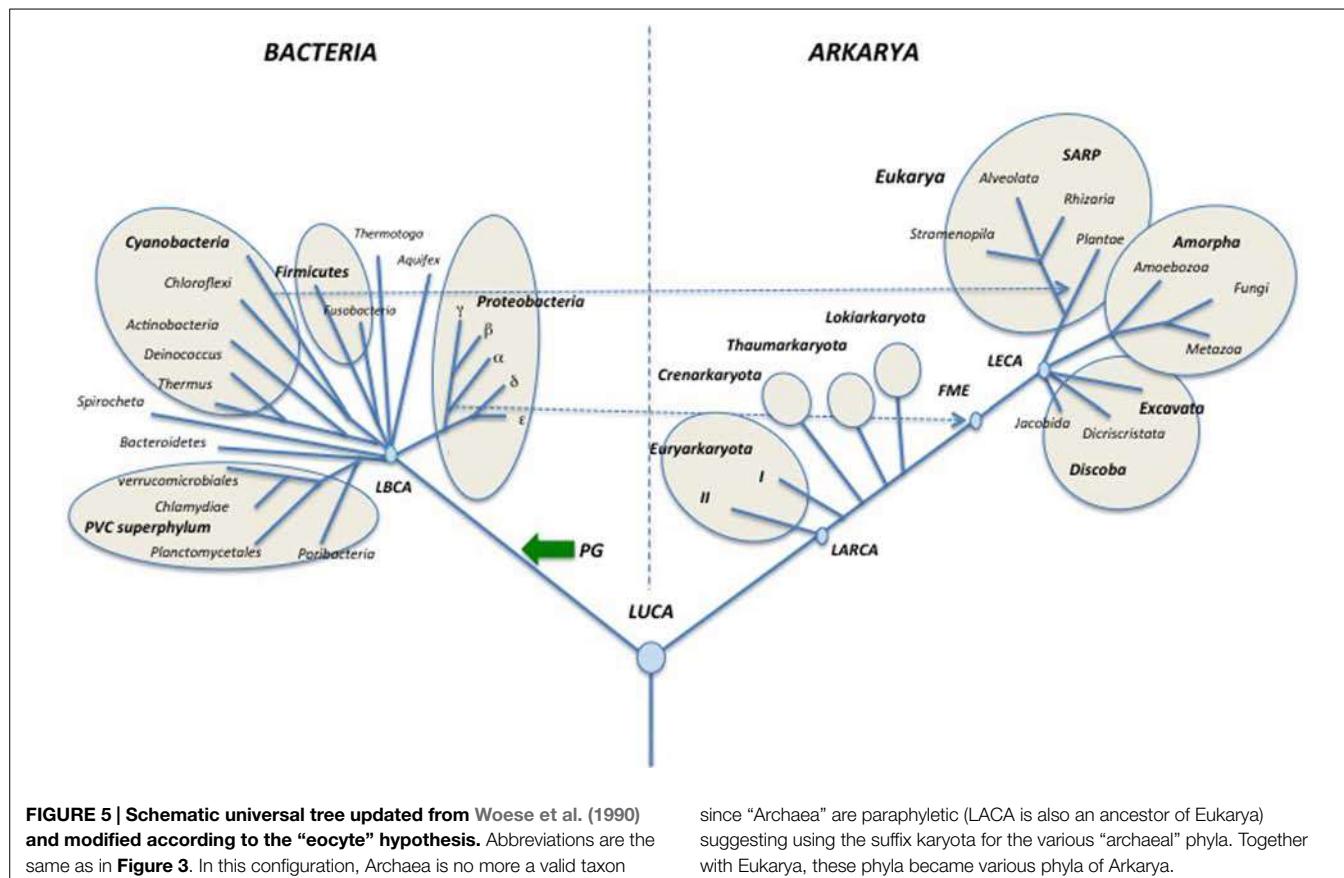
In the tree of **Figure 3**, I use the terms “synkaryote” and “akaryote” (with and without a nucleus, respectively) instead of eukaryotes and prokaryotes (Forterre, 1992; Harish et al., 2013; Penny et al., 2014). This is because the latter terms are the hallmark of the traditional (pre-Woesian) view of the evolution of life from primitive bacteria (“pro” karyotypes) to lower and finally higher eukaryotes (Forterre, 1992; Pace, 2006; Penny et al., 2014).

Some major events that shaped modern domains are indicated, such as the introduction of peptidoglycan (PG) in the lineage leading to Bacteria. The last bacterial and archaeal common ancestors (LBKA and LACA) are colored in pink and red, respectively, to indicate their probable thermophilic and hyperthermophilic nature based on ancestral protein and rRNA sequence reconstruction (Boussau et al., 2008a; Groussin and Gouy, 2011; Groussin et al., 2013). The grouping of hyperthermophiles at the base of the archaeal tree also suggests that LACA was a hyperthermophile (Brochier-Armanet et al., 2011; Petitjean et al., 2015), whereas LUCA was probably a mesophile or a moderate thermophile (Boussau et al., 2008a;

Groussin and Gouy, 2011). Some proposed events are more speculative but supported by theoretical arguments, such as the independent introduction of DNA (blue arrows) from viruses, into the lineages leading to Bacteria and to Archaea/Eukarya (Forterre, 2002) and the thermoreduction (red arrows) at the origin of the modern “akaryotic” phenotype (Forterre, 1995).

Rooting of the domain Eukarya and internal branching in this domain have been adapted from the recent tree of Baldauf and colleagues which is based on a concatenation of mitochondrial proteins rooted with their bacterial homologs (a rather close outgroup compared to Archaea; He et al., 2014). This tree is rooted between Discoba (Jakobida plus Discrititata) and all other eukaryotic megagroups. This rooting has been criticized by Derelle et al. (2015) who found distant paralogs in the data set used by He et al. (2014). These authors located the root of the eukaryotic tree between Amorpha and other eukaryotic groups in mitochondrial proteins trees. This rooting corresponds to the previous division between Unikonta and Bikonta originally proposed by (Stechmann and Cavalier-Smith, 2003; Cavalier-Smith, 2010). Derelle et al. (2015) have now suggested naming these two assemblages Optimoda (Amorpha in **Figures 3–5**) and Diphoda.

The position of the root of the domain Eukarya has been constantly changing with further phylogenetic analyses, raising doubt about the possibility of settling this issue using molecular phylogenetic methods based on protein sequences. I thus decided to root this domain between Jakobida and all other eukaryotes in



the universal tree of **Figure 3**, as recently suggested by Kannan et al. (2014), because Jakobida contain large mitochondrial genomes that still encode the bacterial RNA polymerase genes (Burger et al., 2013; Kamikawa et al., 2014; Kannan et al., 2014). The mitochondria of the LECA probably still had this RNA polymerase, which was subsequently replaced with a viral RNA polymerase in all Eukaryotes, except Jakobida (Kannan et al., 2014). These viral RNA polymerases have been recruited from a provirus integrated into the genome of the alpha-proteobacterium, which gave rise to mitochondria (Filée and Forterre, 2005). It seems unlikely that this non-orthologous replacement occurred twice independently in the history of mitochondria. Accordingly, the rooting between Jakobida and other eukaryotes is reasonable (more parsimonious) because it requires a single NOR of RNA polymerase in mitochondrial evolution, whereas the rooting between Amorpha and other eukaryotes would require several independent non-orthologous replacements.

Rooting of the domain Bacteria and internal branching in this domain have been adapted from the ribosomal protein trees of Koonin and colleagues (Yutin et al., 2012). These authors have suggested several superphyla beside the previously recognized PVC superphylum, which includes Planctomycetes, Verrucomicrobia and Chlamydiae (Kamke et al., 2014). These putative superphyla are indicated by circles in the tree of **Figure 3**. Branchings within Proteobacteria are drawn according to the ribosomal protein tree of Brochier-Armanet and colleagues

(Ramulu et al., 2014). The tree is tentatively rooted between the PVC superphylum and all other Bacteria, according to the basal rooting of Planctomycetes obtained by Brochier and Philippe using slowly evolving positions in ribosomal RNA sequences (Brochier and Philippe, 2002). Bacteroidetes are indicated in the second branching because these Bacteria are grouped with PVC bacteria in the phylogenetic analysis based on ribosomal proteins (Yutin et al., 2012).

The basal position of PVC in the bacterial tree of **Figure 3**, which remains to be confirmed, is appealing because the ancestor of PVC bacteria contained several genes encoding proteins structurally analogous to various eukaryotic coat proteins involved in vesicle and nuclear pore formation (Santarella-Mellwig et al., 2010). These proteins are probably involved in the invagination of the cytoplasmic membrane that led to the formation of the intracellular cytoplasmic membrane (ICM) in most PVC bacteria. This mimics the role of coat proteins in eukaryotes that are involved in the formation of the endoplasmic reticulum and nuclear membranes. The basal position of PVC bacteria suggests a parsimonious scenario in which these proteins were present in LUCA, and later lost in Archaea and most Bacteria (Forterre and Gribaldo, 2010). However, this scenario is still controversial since it is presently unclear whether the structurally analogous proteins of PVC Bacteria and Eukarya are also homologous (McInerney et al., 2011; Devos, 2012). These proteins are formed by the fusion of two domains (one rich in alpha helices, the other in beta strands) that are each present in

the three domains. Accordingly, they can also have originated independently by the fusion of these domains in the branches leading to Eukarya and PVC bacteria.

Archaea have been tentatively rooted in the branch leading to Lokiarchaeota in the tree of **Figure 3**, because this candidate phylum contains most eukaryotic features present in Archaea and branches closer to Eukarya than to other archaea in phylogenetic analyses of universal proteins, even when bacterial proteins are removed from the analysis (Spang et al., 2015, Figure S13D). Previously, the archaeal ribosomal tree was rooted in the branch leading to Thaumarchaeota when eukaryotic proteins were used as outgroup (Brochier-Armanet et al., 2008a). This rooting was also observed in a phylogenetic analysis of the archaeal replicative helicase MCM, which is a good phylogenetic marker for the archaeal domain (Krupovic et al., 2010), and in a phylogeny of five informational proteins present in deeply branching Thaumarchaeota from Kamchatkan thermal springs (Eme et al., 2013). I thus place Thaumarchaeota as the second branch in the archaeal subtree.

Moreira and colleagues, using bacterial proteins (including ribosomal proteins) as outgroup, have recently proposed to root the archaeal tree in the branch leading to Euryarchaeota (Petitjean et al., 2014). As a consequence, they propose to create a new phylum, Proteoarchaeota, grouping Crenarchaeota and Thaumarchaeota, together with the putative phyla Aigarchaeota and Korarchaeota. Proteoarchaea thus corresponds to the previously so-called “TACK superphylum” (Thaumarchaeota, “Aigarchaeota,” Crenarchaeota, Korarchaeota). However, as previously mentioned, using the same strategy, Gribaldo and colleagues obtained a root located within Euryarchaeota, more precisely in between subphyla I and II (Raymann et al., 2015). In contrast, their archaeal tree is rooted between “TACK/Proteoarchaeota” and Euryarchaeota when they used eukaryotic proteins as an outgroup. It will be interesting to see if they recover the root in the branch leading to Lokiarchaeota in future analyses using eukaryotic sequences as an outgroup.

Moreira and colleagues argue that eukaryotic proteins cannot be used to root the archaeal tree if Eukarya emerged from within Archaea. However, in the framework of the classical Woese tree, it makes more sense to root the archaeal tree using eukaryotic proteins as outgroup, because these proteins are much more closely related than bacterial proteins to their archaeal orthologs. Notably, the rooting between Lokiarchaeota/Thaumarchaeota and other Archaea, obtained in that case is more parsimonious than the rooting between Euryarchaeota and other Archaea in explaining the presence in Lokiarchaeota/Thaumarchaeota (including “Aigarchaeota,” see below) of many eukaryotic features lacking in other Archaea (Brochier-Armanet et al., 2008b; Spang et al., 2010, 2015; Koonin and Yutin, 2014).

Euryarchaeota are divided in two sub-phyla I and II, according to the presence/absence of DNA gyrase, a bacterial DNA topoisomerase that was transferred once in the phylum Euryarchaeota (Raymann et al., 2014). The sub-phylum I Euryarchaeota corresponds to those lacking DNA gyrase and encompasses Thermococcales, *Nanoarchaeum*, and class I methanogens, whereas sub-phylum II corresponds to those containing DNA gyrase and encompasses Archaeoglobales,

Thermoplasmatales, Halobacteriales, and class II methanogens (Forterre et al., 2014b).

Phylogenetic analyses have shown that DNA gyrase has been transferred from Bacteria to Archaea (Raymann et al., 2014). This transfer was an important and unique event that had a critical impact on chromosome structure and patterns of gene expression. Indeed, plasmids from all archaea are relaxed or slightly positively supercoiled, whereas plasmids from member of sub-phylum II Euryarchaeota containing gyrase are negatively supercoiled (Forterre et al., 2014b). Once transferred, DNA gyrase became most likely essential, as demonstrated in the case of Halobacteriales and Methanococcales, because such drastic modification in DNA topology modifies all protein DNA interactions involved in replication and transcription (for review and discussion, see Forterre and Gadelle, 2009). Indeed, to date, the loss of DNA gyrase has not been reported in any organism. Importantly, the phylogeny of archaeal DNA gyrase is fully congruent with the phylogeny of sub-phylum II Euryarchaeota, suggesting that, once transferred to the ancestor of this group, DNA gyrase has co-evolved with sub-phylum II Euryarchaeota (Raymann et al., 2014). Accordingly, considering the importance of DNA gyrase in cell physiology (DNA topology controlling all gene expression patterns) I suggest calling sub-phylum II Euryarchaeota, the neo-euryarchaeota. This name emphasizes the fact that the ancestor of this sub-phylum lived after the formation of the major bacterial phyla, since archaeal DNA gyrases branch within bacterial ones (Raymann et al., 2014).

Since all rooting indicated in the tree of **Figure 3**, as well as most internal nodes within domains, are controversial, I present a second tree (**Figure 4**), in which the information is limited to only that accepted by consensus. Accordingly, each one of the three domains is shown in a radial form without roots and only a few nodes within domains that seem supported by strong phylogenetic analyses are indicated (Brochier-Armanet et al., 2008a; He et al., 2014; Kamke et al., 2014; Ramulu et al., 2014; Spang et al., 2015).

Finally, I also present a third tree in which Eukarya emerged from within Archaea (**Figure 5**). This tree includes the new root proposed by Gribaldo and co-workers for Archaea (Raymann et al., 2015) and shows Lokiarchaeota as sister group of Eukarya (Spang et al., 2015). Notably, if future analyses demonstrate that such a tree is the more likely tree, Archaea will not be a valid taxon anymore, except if one accepts to consider eukaryotes as a particular archaeal phylum (much like *Homo* is a particular lineage of Apes)! In that case, the name Arkarya could be substituted to Archaea. Eukarya will become a particular phylum of Arkarya, beside Euryarkaryota, Crenarkaryota, Thaumarkaryota, and Lokiarkaryota (**Figure 5**).

As can be seen, **Figures 4** and **5** can be easily deduced from **Figure 3**. This indicates that it will be easy to update and modify these trees following the accumulation of new data from comparative genomics and phylogenetic analyses.

The Archaeal Tree

Figure 6 illustrates a rather detailed, but schematic, archaeal tree as a tribute to this issue devoted to Archaea. This tree has been adapted from the ribosomal protein tree of

Brochier-Armanet et al. (2011) and from a recent phylogeny based on the concatenation of 273 proteins conserved in at least 119 archaeal species out of 129 (Petitjean et al., 2015; thereafter called the archaeal protein tree). I also include the recently described candidate phylum “Lokiarchaeota” (corresponding to the DSAG clade) considering its importance for the discussions about the origin of Eukarya. The various roots that have been proposed for the domain Archaea are indicated by orange circles.

Aigarchaeota are included within Thaumarchaeota, because the latter were originally defined as a major archaeal phylum encompassing all archaeal lineages that are sister groups of Crenarchaeota in rDNA analyses (Brochier-Armanet et al., 2011). In the original paper in which we propose this new phylum, we noticed that: “*The diversity of mesophilic crenarchaeota—that we proposed to rename Thaumarchaeota—based on SSU rRNA sequence is comparable to that of hyperthermophilic Crenarchaeota and Euryarchaeota, which suggests that they represent a major lineage that has equal status to Euryarchaeota and Crenarchaeota*” (Brochier-Armanet et al., 2008a). We also predicted that the mesophily of Thaumarchaeota “could be challenged by the future identification of non-mesophilic organisms that belong to this phylum.” This suggests considering *candidatus Caldarchaeum subterraneum* as a thermophilic member of the Thaumarchaeota and not as the prototype of a new phylum (Aigarchaeota). In agreement with this proposal, *candidatus C. subterraneum* emerges as sister group of other thaumarchaea in most phylogenies (Brochier-Armanet et al., 2011; Nunoura et al., 2011; Eme et al., 2013; Guy et al., 2014; Petitjean et al., 2014, 2015; Raymann et al., 2015; Spang et al., 2015). Furthermore, *candidatus C. subterraneum* exhibits all molecular features first used to define the phylum Thaumarchaeota (Brochier-Armanet et al., 2008a; Spang et al., 2010), such as a eukaryotic-like Topo IB, which is absent from all other Archaea (Brochier-Armanet et al., 2008b). Topo IB is absent in Lokiarchaeota, but one must bear in mind that the reconstituted genome is only 92% complete (Spang et al., 2015). The phylum Thaumarchaeota should also include uncultivated archaea from the clade MCG (the Miscellaneous Crenarchaeal Group), since these organisms systematically form monophyletic groups with Thaumarchaea and “Aigarchaeota” in phylogenetic analyses (Spang et al., 2015).

The recently proposed phylum “Geoarchaeota” is included as a sister group of Thermoproteales (without phylum status) as suggested by the analysis of Ettema and co-workers (Guy et al., 2014; Spang et al., 2015). *Thermofilum* always branch very early as a sister group of Thermoproteales. This suggests that *Thermofilum*, as well as Geoarchaeota, could have an order status.

Korarchaeota branch in-between Euryarchaeota and other archaea (Crenarchaeota, Thaumarchaeota) in the ribosomal proteins tree. Unfortunately, this phylum is presently represented by a single species whose genome has been sequenced, *candidatus Korarchaeum cryptofilum*. The genome of *candidatus K. cryptofilum* harbors a mixture of features characteristic of the three other archaeal phyla. This can justify maintaining a phylum status for this group for the moment. More genome sequences of Korarchaeota are nevertheless required to confirm this point.

As previously discussed, Euryarchaeota are divided into two groups depending of the presence of DNA gyrase.

Neo-euryarchaeota (sub-phylum II) is a monophyletic group in all phylogenetic analyses. In contrast, sub-phylum I is paraphyletic in the ribosomal and archaeal protein trees (Brochier-Armanet et al., 2011; Petitjean et al., 2015). However, they form a monophyletic assemblage in a tree based on replication proteins (Raymann et al., 2014) and in a recent phylogenomic analysis performed by Makarova et al. (2015) that involved both comparison of multiple phylogenetic trees and a search for putative synapomorphies. I thus decided to favor the monophyly of sub-phylum I in the tree of **Figure 6**.

The close relationship between plasmids of Thermococcales and Methanococcales also suggests that these two orders could be closely related (Soler et al., 2010). It is possible that the emergence of pseudomurein in Methanobacteriales and Methanopyrales allowed these archaea to get rid of mobile elements that used to infect the ancestors of Thermococcales and Methanococcales. Notably, Methanopyrales and Methanobacteriales are monophyletic in the archaeal protein tree, suggesting that the presence of pseudomurein is a synapomorphy for these sister groups (Petitjean et al., 2015). Petitjean and coworkers recently proposed the name Methanomada (superclass) for class I methanogens, that are monophyletic in their protein tree (Petitjean et al., 2015) and in the DNA replication tree (Raymann et al., 2014).

In contrast to class I, the four orders of class II methanogens that are included within neo-euryarchaeota are always paraphyletic in phylogenetic analyses. Methanogens of the recently described order Methanomassiliicoccus form a monophyletic assemblage with Thermoplasmatales, the moderate thermoacidophilic strain *Aciduliprofundum boonei* and several lineages of uncultivated archaea in a ribosomal protein tree (Borrel et al., 2013). The name Diafoarchaea has been proposed for this major subgroup (superclass) of neo-Euryarchaeota (Petitjean et al., 2015).

Altarchaeales correspond to a recently described mesophilic archaeum, *Candidatus Altarchaeum hamiconexum*, characterized by fascinating appendages (Hami) that groups with Methanococcales in a ribosomal protein tree, but between Euryarchaeota of sub-phyla I and II in a tree based on several other universal proteins (Probst et al., 2014). Since *Candidatus A. hamiconexum* contain the two DNA gyrase genes, it is located at the base of the neo-euryarchaeota in the tree of **Figure 6**.

Rinke et al. (2013) have recently proposed promoting the nanosized archaea *Nanoarchaeum*, *Parvarchaeum* (ARMAN 4 and 5) and Nanohaloarchaea to phylum level and grouping them with two other putative new phyla of Archaea (“Aenigmarchaeota” and “Diapherotrites”) into a new superphylum, called DPANN (Rinke et al., 2013). In their phylogenetic analysis based on 38 “universal proteins,” the root of the archaeal tree is located between this putative DPANN superphylum and all other Archaea (see Figure S11 in Rinke et al., 2013). However, their universal protein data set is confusing because it contains eukaryotic proteins of bacterial origin (Williams and Embley, 2014). In the recent tree of Williams and Embley supporting the archaeal origin of eukaryotes, the archaeal tree is rooted between DPANN and Euryarchaeota (see Figure 3 in Williams and Embley, 2014). The basal position of the nanosized archaea in these trees confirms the difficulty of using universal proteins to resolve ancient phylogenies, because

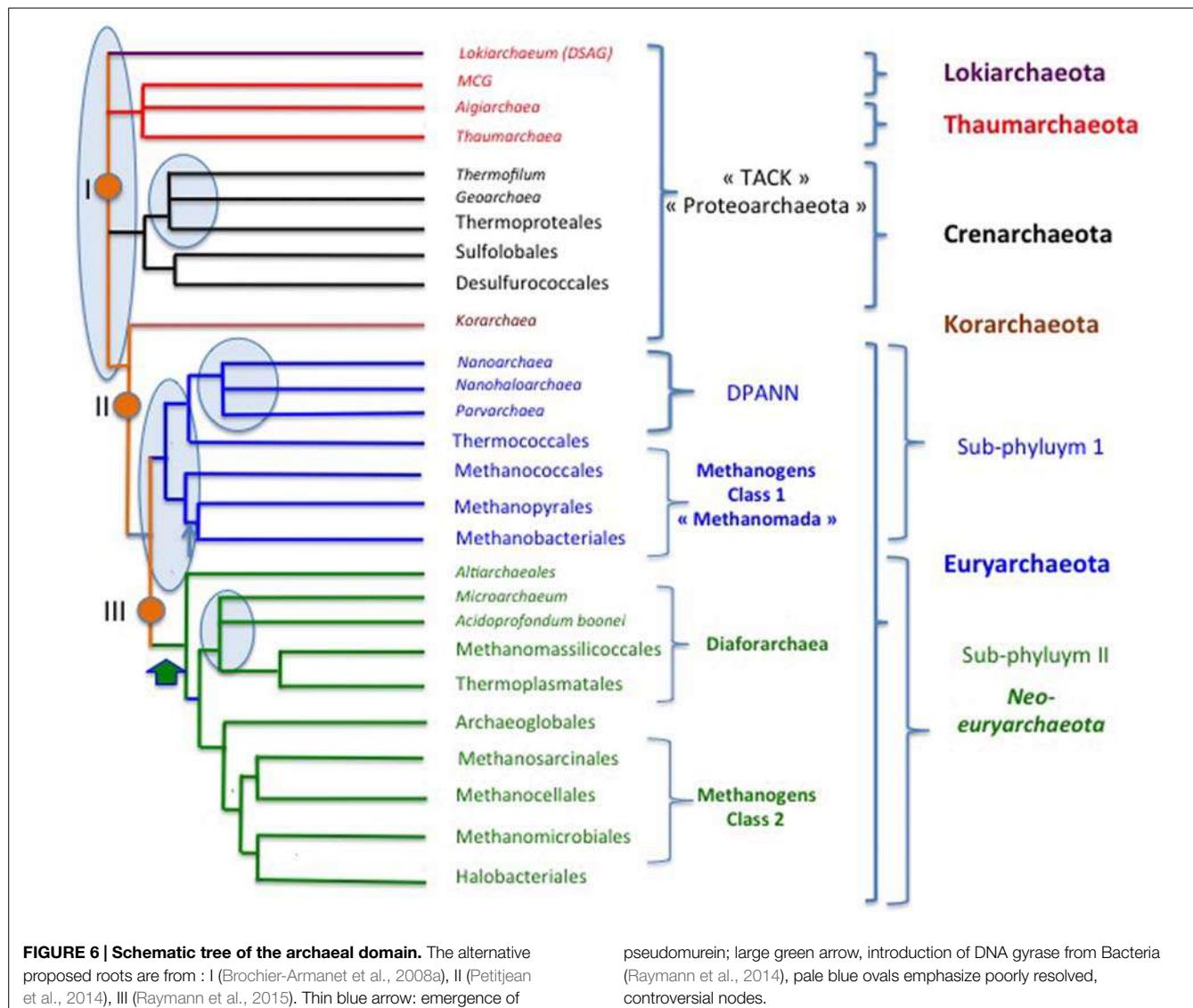


FIGURE 6 | Schematic tree of the archaeal domain. The alternative proposed roots are from: I (Brochier-Armanet et al., 2008a), II (Petitjean et al., 2014), III (Raymann et al., 2015). Thin blue arrow: emergence of

pseudomurein; large green arrow, introduction of DNA gyrase from Bacteria (Raymann et al., 2014), pale blue ovals emphasize poorly resolved, controversial nodes.

this position is clearly misleading (see below and Petitjean et al., 2014). In the case of nanosized archaea, phylogenetic analyses are even more challenging because their proteins tend to evolve rapidly, producing very long branches in phylogenetic trees. This suggests that these nanosized archaea evolve mainly by streamlining (Brochier-Armanet et al., 2011; Petitjean et al., 2014). They have not been included in the protein tree of Petitjean and coworkers (Petitjean et al., 2015).

Previous phylogenetic and comparative genomic analyses focusing on *Nanoarchaeum equitans* have suggested that this fascinating archaeal symbiont belongs to the Euryarchaeota and could be distant relatives of Thermococcales (Brochier et al., 2005b). This sister relationship was later on supported by the discovery of the shared presence of a tRNA modification protein that was recently transferred from Bacteria to both *N. equitans* and Thermococcales (Urbonavicius et al., 2008). The sisterhood of *N. equitans* and Thermococcales has been observed again in more recent analyses based on ribosomal proteins (Brochier-Armanet

et al., 2011) or in the archaeal tree of Moreira and colleagues (Petitjean et al., 2014).

In the tree of **Figure 6**, Thermococcales, *Nanoarchaeum*, *Parvarchaeum*, and *Nanohaloarchaea* tentatively form a monophyletic group. This is supported by several lines of converging (but weak and controversial) evidences. The grouping of Thermococcales and *Nanoarchaeum* with *Parvarchaeum* (ARMAN 4 and 5) is supported by the phylogeny based on ribosomal proteins (Brochier-Armanet et al., 2011), whereas the grouping of *Nanoarchaeum* and *Parvarchaeum* with *Nanohaloarchaea* is supported by the phylogeny of DNA replication proteins (Raymann et al., 2014). Interestingly, the grouping of *Parvarchaeum*, *Nanoarchaeum*, and *Nanohaloarchaea* is supported by the shared presence of an atypical small primase corresponding to the fusion of the two subunits of the *bona fide* archaeal/eukaryal primases, PriS, and PriL (Raymann et al., 2014). It has been suggested that this fusion corresponds to a convergent evolution associated to streamlining (Petitjean et al.,

2014). However, this is unlikely because these unusual monomeric primase primases are also highly divergent in terms of amino-acid sequence from the classical archaeal/eukaryal primases and very similar in all nanosized archaea. Alternatively, it has been suggested that this primase has been distributed between nanosized archaea by horizontal gene transfers (Raymann et al., 2014). This seems also unlikely because nanosized archaea live in very different types of environments (high temperature for known *Nanoarchaeum*, high salt for Nanohaloarchaea, high acidity for *Parvarchaeum*). It seems more likely that this primase was acquired by a common ancestor to *Nanoarchaeum*, *Parvarchaeum*, and Nanohaloarchaea from a mobile genetic element (Raymann et al., 2014). It has indeed been shown that some archaeal plasmids encode unusual primases from the PriS/PriL superfamily that are only distantly related to *bona fide* archaeal and eucaryal primases (Lipps et al., 2004; Krupovic et al., 2013; Gill et al., 2014).

The grouping of Nanohaloarchaea with other nanosized archaea, as in **Figure 6**, is especially controversial because they emerge as sister group of Halobacteriales in a ribosomal protein tree (Narasingarao et al., 2012). However, if Nanohaloarchaea are sister group of Halobacteriales, they would be the only members of neo-euryarchaeota lacking DNA gyrase. Nanohaloarchaea might have lost this enzyme during the streamlining process related to their small size. However, the loss of DNA gyrase has not been reported until now in any other free-living organisms. This is why I finally choose to group Nanohaloarchaea with other nanosized archaea in the tree of **Figure 6**. It is possible that the atypical amino acid composition of the nanohaloarchaeal proteome (Narasingarao et al., 2012), linked to salt adaptation, introduces a bias favoring the artificial grouping of Nanohaloarchaea and Haloarchaea. This would also explain why Nanohaloarchaea attracts *Nanoarchaeum* and *Parvarchaeum* away from Thermococcales and closer to Haloarchaea in the DNA replication tree (Raymann et al., 2014).

The nanosized archaeon *Candidatus Micrarchaeum acidophilum* (ARMAN 2) branches together with *Parvarchaeum* (ARMAN 4, 5) in a rDNA tree (Baker et al., 2006) and in the tree of Moreira and coworkers (Petitjean et al., 2014). However, it branches away from *Parvarchaeum* in the ribosomal tree, as an early branching neo-euryarchaeon. The latter position is supported by the presence of DNA gyrase in *Candidatus Micrarchaeum acidophilum* and the absence of the single subunit primase characteristic of other nanosized archaea (Raymann et al., 2014). I thus included *Micrarchaeum* among the superclass Diafoarchaea in the tree of **Figure 6**. Finally, the phylogenetic position and status of “Aenigmarchaeota” and “Diapherotrites” cannot presently be determined because these groups have only been defined from single cell genomic analyses and their genomes are probably incomplete (Petitjean et al., 2014).

Further analyses and (hopefully) many more isolates are clearly required to determine the correct position of the various groups of nanosized archaea in the archaeal tree. From all these considerations, it is clear that the “DPNN superphylum” is an artificial construction. The same is true for the “TACK superphylum” and the candidate phylum

“Proteoarchaea” if the root of the archaeal tree is located within Euryarchaeota (Raymann et al., 2015) or between Thaumarchaeota/Lokiarchaeota and other Archaea. As we previously discussed, defining the root of the archaeal tree strongly depends on choosing between different scenarios for the universal tree. Since the rooting of the archaeal tree is still in debate, I would not recommend at the moment to use names such as “Proteoarchaea” or “TACK superphylum” in archaeal phylogeny.

This section on archaeal phylogeny has illustrated the fact that, in addition to the root of the tree itself, several nodes in the archaeal tree are still controversial and require more data and more work to be carried out. These nodes have been marked by circles in blue in the tree of **Figure 6**. Future progress will probably come from the sequencing of more genomes, especially in poorly represented groups and in the many groups that are presently only known from environmental rDNA sequences.

Conclusion

I hope that the universal and archaeal trees proposed here will be useful as new metaphors illustrating the history of life on our planet. From the above review, it should be clear that there is no protein or groups of proteins that can give the real species tree, i.e., allow us to recapitulate safely the exact path of life evolution. In particular, one should be cautious with composite trees based on the concatenation of protein sequences or addition of individual trees. The results obtained should always be compared to the result of individual trees. Martin and colleagues recently reported a lack of correspondence between individual protein trees and the concatenation tree in several datasets of archaeal and bacterial proteins (Thiergart et al., 2014). This can reflect either LGT and/or the absence of real phylogenetic signal. Importantly, this lack of correspondence between individual protein trees and concatenation trees is also observed in the analyses that place Eukarya within Archaea (Cox et al., 2008; Williams and Embley, 2014), raising doubts on results supporting the tree of **Figure 5**. Unfortunately, individual trees are not always available in the supplementary data of published studies (Katz and Grant, 2014; Raymann et al., 2015). In our previous work on archaeal phylogeny, careful analysis of individual trees in parallel to their concatenation was critical to obtain a rather confident tree for Archaea based on ribosomal proteins (Matte-Tailliez et al., 2002) and to find the (probably) correct position of Nanoarchaeota, as sister group of Thermococcales (Brochier et al., 2005b). Notably, to obtain this result in Brochier et al. (2005a), it was necessary to remove the proteins from the large ribosome subunit because several of them exhibit a surprising affinity with their crenarchaeal homologs in individual trees, possibly indicating LGT between *N. equitans* and its host, the crenarchaeon *Ignicoccus*. It is also very important to compare trees obtained with different datasets, such as translation, transcription and DNA replication trees, to point-point discrepancies and identify their causes (Brochier et al., 2004, 2005a; Raymann et al., 2014).

In summary, it is (and it will be) only possible to draw schematic (theoretical) trees by combining data from multiple

phylogenetic protein trees with information deduced from probable synapomorphies. It is the exercise that I have tried to do here in drawing the trees of **Figures 3–6**. These organismal trees, which are supposed to recapitulate the history of ribosome encoding organisms, will probably evolve themselves, with the availability of new genomes (especially from poorly sampled groups), better phylogenetic analyses, and the identification of new synapomorphies defining specific domains, sub-phyla groups and superphyla. For instance, my preliminary analyses of universal proteins sequence alignments indicate that the Lokiarchaeon is probably neither an early branching archaeon nor a missing link between Archaea and Eukarya (see also Nasir et al., 2015). It should be relatively easy to update these trees as new data

accumulate in the future and use them in discussions of various controversial scenarios regarding the evolution of ancient life.

Acknowledgments

I thank Mechthild Pohlschroeder and Sonja-Verena Albers for inviting me to draw an updated version of the universal tree of life and David Prangishvili for suggesting the name Arkarya. I am grateful to Sukhvinder Gill for English corrections and a referee for extensive critical analysis. I am supported by an ERC grant from the European Union's Seventh Framework Program (FP/2007–2013)/Project EVOMOBIL—ERC Grant Agreement no.340440.

References

- Albers, S. V., Forterre, P., Prangishvili, D., and Schleper, C. (2013). The legacy of Carl Woese and Wolfram Zillig: from phylogeny to landmark discoveries. *Nat. Rev. Microbiol.* 11, 713–719. doi: 10.1038/nrmicro3124
- Baker, B. J., Tyson, G. W., Webb, R. I., Flanagan, J., Hugenholtz, P., Allen, E. E., et al. (2006). Lineages of acidophilic archaea revealed by community genomic analysis. *Science* 314, 1933–1935. doi: 10.1126/science.1132690
- Borrel, G., O'Toole, P. W., Harris, H. M., Peyret, P., Brugère, J. F., and Gribaldo, S. (2013). Phylogenomic data support a seventh order of methylotrophic methanogens and provide insights into the evolution of methanogenesis. *Genome Biol. Evol.* 5, 1769–1780. doi: 10.1093/gbe/evt128
- Boussau, B., Blanquart, S., Necsulea, A., Lartillot, N., and Gouy, M. (2008a). Parallel adaptations to high temperatures in the Archaean eon. *Nature* 456, 942–945. doi: 10.1038/nature07393
- Boussau, B., Guégan, L., and Gouy, M. (2008b). Accounting for horizontal gene transfers explains conflicting hypotheses regarding the position of Aquificales in the phylogeny of Bacteria. *BMC Evol. Biol.* 8:272. doi: 10.1186/1471-2148-8-272
- Braakman, R., and Smith, E. (2014). Metabolic evolution of a deep-branching hyperthermophilic chemoautotrophic bacterium. *PLoS ONE* 9:e87950. doi: 10.1371/journal.pone.0087950
- Brinkmann, H., and Philippe, H. (1999). Archaea sister group of Bacteria? Indications from tree reconstruction artefacts in ancient phylogenies. *Mol. Biol. Evol.* 16, 817–825. doi: 10.1093/oxfordjournals.molbev.a026166
- Brochier-Armanet, C., Boussau, B., Gribaldo, S., and Forterre, P. (2008a). Mesophilic crenarchaeota: proposal for a third archaeal phylum, the Thaumarchaeota. *Nat. Rev. Microbiol.* 6, 245–252. doi: 10.1038/nrmicro1852
- Brochier-Armanet, C., Gribaldo, S., and Forterre, P. (2008b). A DNA topoisomerase IB in Thaumarchaeota testifies for the presence of this enzyme in the last common ancestor of Archaea and Eucarya. *Biol. Direct.* 3, 54. doi: 10.1186/1745-6150-3-54
- Brochier-Armanet, C., and Forterre, P. (2007). Widespread distribution of archaeal reverse gyrase in thermophilic bacteria suggests a complex history of vertical inheritance and lateral gene transfers. *Archaea* 2, 83–93. doi: 10.1155/2006/582916
- Brochier-Armanet, C., Forterre, P., and Gribaldo, S. (2011). Phylogeny and evolution of the Archaea: one hundred genomes later. *Curr. Opin. Microbiol.* 14, 274–281. doi: 10.1016/j.mib.2011.04.015
- Brochier, C., Forterre, P., and Gribaldo, S. (2004). Archaeal phylogeny based on proteins of the transcription and translation machineries: tackling the *Methanopyrus kandleri* paradox. *Genome Biol.* 5:R17. doi: 10.1186/gb-2004-5-3-r17
- Brochier, C., Forterre, P., and Gribaldo, S. (2005a). An emerging phylogenetic core of Archaea: phylogenies of transcription and translation machineries converge following addition of new genome sequences. *BMC Evol. Biol.* 5:36. doi: 10.1186/1471-2148-5-36
- Brochier, C., Gribaldo, S., Zivanovic, Y., Confalonieri, F., and Forterre, P. (2005b). Nanoarchaea: representatives of a novel archaeal phylum or a fast-evolving euryarchaeal lineage related to Thermococcales? *Genome Biol.* 6:R42. doi: 10.1186/gb-2005-6-5-r42
- Brochier, C., and Philippe, H. (2002). Phylogeny: a non-hyperthermophilic ancestor for bacteria. *Nature* 417, 244. doi: 10.1038/417244a
- Burger, G., Gray, M. W., Forget, L., and Lang, B. F. (2013). Strikingly bacteria-like and gene-rich mitochondrial genomes throughout Jakobid protists. *Genome Biol. Evol.* 5, 418–438. doi: 10.1093/gbe/evt008
- Cavalier-Smith, T. (2010). Origin of the cell nucleus, mitosis and sex: roles of intracellular coevolution. *Biol. Direct.* 5, 7. doi: 10.1186/1745-6150-5-7
- Ciccarelli, F. D., Doerks, T., von Mering, C., Creevey, C. J., Snel, B., and Bork, P. (2006). Toward automatic reconstruction of a highly resolved tree of life. *Science* 311, 1283–1287. doi: 10.1126/science.1123061
- Cox, C. J., Foster, P. G., Hirt, R. P., Harris, S. R., and Embley, T. M. (2008). The archaeabacterial origin of eukaryotes. *Proc. Natl. Acad. Sci. U.S.A.* 105, 20356–20361. doi: 10.1073/pnas.0810647105
- Csurös, M., and Miklos, I. (2009). Streamlining and large ancestral genomes in Archaea inferred with a phylogenetic birth-and-death model. *Mol. Biol. Evol.* 26, 2087–2095. doi: 10.1093/molbev/msp123
- De Duve C. (2007). The origin of eukaryotes: a reappraisal. *Nat. Rev. Genet.* 8, 395–403. doi: 10.1038/nrg2071
- Derelle, R., Torruella, G., Klimeš, V., Brinkmann, H., Kim, E., Vlček, Č., et al. (2015). Bacterial proteins pinpoint a single eukaryotic root. *Proc. Natl. Acad. Sci. U.S.A.* 112, E693–E699. doi: 10.1073/pnas.1420657112
- Desmond, E., Brochier-Armanet, C., Forterre, P., and Gribaldo, S. (2011). On the last common ancestor and early evolution of eukaryotes: reconstructing the history of mitochondrial ribosomes. *Res. Microbiol.* 162, 53–70. doi: 10.1016/j.resmic.2010.10.004
- Deutsch, C., El Yacoubi, B., de Crécy-Lagard, V., and Iwata-Reuyl, D. (2012). Biosynthesis of threonylcarbamoyl adenosine (t6A), a universal tRNA nucleoside. *J. Biol. Chem.* 287, 13666–13673. doi: 10.1074/jbc.M112.344028
- Devos, D. P. (2012). Regarding the presence of membrane coat proteins in bacteria: confusion? What confusion? *Bioessays* 34, 38–39. doi: 10.1002/bies.201100147
- Doolittle, W. F. (1998). You are what you eat: a gene transfer ratchet could account for bacterial genes in eukaryotic nuclear genomes. *Trends Genet.* 14, 307–311. doi: 10.1016/S0168-9525(98)01494-2
- Doolittle, W. F. (2000). Uprooting the tree of life. *Sci. Am.* 282, 90–95. doi: 10.1038/scientificamerican0200-90
- Edgell, D. R., and Doolittle, W. F. (1997). Archaea and the origin(s) of DNA replication proteins. *Cell* 89, 995–998. doi: 10.1016/S0022-8674(00)80285-8
- El Albani, A., Bengtson, S., Canfield, D. E., Bekker, A., Macchiarelli, R., Mazurier, A., et al. (2010). Large colonial organisms with coordinated growth in oxygenated environments 2.1 Gyr ago. *Nature* 466, 100–104. doi: 10.1038/nature09166
- Embley, T. M., and Hirt, R. P. (1998). Early branching eukaryotes? *Curr. Opin. Genet. Dev.* 8, 624–662. doi: 10.1016/S0959-437X(98)80029-4
- Eme, L., Reigstad, L. J., Spang, A., Lanzén, A., Weinmaier, T., Rattei, T., et al. (2013). Metagenomics of Kamchatkan hot spring filaments reveal two new major (hyper)thermophilic lineages related to Thaumarchaeota. *Res. Microbiol.* 164, 425–438. doi: 10.1016/j.resmic.2013.02.006
- Eveleigh, R. J., Meehan, C. J., Archibald, J. M., and Beiko, R. G. (2013). Being Aquifex aeolicus: untangling a hyperthermophile's checkered past. *Genome Biol. Evol.* 5, 2478–2497. doi: 10.1093/gbe/evt195

- Filée, J., and Forterre, P. (2005). Viral proteins functioning in organelles: a cryptic origin? *Trends Microbiol.* 13, 510–513. doi: 10.1016/j.tim.2005.08.012
- Forterre, P. (1992). Neutral terms. *Nature* 335, 305. doi: 10.1038/355305c0
- Forterre, P. (1995). Thermoreduction, a hypothesis for the origin of prokaryotes. *C. R. Acad. Sci. III.* 318, 415–422.
- Forterre, P. (1996). A hot topic: the origin of hyperthermophiles. *Cell* 85, 789–792. doi: 10.1016/S0092-8674(00)81262-3
- Forterre, P. (1999). Displacement of cellular proteins by functional analogues from plasmids or viruses could explain puzzling phylogenies of many DNA informational proteins. *Mol. Microbiol.* 33, 457–465. doi: 10.1046/j.1365-2958.1999.01497.x
- Forterre, P. (2002). The origin of DNA genomes and DNA replication proteins. *Curr. Opin. Microbiol.* 5, 525–532. doi: 10.1016/S1369-5274(02)00360-0
- Forterre, P. (2006). Three RNA cells for ribosomal lineages and three DNA viruses to replicate their genomes: a hypothesis for the origin of cellular domain. *Proc. Natl. Acad. Sci. U.S.A.* 103, 3669–3674. doi: 10.1073/pnas.0510333103
- Forterre, P. (2011). A new fusion hypothesis for the origin of Eukarya: better than previous ones, but probably also wrong. *Res. Microbiol.* 162, 77–91. doi: 10.1016/j.resmic.2010.10.005
- Forterre, P. (2012). Darwin's goldmine is still open: variation and selection run the world. *Front. Cell Infect. Microbiol.* 2:106. doi: 10.3389/fcimb.2012.00106
- Forterre, P. (2013a). The common ancestor of archaea and eukarya was not an archaeon. *Archaea* 2013, 18. doi: 10.1155/2013/372396
- Forterre, P. (2013b). Why are there so many diverse replication machineries? *J. Mol. Biol.* 425, 4714–4726. doi: 10.1016/j.jmb.2013.09.032
- Forterre, P., Benachenhou-Lahfa, N., Confalonieri, F., Duguet, M., Elie, C., and Labedan, B. (1992). The nature of the last universal ancestor and the root of the tree of life, still open questions. *Biosystems* 28, 15–32. doi: 10.1016/0303-2647(92)90004-I
- Forterre, P., and Gadelle, D. (2009). Phylogenomics of DNA topoisomerases: their origin and putative roles in the emergence of modern organisms. *Nucleic Acids Res.* 37, 679–692. doi: 10.1093/nar/gkp032
- Forterre, P., and Gribaldo, S. (2010). Bacteria with a eukaryotic touch: a glimpse of ancient evolution? *Proc. Natl. Acad. Sci. U.S.A.* 107, 12739–12740. doi: 10.1073/pnas.1007720107
- Forterre, P., Krupovic, M., and Prangishvili, D. (2014a). Cellular domains and viral lineages. *Trends Microbiol.* 22, 554–558. doi: 10.1016/j.tim.2014.07.004
- Forterre, P., Krupovic, M., Raymann, K., and Soler, N. (2014b). Plasmids from euryarchaeota. *Microbiol. Spectr.* 2, 1–27. doi: 10.1128/microbiolspec.PLAS-0027-2014
- Forterre, P., and Philippe, H. (1999). Where is the root of the universal tree of life? *Bioessays* 21, 871–879. doi: 10.1002/(SICI)1521-1878(199910)21:10<871::AID-BIES10>3.0.CO;2-Q
- Fritz-Laylin, L. K., Prochnik, S. E., Ginger, M. L., Dacks, J. B., Carpenter, M. L., Field, M. C., et al. (2010). The genome of *Naegleria gruberi* illuminates early eukaryotic versatility. *Cell* 140, 631–642. doi: 10.1016/j.cell.2010.01.032
- Galtier, N., Tourasse, N., and Gouy, M. (1999). A non-hyperthermophilic common ancestor to extant life forms. *Science* 283, 220–221. doi: 10.1126/science.283.5399.220
- Gill, S., Krupovic, M., Desnoues, N., Béguin, P., Sezonov, G., and Forterre, P. (2014). A highly divergent archaeo-eukaryotic primase from the *Thermococcus nautilus* plasmid, pTN2. *Nucleic Acids Res.* 42, 3707–3719. doi: 10.1093/nar/gkt1385
- Gogarten, J. P., Kibak, H., Dittrich, P., Taiz, L., Bowman, E. J., Bowman, B. J., et al. (1989). Evolution of the vacuolar H⁺-ATPase: implications for the origin of eukaryotes. *Proc. Natl. Acad. Sci. U.S.A.* 86, 6661–6665. doi: 10.1073/pnas.86.17.6661
- Glansdorff, N., Xu, Y., and Labedan, B. (2008). The last universal common ancestor: emergence, constitution and genetic legacy of an elusive forerunner. *Biol. Direct.* 3:29. doi: 10.1186/1745-6150-3-29
- Gribaldo, S., and Brochier, C. (2009). Phylogeny of prokaryotes: does it exist and why should we care? *Res. Microbiol.* 160, 513–521. doi: 10.1016/j.resmic.2009.07.006
- Gribaldo, S., and Philippe, H. (2002). Ancient phylogenetic relationships. *Theor. Popul. Biol.* 61, 391–408. doi: 10.1006/tbei.2002.1593
- Gribaldo, S., Poole, A. M., Daubi, V., Forterre, P., and Brochier-Armanet, C. (2010). The origin of eukaryotes and their relationship with the Archaea: are we at a phylogenomic impasse? *Nat. Rev. Microbiol.* 8, 743–752. doi: 10.1038/nrmicro2426
- Groussin, M., Boussau, B., Charles, S., Blanquart, S., and Gouy, M. (2013). The molecular signal for the adaptation to cold temperature during early life on Earth. *Biol. Lett.* 9, 20130608. doi: 10.1098/rsbl.2013.0608
- Groussin, M., and Gouy, M. (2011). Adaptation to environmental temperature is a major determinant of molecular evolutionary rates in archaea. *Mol. Biol. Evol.* 28, 2661–2674. doi: 10.1093/molbev/msr098
- Guy, L., and Ettema, T. J. (2011). The archaeal 'TACK' superphylum and the origin of eukaryotes. *Trends Microbiol.* 19, 580–587. doi: 10.1016/j.tim.2011.09.002
- Guy, L., Spang, A., Saw, J. H., and Ettema, T. J. (2014). 'Geoarchaeote NAGI' is a deeply rooting lineage of the archaeal order Thermoproteales rather than a new phylum. *ISME J.* 8, 1353–1357. doi: 10.1038/ismej.2014.6
- Harish, A., Tunlid, A., and Kurland, C. G. (2013). Rooted phylogeny of the three superkingdoms. *Biochimie* 95, 1593–1604. doi: 10.1016/j.biochi.2013.04.016
- Hashimoto, T., Nakamura, Y., Kamaishi, T., Nakamura, F., Adachi, J., Okamoto, K., et al. (1995). Phylogenetic place of mitochondrion-lacking protozoan, *Giardia lamblia*, inferred from amino acid sequences of elongation factor 2. *Mol. Biol. Evol.* 12, 782–793.
- Hecker, A., Leulliot, N., Gadelle, D., Graille, M., Justome, A., Dorlet, P., et al. (2007). An archaeal orthologue of the universal protein Kae1 is an iron metalloprotein which exhibits atypical DNA-binding properties and apurinic-endonuclease activity *in vitro*. *Nucleic Acids Res.* 35, 6042–6051. doi: 10.1093/nar/gkm554
- He D., Fiz-Palacios, O., Fu, C. J., Fehling, J., Tsai, C. C., and Baldauf, S. L. (2014). An alternative root for the eukaryote tree of life. *Curr. Biol.* 24, 465–470. doi: 10.1016/j.cub.2014.01.036
- Iwabe, N., Kuma, K., Hasegawa, M., Osawa, S., and Miyata, T. (1989). Evolutionary relationship of archaebacteria, eubacteria, and eukaryotes inferred from phylogenetic trees of duplicated genes. *Proc. Natl. Acad. Sci. U.S.A.* 86, 9355–8359. doi: 10.1073/pnas.86.23.9355
- Jørgensen, S. L., Thorseth, I. H., Pedersen, R. B., Baumberger, T., and Schleper, C. (2013). Quantitative and phylogenetic study of the Deep Sea Archaeal Group in sediments of the Arctic mid-ocean spreading ridge. *Front. Microbiol.* 4:299. doi: 10.3389/fmicb.2013.00299
- Kamaishi, T., Hashimoto, T., Nakamura, Y., Nakamura, F., Murata, S., Okada, N., et al. (1996). Protein phylogeny of translation elongation factor EF-1 alpha suggests microsporidians are extremely ancient eukaryotes. *J. Mol. Evol.* 42, 257–263. doi: 10.1007/BF02198852
- Kamke, J., Rinke, C., Schwientek, P., Mavromatis, K., Ivanova, N., Sczyrba, A., et al. (2014). The candidate phylum Poribacteria by single-cell genomics: new insights into phylogeny, cell-compartmentation, eukaryote-like repeat proteins, and other genomic features. *PLoS ONE* 9:e87353. doi: 10.1371/journal.pone.0087353
- Kamikawa, R., Kolisko, M., Nishimura, Y., Yabuki, A., Brown, M. W., Ishikawa, S. A., et al. (2014). Gene content evolution in Discobid mitochondria deduced from the phylogenetic position and complete mitochondrial genome of *Tsukubamonas globosa*. *Genome Biol. Evol.* 6, 306–315. doi: 10.1093/gbe/evu015
- Kannan, S., Rogozin, I. B., and Koonin, E. V. (2014). MitoCOGs: clusters of orthologous genes from mitochondria and implications for the evolution of eukaryotes. *BMC Evol. Biol.* 14:237. doi: 10.1186/s12862-014-0237-5
- Katz, L. A., and Grant, J. R. (2014). Taxon-rich phylogenomic analyses resolve the eukaryotic tree of life and reveal the power of subsampling by sites. *Syst. Biol.* [Epub ahead of print].
- Keeling, P. J., and McFadden, G. I. (1998). Origins of microsporidia. *Trends Microbiol.* 6, 19–23. doi: 10.1016/S0966-842X(97)01185-2
- Koonin, E. V., and Yutin, N. (2014). The dispersed archaeal eukaryome and the complex archaeal ancestor of eukaryotes. *Cold Spring Harb. Perspect. Biol.* 6:a016188. doi: 10.1101/cshperspect.a016188
- Krupovic, M., Gribaldo, S., Bamford, D. H., and Forterre, P. (2010). The evolutionary history of archaeal MCM helicases: a case study of vertical evolution combined with hitchhiking of mobile genetic elements. *Mol. Biol. Evol.* 27, 2716–2732. doi: 10.1093/molbev/msq161
- Krupovic, M., Gonnet, M., Hania, W. B., Forterre, P., and Eraso, G. (2013). Insights into dynamics of mobile genetic elements in hyperthermophilic environments from five new *Thermococcus* plasmids. *PLoS ONE* 8:e49044. doi: 10.1371/journal.pone.0049044
- Kurland, C. G., Collins, L. J., and Penny, D. (2006). Genomics and the irreducible nature of eukaryote cells. *Science* 312, 1011–1014. doi: 10.1126/science.1121674
- Lake, J. A., Henderson, E., Oakes, M., and Clark, M. W. (1984). Eocytes: a new ribosome structure indicates a kingdom with a close relationship to eukaryotes. *Proc. Natl. Acad. Sci. U.S.A.* 81, 3786–3790. doi: 10.1073/pnas.81.12.3786

- Lane, N., and Martin, W. (2010). The energetics of genome complexity. *Nature* 467, 929–934. doi: 10.1038/nature09486
- Lasek-Nesselquist, E., and Gogarten, J. P. (2013). The effects of model choice and mitigating bias on the ribosomal tree of life. *Mol. Phylogenet. Evol.* 69, 17–38. doi: 10.1016/j.ympev.2013.05.006
- Lecompte, O., Ripp, R., Thierry, J. C., Moras, D., and Poch, O. (2002). Comparative analysis of ribosomal proteins in complete genomes: an example of reductive evolution at the domain scale. *Nucleic Acids Res.* 30, 5382–5390. doi: 10.1093/nar/gkf693
- Leipe, D. D., Aravind, L., and Koonin, E. V. (1999). Did DNA replication evolve twice independently? *Nucleic Acids Res.* 27, 3389–3401. doi: 10.1093/nar/27.17.3389
- Lipps, G., Weinzierl, A. O., von Scheven, G., Buchen, C., and Cramer, P. (2004). Structure of a bifunctional DNA primase-polymerase. *Nat. Struct. Mol. Biol.* 11, 157–162. doi: 10.1038/nsmb723
- López-García, P., and Moreira, D. (2008). Tracking microbial biodiversity through molecular and genomic ecology. *Res. Microbiol.* 159, 67–73. doi: 10.1016/j.resmic.2007.11.019
- Makarova, K. S., Yutin, N., Bell, S. D., and Koonin, E. V. (2010). Evolution of diverse cell division and vesicle formation systems in Archaea. *Nat. Rev. Microbiol.* 8, 731–741. doi: 10.1038/nrmicro2406
- Makarova, K. S., Wolf, Y. I., and Koonin, E. V. (2015). Archaeal clusters of orthologous genes (arCOGs): an update and application for analysis of shared features between thermococcales, methanococcales, and methanobacteriales. *Life* 5, 818–840. doi: 10.3390/life5010818
- Martijn, J., and Ettema, T. J. (2013). From archaeon to eukaryote: the evolutionary dark ages of the eukaryotic cell. *Biochem. Soc. Trans.* 41, 451–457. doi: 10.1042/BST20120292
- Martin, W., Rujan, T., Richly, E., Hansen, A., Cornelsen, S., Lins, T., et al. (2002). Evolutionary analysis of Arabidopsis, cyanobacterial, and chloroplast genomes reveals plastid phylogeny and thousands of cyanobacterial genes in the nucleus. *Proc. Natl. Acad. Sci. U.S.A.* 99, 12246–12251. doi: 10.1073/pnas.182432999
- Matte-Taillez, O., Brochier, C., Forterre, P., and Philippe, H. (2002). Archaeal phylogeny based on ribosomal proteins. *Mol. Biol. Evol.* 19, 631–639. doi: 10.1093/oxfordjournals.molbev.a04122
- McInerney, J. O., Martin, W. F., Koonin, E. V., Allen, J. F., Galperin, M. Y., Lane, N., et al. (2011). Planctomycetes and eukaryotes: a case of analogy not homology. *Bioessays* 33, 810–817. doi: 10.1002/bies.201100045
- McInerney, J. O., O'Connell, M. J., and Pisani, D. (2014). The hybrid nature of the Eukarya and a consilient view of life on Earth. *Nat. Rev. Microbiol.* 12, 449–455. doi: 10.1038/nrmicro3271
- Mossel, E. (2003). On the impossibility of reconstructing ancestral data and phylogenies. *J. Comput. Biol.* 10, 669–676. doi: 10.1089/106652703322539015
- Mulkidjanian, A. Y., Makarova, K. S., Galperin, M. Y., and Koonin, E. V. (2007). Inventing the dynamo machine: the evolution of the F-type and V-type ATPases. *Nat. Rev. Microbiol.* 5, 892–899. doi: 10.1038/nrmicro1767
- Mushegian, A. R., and Koonin, E. V. (1996). A minimal gene set for cellular life derived by comparison of complete bacterial genomes. *Proc. Natl. Acad. Sci. U.S.A.* 93, 10268–10273. doi: 10.1073/pnas.93.19.10268
- Nasir, A., Kim, K. M., and Caetano-Anollés, G. (2015). Lokiarchaeota: eukaryote-like missing links from microbial dark matter? *Trends Microbiol.* doi: 10.1016/j.tim.2015.06.001 [Epub ahead of print].
- Narasingarao, P., Podell, S., Ugalde, J. A., Brochier-Armanet, C., Emerson, J. B., Brocks, J. J., et al. (2012). *De novo* metagenomic assembly reveals abundant novel major lineage of Archaea in hypersaline microbial communities. *ISME J.* 6, 81–93. doi: 10.1038/ismej.2011.78
- Nunoura, T., Takaki, Y., Kakuta, J., Nishi, S., Sugahara, J., Kazama, H., et al. (2011). Insights into the evolution of Archaea and eukaryotic protein modifier systems revealed by the genome of a novel archaeal group. *Nucleic Acids Res.* 39, 3204–3223. doi: 10.1093/nar/gkq1228
- Olsen, G. J., and Woese, C. R. (1997). Archaeal genomics: an overview. *Cell* 89, 991–994. doi: 10.1016/S0092-8674(00)80284-6
- Pace, N. R. (1997). A molecular view of microbial diversity and the biosphere. *Science* 276, 734–740. doi: 10.1126/science.276.5313.734
- Pace, N. R. (2006). Time for a change. *Nature* 441, 289. doi: 10.1038/441289a
- Pace, N. R., Olsen, G. J., and Woese, C. R. (1986). Ribosomal RNA phylogeny and the primary lines of evolutionary descent. *Cell* 45, 325–326. doi: 10.1016/0092-8674(86)90315-6
- Penny, D., Collins, L. J., Daly, T. K., and Cox, S. J. (2014). The relative ages of eukaryotes and akaryotes. *J. Mol. Evol.* 79, 228–239. doi: 10.1007/s00239-014-9643-y
- Perrochia, L., Guetta, D., Hecker, A., Forterre, P., and Basta, T. (2013a). Functional assignment of KEOPS/EKC complex subunits in the biosynthesis of the universal t6A tRNA modification. *Nucleic Acids Res.* 41, 9484–9499. doi: 10.1093/nar/gkt720
- Perrochia, L., Crozat, E., Hecker, A., Zhang, W., Bareille, J., Collinet, B., et al. (2013b). *In vitro* biosynthesis of a universal t6A tRNA modification in Archaea and Eukarya. *Nucleic Acids Res.* 41, 1953–1964. doi: 10.1093/nar/gks1287
- Petitjean, C., Deschamps, P., López-García, P., and Moreira, D. (2014). Rooting the Domain archaea by phylogenomic analysis supports the foundation of the new kingdom proteoarchaeota. *Genome Biol. Evol.* 7, 191–204. doi: 10.1093/gbe/evu274
- Petitjean, C., Deschamps, P., López-García, P., Moreira, D., and Brochier-Armanet, C. (2015). Extending the conserved phylogenetic core of archaea disentangles the evolution of the third domain of life. *Mol. Biol. Evol.* doi: 10.1093/molbev/msv015 [Epub ahead of print].
- Philippe, H., and Adoutte, A. (1998). “The molecular phylogeny of protozoa: solid facts and uncertainties,” in *Evolutionary Relationships Among Protozoa*, eds G. H. Coombs, K. Vickerman, M. A. Sleigh, and A. Warren (Dordrecht: Kluwer Academic Publishers), 25–56.
- Philippe, H., and Forterre, P. (1999). The rooting of the universal tree of life is not reliable. *J. Mol. Evol.* 49, 509–523. doi: 10.1007/PL00006573
- Poole, A. M. (2009). Horizontal gene transfer and the earliest stages of the evolution of life. *Res. Microbiol.* 160, 473–480. doi: 10.1016/j.resmic.2009.07.009
- Probst, A. J., Weinmaier, T., Raymann, K., Perris, A., Emerson, J. B., Rattei, T., et al. (2014). Biology of a widespread uncultivated archaeon that contributes to carbon fixation in the subsurface. *Nat. Commun.* 5, 5497. doi: 10.1038/ncomms6497
- Puigbo, P., Wolf, Y. I., and Koonin, E. V. (2013). Seeing the Tree of Life behind the phylogenetic forest. *BMC Biol.* 11:46. doi: 10.1186/1741-7007-11-46
- Ramulu, H. G., Groussin, M., Talla, E., Planell, R., Daubin, V., and Brochier-Armanet, C. (2014). Ribosomal proteins: toward a next generation standard for prokaryotic systematics? *Mol. Phylogenet. Evol.* 75, 103–117. doi: 10.1016/j.ympev.2014.02.013
- Raoult, D., and Forterre, P. (2008). Redefining viruses: lessons from Mimivirus. *Nat. Rev. Microbiol.* 6, 315–319. doi: 10.1038/nrmicro1858
- Raymann, K., Brochier-Armanet, C., and Gribaldo, S. (2015). The two-domains tree of life is linked to a new root for the Archaea. *Proc. Natl. Acad. Sci.* 112, 6670–6675. doi: 10.1073/pnas.1420858112
- Raymann, K., Forterre, P., Brochier-Armanet, C., and Gribaldo, S. (2014). Global phylogenomic analysis disentangles the complex evolutionary history of DNA replication in archaea. *Genome Biol. Evol.* 6, 192–212. doi: 10.1093/gbe/evu004
- Rinke, C., Schwientek, P., Sczyrba, A., Ivanova, N. N., Anderson, I. J., Cheng, J. F., et al. (2013). Insights into the phylogeny and coding potential of microbial dark matter. *Nature* 499, 431–437. doi: 10.1038/nature12352
- Rivera, M. C., and Lake, J. A. (2004). The ring of life provides evidence for a genome fusion origin of eukaryotes. *Nature* 431, 152–155. doi: 10.1038/nature02848
- Rochette, N. C., Brochier-Armanet, C., and Gouy, M. (2014). Phylogenomic test of the hypotheses for the evolutionary origin of eukaryotes. *Mol. Biol. Evol.* 31, 832–845. doi: 10.1093/molbev/mst272
- Santarella-Mellwig, R., Franke, J., Jaedicke, A., Gorjanacz, M., Bauer, U., Budd, A., et al. (2010). The compartmentalized bacteria of the planctomycetes-verrucomicrobia-chlamydiae superphylum have membrane coat-like proteins. *PLoS Biol.* 8:e1000281. doi: 10.1371/journal.pbio.1000281
- Sapp, J. (2009). *The New Foundation of Evolution*. New York: Oxford University Press.
- Soler, N., Marguet, E., Cortez, D., Desnoues, N., Keller, J., van Tilbeurgh, H., et al. (2010). Two novel families of plasmids from hyperthermophilic archaea encoding new families of replication proteins. *Nucleic Acids Res.* 38, 5088–5104. doi: 10.1093/nar/gkq236
- Spang, A., Hatzenpichler, R., Brochier-Armanet, C., Rattei, T., Tischler, P., Speck, E., et al. (2010). Distinct gene set in two different lineages of ammonia-oxidizing archaea supports the phylum Thaumarchaeota. *Trends Microbiol.* 18, 331–340. doi: 10.1016/j.tim.2010.06.003
- Spang, A., Saw, J. H., Jorgensen, S. L., Zaremba-Niedzwiedzka, K., Martijn, J., Lind, A. E., et al. (2015). Complex archaea that bridge the gap between prokaryotes and eukaryotes. *Nature* 521, 173–179. doi: 10.1038/nature14447

- Stechmann, A., and Cavalier-Smith, T. (2003). The root of the eukaryote tree pinpointed. *Curr Biol.* 13:R665–R666. doi: 10.1016/S0960-9822(03)00602-X
- Stetter, K. O. (1996). Hyperthermophiles in the history of life. *Ciba Found. Symp.* 202, 1–10.
- Thiaville, P. C., El Yacoubi, B., Perrochia, L., Hecker, A., Prigent, M., Thiaville, J. J., et al. (2014). Cross kingdom functional conservation of the core universally conserved threonylcarbamoyl adenosine tRNA synthesis enzymes. *Eukaryot. Cell* 13, 1222–1231. doi: 10.1128/EC.00147-14
- Thiergart, T., Landan, G., and Martin, W. F. (2014). Concatenated alignments and the case of the disappearing tree. *BMC Evol. Biol.* 14:266. doi: 10.1186/s12862-014-0266-0
- Urbonavicius, J., Auxilien, S., Walbott, H., Trachana, K., Golinelli-Pimpaneau, B., Brochier-Armanet, C., et al. (2008). Acquisition of a bacterial RumA-type tRNA(uracil-54, C5)-methyltransferase by Archaea through an ancient horizontal gene transfer. *Mol. Microbiol.* 67, 323–335. doi: 10.1111/j.1365-2958.2007.06047.x
- Wan, L. C., Mao, D. Y., Neculai, D., Strecker, J., Chiovitti, D., Kurinov, I., et al. (2013). Reconstitution and characterization of eukaryotic N6-threonylcarbamoylation of tRNA using a minimal enzyme system. *Nucleic Acids Res.* 41, 6332–6346. doi: 10.1093/nar/gkt322
- Williams, T. A., Foster, P. G., Cox, C. J., and Embley, T. M. (2013). An archaeal origin of eukaryotes supports only two primary domains of life. *Nature* 504, 231–236. doi: 10.1038/nature12779
- Williams, T. A., and Embley, T. M. (2014). Archaeal “dark matter” and the origin of eukaryotes. *Genome Biol. Evol.* 6, 474–481. doi: 10.1093/gbe/evu031
- Woese, C. (1998). The universal ancestor. *Proc. Natl. Acad. Sci. U.S.A.* 95, 6854–6859. doi: 10.1073/pnas.95.12.6854
- Woese, C. R. (2000). Interpreting the universal phylogenetic tree. *Proc. Natl. Acad. Sci. U.S.A.* 97, 8392–8396. doi: 10.1073/pnas.97.15.8392
- Woese, C. R. (2002). On the evolution of cells. *Proc. Natl. Acad. Sci. U.S.A.* 99, 8742–8747. doi: 10.1073/pnas.132266999
- Woese, C. R., and Fox, G. E. (1977). The concept of cellular evolution. *J. Mol. Evol.* 10, 1–6. doi: 10.1007/BF01796132
- Woese, C. R., Kandler, O., and Wheelis, M. L. (1990). Towards a natural system of organisms: proposal for the domains Archaea, Bacteria, and Eucarya. *Proc. Natl. Acad. Sci. U.S.A.* 87, 4576–4579. doi: 10.1073/pnas.87.12.4576
- Yutin, N., Puigbò, P., Koonin, E. V., and Wolf, Y. I. (2012). Phylogenomics of prokaryotic ribosomal proteins. *PLoS ONE* 7:e36972. doi: 10.1371/journal.pone.0036972
- Zhaxybayeva, O., Swithers, K. S., Lapierre, P., Fournier, G. P., Bickhart, D. M., DeBoy, R. T., et al. (2009). On the chimeric nature, thermophilic origin, and phylogenetic placement of the Thermotogales. *Proc. Natl. Acad. Sci. U.S.A.* 106, 5865–5870. doi: 10.1073/pnas.0901260106

Conflict of Interest Statement: The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2015 Forterre. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Archaeal membrane-associated proteases: insights on *Haloferax volcanii* and other haloarchaea

María I. Giménez*, Micaela Cerletti and Rosana E. De Castro*

Instituto de Investigaciones Biológicas, Facultad de Ciencias Exactas y Naturales, Universidad Nacional de Mar del Plata, Consejo Nacional de Investigaciones Científicas y Técnicas, Mar del Plata, Argentina

Edited by:

Mechthild Pohlschroder, University of Pennsylvania, USA

Reviewed by:

Kelly Bidle, Rider University, USA

Maria-Jose Bonete, University of Alicante, Spain

***Correspondence:**

Maria I. Giménez and Rosana E. De Castro, Instituto de Investigaciones Biológicas, Facultad de Ciencias Exactas y Naturales, Universidad Nacional de Mar del Plata, Consejo Nacional de Investigaciones Científicas y Técnicas, Funes 3250 4to nivel, Mar del Plata 7600, Argentina
e-mail: migimen@mdp.edu.ar; decastro@mdp.edu.ar

The function of membrane proteases range from general house-keeping to regulation of cellular processes. Although the biological role of these enzymes in archaea is poorly understood, some of them are implicated in the biogenesis of the archaeal cell envelope and surface structures. The membrane-bound ATP-dependent Lon protease is essential for cell viability and affects membrane carotenoid content in *Haloferax volcanii*. At least two different proteases are needed in this archaeon to accomplish the posttranslational modifications of the S-layer glycoprotein. The rhomboid protease Rholl is involved in the N-glycosylation of the S-layer protein with a sulfoquinovose-containing oligosaccharide while archaeosortase ArtA mediates the proteolytic processing coupled-lipid modification of this glycoprotein facilitating its attachment to the archaeal cell surface. Interestingly, two different signal peptidase I homologs exist in *H. volcanii*, Sec11a and Sec11b, which likely play distinct physiological roles. Type IV prepilin peptidase PibD processes flagellin/pilin precursors, being essential for the biogenesis and function of the archaeallum and other cell surface structures in *H. volcanii*.

Keywords: archaeal proteolysis, membrane-associated proteases, *Haloferax volcanii*, cell envelope, S-layer glycoprotein

INTRODUCTION

Membrane-associated proteases participate in a variety of processes essential for cell physiology including membrane protein quality control, processing of exported and/or membrane-anchored polypeptides, regulatory circuits, cell-signaling, the stress response and assembly of cell surface structures (Akiyama, 2009; Dalbey et al., 2011; Schneewind and Missiakas, 2012; Konovalova et al., 2014). Their targets are mainly membrane-bound or secreted proteins which account for 20–30% of total proteins encoded in most genomes (Wallin and von Heijne, 1998) and include membrane receptors, structural proteins, transporters and enzymes such as transferases, oxidoreductases, and hydrolases.

Integral membrane proteases comprise two distinct groups. The first group is represented by peptidases anchored to the cytoplasmic membrane that exert their catalytic activity in an aqueous compartment (cytoplasm, periplasm, or extracellular milieu) either at the aqueous-membrane boundary or after the substrate has been released or extracted from the membrane. Within this category are signal peptidases (SP), site 1 proteases (S1P) or sheddases, signal peptide hydrolases SPPA, HtpX, sortases and the energy-dependent proteases FtsH and LonB. The second group is represented by the so-called intramembrane cleaving proteases (ICliPs) which have their active sites immersed in the hydrophobic environment of the membrane (Wolfe, 2009; Dalbey et al., 2012). This group includes GxGD-aspartyl proteases (eukaryal signal peptide peptidase SPP and presenilin families), rhomboids and site 2 proteases (S2P).

Archaea, one of the three domains of life, are widespread in nature but predominate in environments with extreme values of pH, temperature, salt concentration and pressure (Robertson et al.,

2005). Studies on archaeal biology are encouraged as they provide the opportunity to better understand cell physiology as well as extend the resources for biotechnology.

The genome sequences of archaea show that these unusual organisms encode a variety of proteolytic enzymes some of which have been characterized (Ward et al., 2002; Maupin-Furlow et al., 2005; De Castro et al., 2006; Ng et al., 2007). Most of the membrane protease families found in bacteria and/or eukaryotic cells also occur in archaea, however, the role of these enzymes in the context of the archaeal cell is poorly understood. In the last decade a number of studies have started to advance the knowledge on this field (see references in Table 1). This mini review describes what is known about proteases associated with the cell surface of archaeal cells on the basis of complete genome sequences and biochemical and/or genetic studies. Emphasis will be placed on the proteolytic enzymes affecting the cell envelope and surface structures of the euryarchaeon *Haloferax volcanii* and other haloarchaea. *H. volcanii* grows in a wide range of salinity (1.5–3.5 M NaCl) and is a model organism to study archaeal biology due to a number of advantages including the simplicity of its culture conditions, availability of complete genome sequences and feasibility of its genetic manipulation.

MEMBRANE-ASSOCIATED PROTEASES OF ARCHAEA

An overview of the repertoire of membrane proteases that occur in archaeal cells is shown in Table S1 based on *in silico* examination of the complete genome sequences of some representative archaea members. Some protease families are widely represented among archaeal genomes such as HtpX homologs, LonB, SP, and Site 2

Table 1 | Predicted membrane proteases of representative archaeal genomes.

Protease name	MEROPS family	Cleavage site	Topology	Description	Organism	Reference
Signal peptidases (SP)				Cleaves signal peptides of preproteins		
SPI	S26B	Active site on external face of cytoplasmic membrane		Resemble eukaryotic SPI. Cleaves the majority of signal peptides from preproteins exported by Sec (Tat?) system		
SPI (MJ0260)			2 TMS	First SPI characterized from archaea	<i>Methanococcus voltae</i>	Ng and Jarrell (2003); Ng et al. (2007)
Sec11a (HVO_2603); Sec11b (HVO_0002)			1 TMS	Both active <i>in vitro</i> , different efficiency for substrate cleavage. Sec11b is essential for viability	<i>Halofexax volcanii</i>	Fine et al. (2006); Fink-Lavi and Eichler (2008)
TFPP-like (SPIII)		Active site on internal face of cytoplasmic membrane		Novel family of aspartyl proteases (GxGD). Cleaves signal peptides of preflagellins prior to incorporation into the archaeallum		
PibD (SSO0131)	A24		5 TMS	Broad substrate range. Cleaves signal peptides of preflagellins and sugar-binding proteins ("bindosome")	<i>Sulfolobus solfataricus</i>	Albers et al. (2003, 2006); Szabo et al. (2006); Ng et al. (2009)
(Saci_0139)			6 TMS	<i>In vivo</i> activity assays led to redefine PibD as GxHyD (Hy = hydrophobic residue) rather than a GxGD group of proteases	<i>Sulfolobus acidocaldarius</i>	Henche et al. (2014)
(HVO_2993)			6 TMS	Cleaves preflagellins and type IV pilin-like proteins. Required for swimming motility (flagella) and surface adhesion (type V pilin-like structures)	<i>H. volcanii</i>	Tripepi et al. (2010)
EppA (MMP0232; MJ0835.2)	A24A		9–10 TMS	Cleaves signal peptides of pilin-like proteins containing a DUF361 domain	<i>Methanococcus maripaludis</i> , <i>Methanococcus jannaschii</i>	Szabo et al. (2007)
FlaK (MvoL_0164)	A24B		5 TMS	Inactivation of FlaK generates non-flagellated cells. Residues Ser52, His122 and Asp148 crucial for protease activity	<i>M. voltae</i>	Bardy and Jarrell (2003); Bardy et al. (2005)
(MMP0555)			5 TMS	Displays preflagellin peptidase activity. placeCityCrystal structure is available	<i>M. maripaludis</i>	Bardy and Jarrell (2002); Hu et al. (2011)

(Continued)

Table 1 | Continued

Protease name	MEROPS family	Cleavage site	Topology	Description	Organism	Reference
Signal peptide peptidases (SPP)						
SppA (TK1164)	S49	Active site at the soluble extracytoplasmic side	1 TMS	First insight into catalytic mechanism of prokaryotic SPPs	<i>Thermococcus kodakarensis</i>	Matsumi et al. (2005, 2006)
Presenilin/SPP homolog	A22	Active site at the membrane plane (IClP)	9 TMS	Regulatory intramembrane proteolysis (RIP). Aspartyl proteases (GxGD) Crystal structure determined in an archaeon. Insights on catalytic mechanism of Presenilin and SPP intramembrane proteases	<i>Methanoculleus marisnigri</i>	Li et al. (2013)
ATP-dependent proteases						
LonB	S16	ATPase and protease domains facing cytoplasm	2 TMS	ATP-independent activity on unfolded substrates; ATP-dependent activity on folded proteins	<i>T. kodakarensis</i>	Fukui et al. (2002)
(TK1264)			2 TMS	Confirmation of S-K catalytic dyad. Recombinant LonB protease showed ATPase and protease activities	<i>Thermoplasma acidophilum</i>	Besche and Zwickl (2004); Besche et al. (2004)
(Ta1081)			1 TMS	Crystal structure of LonB proteolytic domain solved	<i>Archaeoglobus fulgidus</i>	Botos et al. (2005)
(AF0364)			2 TMS	Crystal structure 2.0 Å resolution solved	<i>Thermococcus onnurineus</i>	Cha et al. (2010)
(TON_0529)			2 TMS	DNA-binding capacity <i>in vitro</i> . Can complement a LonB mutant in <i>H. volcanii</i> suggesting functional conservation of LonB	<i>Natrialba magadii</i>	Sastre et al. (2011); Cerletti et al. (2014)
(Nmag_2822)			2 TMS	Essential for viability. Quality control of proteins. Suboptimal LonB expression affects the content of membrane carotenoids and other lipids	<i>H. volcanii</i>	Cerletti et al. (2014)
(HVO_0783)						

(Continued)

Table 1 | Continued

Protease name	MEROPS family	Cleavage site	Topology	Description	Organism	Reference
Rhomboids	S54	Active site at the membrane plane (IClP)	6TMS	Regulatory intramembrane proteolysis (RIP)		
	Rho II (HVO_0727)	A null mutant in <i>rhol</i> gene showed reduced mobility and higher sensitivity to novobiocin. The mutation affected N-glycosylation of the S-layer glycoprotein with a sulfoquinovose containing oligosaccharide			<i>H. volcanii</i>	Parente et al. (2014)
Site 2 proteases (S2P)	M50B	Active site at the membrane plane (IClP)	6TMS	Regulatory intramembrane proteolysis (RIP)		
	(MJ_0392)	The crystal structure of the transmembrane core was determined (3B4R)			<i>M. jannaschii</i>	Feng et al. (2007)
Archaeosortases	N/A	Predicted active-site Cys oriented toward the extracellular side	7TMS	Involved in C-terminal processing of the S-layer glycoprotein	<i>H. volcanii</i>	Abdul Hallim et al. (2013)
	(HVO_0915)					
Thermopsine (SSO2194)	A5	Active site oriented to the extracellular milieu	1TMS	Overproduced in peptide enriched media. Catalytic domain followed by PKD and Y_Y_Y domains	<i>S. solfataricus</i>	Cannio et al. (2010)
CAAX prenyl proteases	M79	In eukaryotic cells they are involved in prenylation of proteins for membrane localization				
	Abi (HVO_0784)	Putative CAAX protease homolog. Forms a transcription unit with <i>lnb</i> gene. A null mutant showed no evident phenotypes	5TMS		<i>H. volcanii</i>	Cerletti et al. (2014)

proteases (S2Ps) whereas others are restricted to a limited number of organisms (for instance the protease families A5, M10, and PrsW protease). **Table 1** describes the membrane proteases that have been experimentally characterized from the *Archaea* domain. Some of them have been studied in more detail (SPI and TFP-like SP) and at least a few of their endogenous substrates have been identified (e.g., preflagellins, prepilins, and sugar-binding proteins for TFP-like peptidases). However, most families have been examined to a limited extent or remain uncharacterized, and their biological relevance and/or targets are unknown (e.g., rhomboids, LonB, CAAX prenyl protease homologs, S2Ps).

The crystal structures of a number of archaeal membrane proteases have been solved (*Methanococcus maripaludis* FlaK; *Thermococcus onnurineus* and *Archaeoglobus fulgidus* LonB proteolytic domains; S2P transmembrane segments (TMSs) core from *Methanococcus jannaschii*; MCMJR1 peptidase from *Methanoculleus marisnigri*) providing valuable structure/function insights on these protease families (see **Table 1** for references).

MEMBRANE PROTEASES IMPLICATED IN THE ASSEMBLY OF THE ARCHAEOAL CELL ENVELOPE AND SURFACE STRUCTURES

Probably one of the most distinctive features of archaea is their ability to survive in environments with extremely adverse conditions that are lethal for most life forms. To this end, they have adapted their physiology and cellular structures. One such instance is the cell envelope. The archaeal cell envelope is composed of an atypical cellular membrane constituted by isoprenyl ether glycerol phospholipids surrounded by surface S-layer proteins as the major (or sole) component of the cell wall (Albers and Meyer, 2011). These structures maintain the cellular integrity and functionality as well as serve as a shell to cope with the harsh conditions predominating in their surroundings (Claus et al., 2005).

In addition to the S-layer, archaea show very diverse and complex cell surface structures (reviewed in Lassak et al., 2012). The biogenesis of the appendages composed of bacterial type IV pilin subunits, the pili and the archaeal flagellum or archaellum, has been characterized to some extent. These structures play important roles in cell motility as well as in surface attachment, DNA exchange and cell-cell interaction.

Haloarchaea, a very diverse and probably the best characterized group of archaea, flourish in habitats with high salinity (> 2M NaCl) and intense solar irradiation. In the haloarchaeon *H. volcanii* the structure and maturation of the S-layer glycoprotein as well as the biogenesis of pili and flagella have been examined (Jarrell et al., 2010; Kaminski et al., 2013; Kandiba et al., 2013; Tripepi et al., 2013; Esquivel and Pohlschroder, 2014). The adequate localization and functionality of these structures requires the participation of different families of proteases which are immersed in the context of the cytoplasmic membrane. Below we describe the recent advances on the membrane-associated proteases involved in the processes leading to the assembly of the cell envelope and surface structures in the euryarchaeon *H. volcanii*. The currently available information is summarized in **Figure 1**.

PROTEASES INVOLVED IN THE BIOGENESIS OF THE CYTOPLASMIC MEMBRANE AND SECRETION OF PREPROTEINS

The quality control of membrane proteins is essential for proper cell physiology. In bacteria and eukaryotic organelles a major role in this process is performed by the energy-dependent membrane protease FtsH (Dalbey et al., 2012; Langklotz et al., 2012). Archaea possess only two ATP-dependent proteases: the 20S proteasome (soluble enzyme) and an unusually membrane-bound version of the Lon protease (LonB). The archaeal LonB probably resembles functionally to the FtsH protease which is absent in archaea. LonB has been biochemically and/or structurally characterized in several archaeal members (**Table 1**). In agreement with the genomic prediction, LonB has been immunolocalized in association with the cell membrane in the haloarchaea *Natrialba magadii* and *H. volcanii*. The recombinant protease derived from *N. magadii* (NmLon) showed DNA binding capacity *in vitro*, a feature in common with LonA proteases (Sastre et al., 2011). As FtsH is for *Escherichia coli* (Langklotz et al., 2012), LonB is essential for viability of *H. volcanii* cells. On the other hand suboptimal expression of this protease affects growth rate, cell shape, antibiotic sensitivity, and lipid composition (Cerletti et al., 2014). Also, *H. volcanii* mutant cells deficient in Lon content are more sensitive to puromycin compared to wild type cells suggesting that LonB is involved in the disposal of abnormal proteins. A distinctive feature of haloarchaea is the presence of red membrane-bound carotenoid pigments (C50-bacterioruberins) which serve to protect their macromolecules from the damaging effects of UV light (Khanafari et al., 2010). Interestingly, the cellular content of bacterioruberins dramatically increased in *H. volcanii* mutant cells with a suboptimal Lon concentration while overexpression of this protease rendered the cells colorless (Cerletti et al., 2014). This observation suggests that LonB controls carotenoid biosynthesis in *H. volcanii* probably by degrading enzyme/s involved in this pathway. It is likely that deregulation of the cellular concentration of bacterioruberins and other lipids affects membrane stability contributing to the lethal phenotype of the *lon* knock out mutant.

Signal peptidases are central in the protein secretion process as they remove signal peptides from secretory and membrane-bound polypeptides. In archaea, type I signal peptidase (SPI), type IV prepilin peptidase (TFPP)-like enzymes and signal peptide peptidase (SPP) have been characterized. A detailed description on the distribution and properties of these enzymes has been previously reported (Ng et al., 2007). SPIs process the majority of pre-proteins that are translocated through the general secretion pathway (Sec), however, whether this enzyme also cleaves Tat signal peptides remains to be demonstrated. Like all members of the SPI family, archaeal SPIs are serine proteases and based on studies performed in SPI from *M. voltae* (Ng and Jarrell, 2003) and *H. volcanii* (Fink-Lavi and Eichler, 2008) the catalytic mechanism of the archaeal SPI homolog seems to rely on a Ser/His/Asp tryad resembling the eukaryotic enzyme. In *H. volcanii* two different SPIs with distinct efficiency for substrate cleavage exist, Sec11a and Sec11b, however, only Sec11b is essential for viability (Fine et al., 2006). It is likely that these enzymes exert different roles and/or cleave distinct substrates *in vivo*.

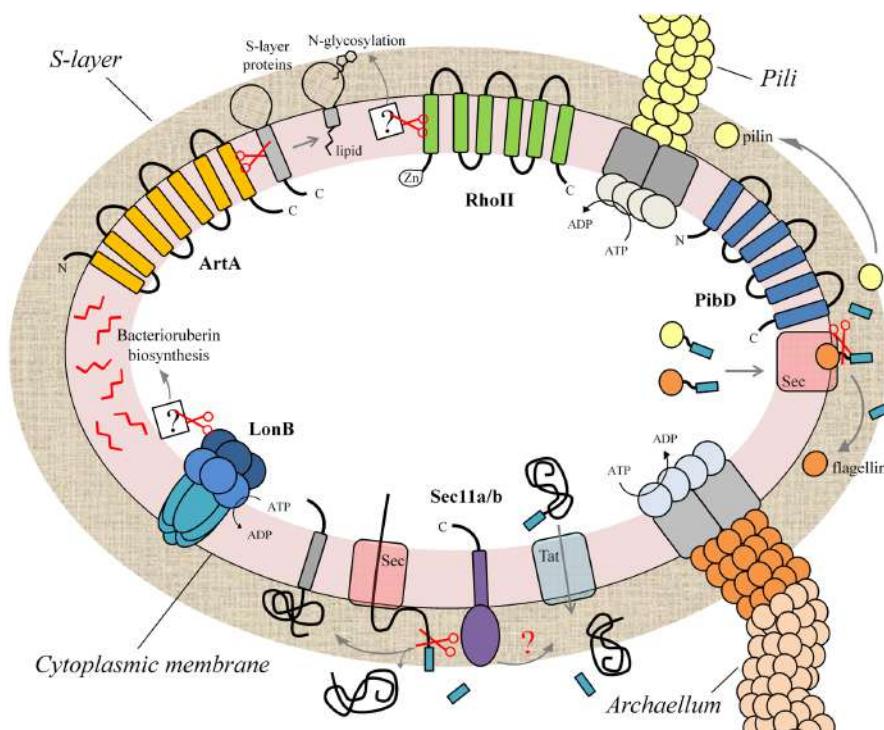


FIGURE 1 | Membrane proteases implicated in the biogenesis of the cell envelope and extracellular appendages of the haloarchaeon *Haloferax volcanii*. Schematic representation of different protease families that participate in processes leading to the biogenesis of the cell envelope (cytoplasmic membrane, S-layer) and cell surface structures (archaellum, pili) in *H. volcanii*, and probably in other archaea. LonB is implicated in bacterioruberin biosynthesis and protein quality control. The signal peptidases

Sec11a/b process the signal peptide of secreted and membrane proteins translocated through the Sec pathway. PibD cleaves preflagellins and prepilins, the protein components of the archaellum and pili. The rhomboid protease RhoII is involved in the N-glycosylation of the S-layer protein with sulfoquinovose-containing oligosaccharide while archaeosortase ArtA mediates the proteolytic processing coupled-lipid modification of this glycoprotein facilitating its attachment to the archaeal cell surface.

SPII removes signal peptides from lipoproteins. Although there are numerous proteins in archaea that contain signal peptides with the lipobox motif, including several predicted to be secreted via the Tat pathway, homologs of bacterial SPII have not been identified in archaeal genomes (Giménez et al., 2007). Thus, it has been proposed that a distinct enzyme may exist in archaea to process prelipoproteins (Ng et al., 2007).

PROTEASES INVOLVED IN MATURATION OF THE CELL WALL (S-LAYER GLYCOPROTEIN)

In *H. volcanii* the S-layer glycoprotein is the sole structure that constitutes the cell wall. This protein has been used to examine the molecular/structural adaptations of haloarchaeal proteins to high salt and has served as a model to study protein glycosylation in archaea (Eichler et al., 2013; Jarrell et al., 2014). In haloarchaea, maturation of the S-layer glycoprotein requires at least three different types of posttranslational modifications: glycosylation, proteolytic cleavage and isoprenylation (Konrad and Eichler, 2002; Eichler, 2003). The glycosylation process of the S-layer has been recently reviewed (Eichler et al., 2013).

Sortases are cysteine proteases from Gram-positive bacteria that “sort” proteins to the cell surface by covalently joining them to the cell wall or polymerize pilins to build pili (Proft and Baker, 2009; Clancy et al., 2010; Hendrickx et al., 2011; Spirig

et al., 2011). These enzymes modify surface proteins by recognizing and cleaving a sorting signal located either in the N or C-terminus of the target protein. Many genomes in bacteria and archaea encode proteins containing a C-terminal domain with structural similarity to the C-terminus of sortase substrates. These proteins coexist in these genomes with at least one member of the protease families denoted as exosortases (bacteria) or archaeosortases (archaea). Exo and archaeosortases are polytopic membrane proteins with no sequence homology to bacterial sortases. However, they contain the conserved cysteine, arginine, and histidine residues found in the active site of sortases suggesting that they may perform similar functions (Haft et al., 2012). Recently it was reported that *H. volcanii* mutant cells with a deletion in the archaeosortase gene *artA* showed growth defects (which were more evident under low-salt conditions), alterations in cell shape and the S-layer organization, impaired motility and decreased conjugation rates (Abdul Halim et al., 2013). This work demonstrated that ArtA is involved in C-terminal processing of the S-layer glycoprotein suggesting that archaeosortases are functional homologs of bacterial sortases. Considering the location of the archaeosortase recognition sequence (PGF) immediately following the TMS of the substrate protein, it was proposed that this enzyme may facilitate the covalent attachment of target proteins (e.g., S-layer glycoprotein) to a membrane lipid in

contrast to sortases which attach proteins to the growing cell wall.

Rhomboids are membrane serine proteases involved in regulatory intramembrane proteolysis (RIP) and are conserved in the three domains of life (Lemberg, 2013). The catalytic mechanism of rhomboids relies on a Ser/His dyad located in different TMS of the protease to cleave membrane protein substrates. In eukaryotic cells the functions of this protease family are very diverse and include cell-cell signaling, development, apoptosis, organelle integrity and parasite invasion (reviewed in Freeman, 2014). The relevance of rhomboids in the prokaryotic cell physiology is scarcely understood. In bacteria, rhomboid null mutants show phenotypes that may be related to defective cell envelope and/or cell-surface structures. In *Bacillus subtilis*, a mutant strain in the rhomboid homolog YqqP displayed a slight decrease in glucose uptake and a defect in cell division leading to the formation of filamentous cells (Mesak et al., 2004); *Mycobacterium smegmatis* rhomboid mutants showed reduced capacity for biofilm formation and increased sensitivity to antibiotics (Kateete et al., 2012). So far only TatA, a protein component of the Tat translocon in the pathogenic bacterium *Providencia stuartii*, has been experimentally confirmed as a rhomboid substrate (Stevenson et al., 2007). In this organism the rhomboid protease AarA cleaves an N-terminal extension of TatA which in turn allows for secretion of an unknown quorum sensing signal. Archaea appear to encode various sequences for rhomboid proteases (Table S1). In haloarchaea, homologs with various topologies can be found including proteins with six or more TMS as well as unusual rhomboids containing an AN-1 Zn-finger domain at the N-terminus. *H. volcanii* has two putative genes for rhomboids, RhoI (nine TMS) and RhoII (six TMS, with N-terminal AN-1 Zn finger domain). A knock-out mutant of *rholI* in *H. volcanii* displayed mild defects in motility and novobiocin sensitivity. This mutant strain was also affected in the glycosylation of the S-layer. In *H. volcanii* wild type cells the S-layer glycoprotein Asn732 is bound to an oligosaccharide containing at least 6 repeating units of sulfoquinovose-hexose (SQ-Hex) while in the mutant strain this residue contained only two SQ-Hex suggesting that RhoII controls (directly or indirectly) the protein glycosylation process in *H. volcanii* (Parente et al., 2014).

PROTEASES INVOLVED IN THE BIOGENESIS OF CELL SURFACE APPENDAGES

In bacteria, the precursors of type IV pilins and related pseudopilins are processed by a special enzyme belonging to a novel aspartic acid protease family, the type IV prepilin signal peptidase (SPIV/TFPP; Ng et al., 2009). In contrast to SPI and SPII, this enzyme cleaves the signal peptides directly after the *n*-region leaving the *h*-region bound to the mature protein facilitating anchoring/assembly of pilin subunits onto the cell surface (Ng et al., 2007). Archaea encode TFPP-like proteins and they have been studied with regard to their role in the assembly of the structures composing the motility apparatus (Table 1). The archaeallum is composed of unique proteins that are unrelated to bacterial flagellins. Archaeal preflagellins contain short signal peptides at the N-terminus which are similar to those of bacterial type IV pilins, the protein components of pili. These filamentous surface

structures facilitate twitching motility in bacteria. TFPP-like proteases process the signal peptides of archaeal preflagellins. The enzymes present in *M. maripaludis* and *Methanococcus voltae* (FlaK), *Sulfolobus solfataricus*, and *H. volcanii* (PibD) are the most extensively characterized TFPPs of archaea (see references in Table 1). FlaK and PibD show some divergences including the length of the signal peptide, key amino acid residues surrounding the cleavage site as well as substrate preference (Ng et al., 2009). PibD from *S. solfataricus* and *H. volcanii* has a broader substrate selection than FlaK, as, in addition to preflagellins, these enzymes can mature prepilins (Albers et al., 2003; Triepi et al., 2010). *S. solfataricus* PibD also processes certain sugar-binding proteins of the “bindosome,” filamentous-like structures that extend from the cell surface (Albers et al., 2003; Szabo et al., 2006).

The *H. volcanii* genome encodes flagellins and contains genes for other type IV pilin-like proteins. Triepi et al. (2010) showed that deletion of *pibD* disrupted preflagellins processing and prevented maturation of type IV pilin-like proteins. The mutant cells were non-motile and were unable to adhere to a glass surface. These results suggest that PibD is needed for maturation of preflagellins and other type IV pilin-like proteins in *H. volcanii*.

Recently, based on *in vivo* analysis of the catalytic activity of *Sulfolobus acidocaldarius* PibD, TFPPs were renamed as GxHyD group of proteases (rather than DxGD; Henche et al., 2014).

CONCLUDING REMARKS

In prokaryotes the assembly and composition of cell surface structures are essential for the adjustment to the varying conditions of the environment and to interact with their surroundings (e.g., establish cell-cell and/or cell-substrate contacts). In the haloarchaeon *H. volcanii*, several membrane-associated proteases are implicated in different processes (protein secretion, processing and sorting) leading to the biogenesis of the cell wall and extracellular appendages (Figure 1), highlighting the importance of these enzymes in the adaptation and interaction of archaea with their environment.

Structural analysis of archaeal membrane proteases (Flak and GxGD proteases) have advanced the knowledge on the catalytic and molecular mechanism of intramembrane cleaving proteases. This will help to understand the mechanism of the eukaryotic homologous enzymes which are implicated in human physiology (regulation of immune response) and/or in the development of diseases (e.g. Alzheimer).

There are still many open questions in this field: e.g., endogenous substrates of most membrane proteases are unknown. Efforts should continue to better understand the role of membrane proteases in archaeal physiology.

AUTHOR CONTRIBUTIONS

All authors made substantial contributions to the acquisition, analysis and interpretation of data for this review. All authors critically reviewed and edited the manuscript, and approved the final version before submission to publication. All authors agree to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

ACKNOWLEDGMENTS

Research at De Castro's laboratory is supported by grants from the National Council of Scientific and Technological Research (CONICET) and the National University of Mar del Plata (UNMDP) Argentina. Micaela Cerletti is a PhD student at the UNMDP supported by a research fellow from CONICET; María I. Giménez and Rosana E. De Castro are researchers from CONICET.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <http://www.frontiersin.org/journal/10.3389/fmicb.2015.00039/abstract>

REFERENCES

- Abdul Halim, M. F., Pfeiffer, F., Zou, J., Frisch, A., Haft, D., Wu, S., et al. (2013). *Haloferax volcanii* archaeosortase is required for motility, mating, and C-terminal processing of the S-layer glycoprotein. *Mol. Microbiol.* 88, 1164–1175. doi: 10.1111/mmi.12248
- Akiyama, Y. (2009). Quality control of cytoplasmic membrane proteins in *Escherichia coli*. *J. Biochem.* 146, 449–454. doi: 10.1093/jb/mvp071
- Albers, S. V., and Meyer, B. H. (2011). The archaeal cell envelope. *Nat. Rev. Microbiol.* 9, 414–426. doi: 10.1038/nrmicro2576
- Albers, S. V., Szabo, Z., and Driessens, A. J. (2003). Archaeal homolog of bacterial type IV prepilin signal peptidases with broad substrate specificity. *J. Bacteriol.* 185, 3918–3925. doi: 10.1128/JB.185.13.3918-3925.2003
- Albers, S. V., Szabo, Z., and Driessens, A. J. (2006). Protein secretion in the archaea: multiple paths towards a unique cell surface. *Nat. Rev. Microbiol.* 4, 537–547. doi: 10.1038/nrmicro1440
- Bardy, S. L., and Jarrell, K. F. (2002). FlaK of the archaeon *Methanococcus maripaludis* possesses preflagellin peptidase activity. *FEMS Microbiol. Lett.* 208, 53–59. doi: 10.1111/j.1574-6968.2002.tb11060.x
- Bardy, S. L., and Jarrell, K. F. (2003). Cleavage of preflagellins by an aspartic acid signal peptidase is essential for flagellation in the archaeon *Methanococcus voltae*. *Mol. Microbiol.* 50, 1339–1347. doi: 10.1046/j.1365-2958.2003.03758.x
- Bardy, S. L., Ng, S. Y., Carnegie, D. S., and Jarrell, K. F. (2005). Site-directed mutagenesis analysis of amino acids critical for activity of the type I signal peptidase of the archaeon *Methanococcus voltae*. *J. Bacteriol.* 187, 1188–1191. doi: 10.1128/JB.187.3.1188-1191.2005
- Besche, H., Tamura, N., Tamura, T., and Zwickl, P. (2004). Mutational analysis of conserved AAA+ residues in the archaeal Lon protease from *Thermoplasma acidophilum*. *FEBS Lett.* 574, 161–166. doi: 10.1016/j.febslet.2004.08.021
- Besche, H., and Zwickl, P. (2004). The *Thermoplasma acidophilum* Lon protease has a Ser-Lys dyad active site. *Eur. J. Biochem.* 271, 4361–4365. doi: 10.1111/j.1432-1033.2004.04421.x
- Botos, I., Melnikov, E. E., Cherry, S., Kozlov, S., Makhovskaya, O. V., Tropea, J. E., et al. (2005). Atomic-resolution crystal structure of the proteolytic domain of *Archaeoglobus fulgidus* Lon reveals the conformational variability in the active sites of Lon proteases. *J. Mol. Biol.* 351, 144–157. doi: 10.1016/j.jmb.2005.06.008
- Cannio, R., Catara, G., Fiume, I., Balestrieri, M., Rossi, M., and Palmieri, G. (2010). Identification of a cell-bound extracellular protease overproduced by *Sulfolobus solfataricus* in peptide-rich media. *Protein Pept. Lett.* 17, 78–85. doi: 10.2174/092986610789909458
- Cerletti, M., Martinez, M. J., Giménez, M. I., Sastre, D. E., Paggi, R. A., and De Castro, R. E. (2014). The LonB protease controls membrane lipids composition and is essential for viability in the extremophilic haloarchaeon *Haloferax volcanii*. *Environ. Microbiol.* 16, 1779–1792. doi: 10.1111/1462-2920.12385
- Cha, S. S., An, Y. J., Lee, C. R., Lee, H. S., Kim, Y. G., Kim, S. J., et al. (2010). Crystal structure of Lon protease: molecular architecture of gated entry to a sequestered degradation chamber. *EMBO J.* 29, 3520–3530. doi: 10.1038/emboj.2010.226
- Clancy, K. W., Melvin, J. A., and McCafferty, D. G. (2010). Sortase transpeptidases: insights into mechanism, substrate specificity, and inhibition. *Biopolymers* 94, 385–396. doi: 10.1002/bip.21472
- Claus, H., Akca, E., Debaerdemaeker, T., Evrard, C., Declercq, J. P., Harris, J. R., et al. (2005). Molecular organization of selected prokaryotic S-layer proteins. *Can. J. Microbiol.* 51, 731–743. doi: 10.1139/w05-093
- Dalbey, R. E., Wang, P., and Kuhn, A. (2011). Assembly of bacterial inner membrane proteins. *Annu. Rev. Biochem.* 80, 161–187. doi: 10.1146/annurev-biochem-060409-092524
- Dalbey, R. E., Wang, P., and Van Dijl, J. M. (2012). Membrane proteases in the bacterial protein secretion and quality control pathway. *Microbiol. Mol. Biol. Rev.* 76, 311–330. doi: 10.1128/MMBR.05019-11
- De Castro, R. E., Maupin-Furlow, J. A., Giménez, M. I., Herrera Seitz, M. K., and Sanchez, J. J. (2006). Haloarchaeal proteases and proteolytic systems. *FEMS Microbiol. Rev.* 30, 17–35. doi: 10.1111/j.1574-6976.2005.00003.x
- Eichler, J. (2003). Facing extremes: archaeal surface-layer (glyco)proteins. *Microbiology* 149, 3347–3351. doi: 10.1099/mic.0.26591-0
- Eichler, J., Arbib, A., Cohen-Rosenzweig, C., Kaminski, L., Kandiba, L., and Konrad, Z. (2013). N-glycosylation in *Haloferax volcanii*: adjusting the sweetness. *Front. Microbiol.* 4:403. doi: 10.3389/fmicb.2013.00403
- Esquivel, R. N., and Pohlschroder, M. (2014). A conserved type IV pilin signal peptide H-domain is critical for the post-translational regulation of flagella-dependent motility. *Mol. Microbiol.* 93, 494–504. doi: 10.1111/mmi.12673
- Feng, L., Yan, H., Wu, Z., Yan, N., Wang, Z., Jeffrey, P. D., et al. (2007). Structure of a site-2 protease family intramembrane metalloprotease. *Science* 318, 1608–1612. doi: 10.1126/science.1150755
- Fine, A., Irihimovich, V., Dahan, I., Konrad, Z., and Eichler, J. (2006). Cloning, expression, and purification of functional Sec11a and Sec11b, type I signal peptidases of the archaeon *Haloferax volcanii*. *J. Bacteriol.* 188, 1911–1919. doi: 10.1128/JB.188.5.1911-1919.2006
- Fink-Lavi, E., and Eichler, J. (2008). Identification of residues essential for the catalytic activity of Sec11b, one of the two type I signal peptidases of *Haloferax volcanii*. *FEMS Microbiol. Lett.* 278, 257–260. doi: 10.1111/j.1574-6968.2007.01000.x
- Freeman, M. (2014). The rhomboid-like superfamily: molecular mechanisms and biological roles. *Annu. Rev. Cell Dev. Biol.* 30, 235–254. doi: 10.1146/annurev-cellbio-100913-012944
- Fukui, T., Eguchi, T., Atomi, H., and Imanaka, T. (2002). A membrane-bound archaeal Lon protease displays ATP-independent proteolytic activity towards unfolded proteins and ATP-dependent activity for folded proteins. *J. Bacteriol.* 184, 3689–3698. doi: 10.1128/JB.184.13.3689-3698.2002
- Giménez, M. I., Dilks, K., and Pohlschroder, M. (2007). *Haloferax volcanii* twin-arginine translocation substrates include secreted soluble, C-terminally anchored and lipoproteins. *Mol. Microbiol.* 66, 1597–1606. doi: 10.1111/j.1365-2958.2007.06034.x
- Haft, D. H., Payne, S. H., and Selengut, J. D. (2012). Archaeosortases and exosortases are widely distributed systems linking membrane transit with posttranslational modification. *J. Bacteriol.* 194, 36–48. doi: 10.1128/JB.06026-6011
- Henché, A. L., Van Wolferen, M., Ghosh, A., and Albers, S. V. (2014). Dissection of key determinants of cleavage activity in signal peptidase III (SPaseIII) PibD. *Extremophiles* 18, 905–913. doi: 10.1007/s00792-014-0675-4
- Hendrickx, A. P., Budzik, J. M., Oh, S. Y., and Schneewind, O. (2011). Architects at the bacterial surface – sortases and the assembly of pili with isopeptide bonds. *Nat. Rev. Microbiol.* 9, 166–176. doi: 10.1038/nrmicro2520
- Hu, J., Xue, Y., Lee, S., and Ha, Y. (2011). The crystal structure of GXGD membrane protease FlaK. *Nature* 475, 528–531. doi: 10.1038/nature10218
- Im, Y. J., Na, Y., Kang, G. B., Rho, S. H., Kim, M. K., Lee, J. H., et al. (2004). The active site of a lon protease from *Methanococcus jannaschii* distinctly differs from the canonical catalytic dyad of Lon proteases. *J. Biol. Chem.* 279, 53451–53457. doi: 10.1074/jbc.M410437200
- Jarrell, K. F., Ding, Y., Meyer, B. H., Albers, S. V., Kaminski, L., and Eichler, J. (2014). N-linked glycosylation in Archaea: a structural, functional, and genetic analysis. *Microbiol. Mol. Biol. Rev.* 78, 304–341. doi: 10.1128/MMBR.00052-13
- Jarrell, K. F., Jones, G. M., Kandiba, L., Nair, D. B., and Eichler, J. (2010). S-layer glycoproteins and flagellins: reporters of archaeal posttranslational modifications. *Archaea* 2010:13. doi: 10.1155/2010/612948
- Kaminski, L., Guan, Z., Yurist-Doutsch, S., and Eichler, J. (2013). Two distinct N-glycosylation pathways process the *Haloferax volcanii* S-layer glycoprotein upon changes in environmental salinity. *MBio* 4, e00716–00713. doi: 10.1128/mBio.00716-13
- Kandiba, L., Guan, Z., and Eichler, J. (2013). Lipid modification gives rise to two distinct *Haloferax volcanii* S-layer glycoprotein populations. *Biochim. Biophys. Acta* 1828, 938–943. doi: 10.1016/j.bbamem.2012.11.023
- Kateete, D. P., Katabazi, F. A., Okeng, A., Okee, M., Musinguzi, C., Asiimwe, B. B., et al. (2012). Rhomboids of *Mycobacteria*: characterization using an aarA mutant

- of *Providencia stuartii* and gene deletion in *Mycobacterium smegmatis*. *PLoS ONE* 7:e45741. doi: 10.1371/journal.pone.0045741
- Khanafari, A., Khavarinejad, D., and Mashinchian, A. (2010). Solar salt lake as natural environmental source for extraction halophilic pigments. *Iran. J. Microbiol.* 2, 103–109.
- Konovalova, A., Sogaard-Andersen, L., and Kroos, L. (2014). Regulated proteolysis in bacterial development. *FEMS Microbiol. Rev.* 38, 493–522. doi: 10.1111/1574-6976.12050
- Konrad, Z., and Eichler, J. (2002). Lipid modification of proteins in archaea: attachment of a mevalonic acid-based lipid moiety to the surface-layer glycoprotein of *Haloferax volcanii* follows protein translocation. *Biochem. J.* 366, 959–964. doi: 10.1042/BJ20020757
- Langklotz, S., Baumann, U., and Narberhaus, F. (2012). Structure and function of the bacterial AAA protease FtsH. *Biochim. Biophys. Acta* 1823, 40–48. doi: 10.1016/j.bbamcr.2011.08.015
- Lassak, K., Neiner, T., Ghosh, A., Klingl, A., Wirth, R., and Albers, S. V. (2012). Molecular analysis of the crenarchaeal flagellum. *Mol. Microbiol.* 83, 110–124. doi: 10.1111/j.1365-2958.2011.07916.x
- Lemberg, M. K. (2013). Sampling the membrane: function of rhomboid-family proteins. *Trends Cell Biol.* 23, 210–217. doi: 10.1016/j.tcb.2013.01.002
- Li, X., Dang, S., Yan, C., Gong, X., Wang, J., and Shi, Y. (2013). Structure of a presenilin family intramembrane aspartate protease. *Nature* 493, 56–61. doi: 10.1038/nature11801
- Matsumi, R., Atomi, H., and Imanaka, T. (2005). Biochemical properties of a putative signal peptide peptidase from the hyperthermophilic archaeon *Thermococcus kodakaraensis* KOD1. *J. Bacteriol.* 187, 7072–7080. doi: 10.1128/JB.187.20.7072-7080.2005
- Matsumi, R., Atomi, H., and Imanaka, T. (2006). Identification of the amino acid residues essential for proteolytic activity in an archaeal signal peptide peptidase. *J. Biol. Chem.* 281, 10533–10539. doi: 10.1074/jbc.M513754200
- Maupin-Furlow, J. A., Gil, M. A., Humbard, M. A., Kirkland, P. A., Li, W., Reuter, C. J., et al. (2005). Archaeal proteasomes and other regulatory proteases. *Curr. Opin. Microbiol.* 8, 720–728. doi: 10.1016/j.mib.2005.10.005
- Mesak, L. R., Mesak, F. M., and Dahl, M. K. (2004). Expression of a novel gene, gluP, is essential for normal *Bacillus subtilis* cell division and contributes to glucose export. *BMC Microbiol.* 4:13. doi: 10.1186/1471-2180-4-13
- Ng, S. Y., Chaban, B., Vandyke, D. J., and Jarrell, K. F. (2007). Archaeal signal peptidases. *Microbiology* 153, 305–314. doi: 10.1099/mic.0.2006/003087-0
- Ng, S. Y., and Jarrell, K. F. (2003). Cloning and characterization of archaeal type I signal peptidase from *Methanococcus voltae*. *J. Bacteriol.* 185, 5936–5942. doi: 10.1128/JB.185.20.5936-5942.2003
- Ng, S. Y., Vandyke, D. J., Chaban, B., Wu, J., Nosaka, Y., Aizawa, S., et al. (2009). Different minimal signal peptide lengths recognized by the archaeal preproline-like peptidases FlaK and PibD. *J. Bacteriol.* 191, 6732–6740. doi: 10.1128/JB.00673
- Parente, J., Casabuono, A., Ferrari, M. C., Paggi, R. A., De Castro, R. E., Couto, A. S., et al. (2014). A rhomboid protease gene deletion affects a novel oligosaccharide N-linked to the S-layer glycoprotein of *Haloferax volcanii*. *J. Biol. Chem.* 289, 11304–11317. doi: 10.1074/jbc.M113.546531
- Proft, T., and Baker, E. N. (2009). Pili in Gram-negative and Gram-positive bacteria – structure, assembly and their role in disease. *Cell. Mol. Life Sci.* 66, 613–635. doi: 10.1007/s00018-008-8477-4
- Robertson, C. E., Harris, J. K., Spear, J. R., and Pace, N. R. (2005). Phylogenetic diversity and ecology of environmental Archaea. *Curr. Opin. Microbiol.* 8, 638–642. doi: 10.1016/j.mib.2005.10.003
- Sastre, D. E., Paggi, R. A., and De Castro, R. E. (2011). The Lon protease from the haloalkaliphilic archaeon *Natrialba magadii* is transcriptionally linked to a cluster of putative membrane proteases and displays DNA-binding activity. *Microbiol. Res.* 166, 304–313. doi: 10.1016/j.micres.2010.07.002
- Schneewind, O., and Missiakas, D. M. (2012). Protein secretion and surface display in Gram-positive bacteria. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 367, 1123–1139. doi: 10.1098/rstb.2011.0210
- Spirig, T., Weiner, E. M., and Clubb, R. T. (2011). Sortase enzymes in Gram-positive bacteria. *Mol. Microbiol.* 82, 1044–1059. doi: 10.1111/j.1365-2958.2011.07887.x
- Stevenson, L. G., Strisovsky, K., Clemmer, K. M., Bhatt, S., Freeman, M., and Rather, P. N. (2007). Rhomboid protease AarA mediates quorum-sensing in *Providencia stuartii* by activating TatA of the twin-arginine translocase. *Proc. Natl. Acad. Sci. U.S.A.* 104, 1003–1008. doi: 10.1073/pnas.0608140104
- Szabo, Z., Albers, S. V., and Driessens, A. J. (2006). Active-site residues in the type IV prepilin peptidase homologue PibD from the archaeon *Sulfolobus solfataricus*. *J. Bacteriol.* 188, 1437–1443. doi: 10.1128/JB.188.4.1437-1443.2006
- Szabo, Z., Stahl, A. O., Albers, S. V., Kissinger, J. C., Driessens, A. J., and Pohlschroder, M. (2007). Identification of diverse archaeal proteins with class III signal peptides cleaved by distinct archaeal prepilin peptidases. *J. Bacteriol.* 189, 772–778. doi: 10.1128/JB.01547-06
- Tripepi, M., Esquivel, R. N., Wirth, R., and Pohlschroder, M. (2013). *Haloferax volcanii* cells lacking the flagellin FlgA2 are hypermotile. *Microbiology* 159, 2249–2258. doi: 10.1099/mic.0.069617-0
- Tripepi, M., Imam, S., and Pohlschroder, M. (2010). *Haloferax volcanii* flagella are required for motility but are not involved in PibD-dependent surface adhesion. *J. Bacteriol.* 192, 3093–3102. doi: 10.1128/JB.00133-10
- Wallin, E., and von Heijne, G. (1998). Genome-wide analysis of integral membrane proteins from eubacterial, archaean, and eukaryotic organisms. *Protein Sci.* 7, 1029–1038. doi: 10.1002/pro.5560070420
- Ward, D. E., Shockley, K. R., Chang, L. S., Levy, R. D., Michel, J. K., Connors, S. B., et al. (2002). Proteolysis in hyperthermophilic microorganisms. *Archaea* 1, 63–74. doi: 10.1155/2002/503191
- Wolfe, M. S. (2009). Intramembrane-cleaving proteases. *J. Biol. Chem.* 284, 13969–13973. doi: 10.1074/jbc.R800039200

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

Received: 27 November 2014; accepted: 12 January 2015; published online: 06 February 2015.

*Citation: Giménez MI, CerlettiM and De CastroRE (2015) Archaeal membrane-associated proteases: insights on *Haloferax volcanii* and other haloarchaea. Front. Microbiol. 6:39. doi: 10.3389/fmicb.2015.00039*

This article was submitted to Microbial Physiology and Metabolism, a section of the journal Frontiers in Microbiology.

Copyright © 2015 Giménez, Cerletti and De Castro. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Biosynthesis of archaeal membrane ether lipids

Samta Jain^{1,2†}, Antonella Caforio^{1,2} and Arnold J. M. Driessen^{1,2*}

¹ Department of Molecular Microbiology, Groningen Biomolecular Sciences and Biotechnology Institute, University of Groningen, Groningen, Netherlands

² The Zernike Institute for Advanced Materials, University of Groningen, Groningen, Netherlands

Edited by:

Sonja-Verena Albers, University of Freiburg, Germany

Reviewed by:

Dong-Woo Lee, Kyungpook National University, South Korea

Jerry Eichler, Ben Gurion University of the Negev, Israel

***Correspondence:**

Arnold J. M. Driessen, Department of Molecular Microbiology, Groningen Biomolecular Sciences and Biotechnology Institute – The Zernike Institute for Advanced Materials, University of Groningen, Nijenborgh 7, 9747 AG Groningen, Netherlands
e-mail: a.j.m.driessen@rug.nl

†Present address:

Samta Jain, Department of Medicine, Section of Infectious Diseases, Boston University School of Medicine, 02118 Boston, MA, USA

A vital function of the cell membrane in all living organism is to maintain the membrane permeability barrier and fluidity. The composition of the phospholipid bilayer is distinct in archaea when compared to bacteria and eukarya. In archaea, isoprenoid hydrocarbon side chains are linked via an ether bond to the *sn*-glycerol-1-phosphate backbone. In bacteria and eukarya on the other hand, fatty acid side chains are linked via an ester bond to the *sn*-glycerol-3-phosphate backbone. The polar head groups are globally shared in the three domains of life. The unique membrane lipids of archaea have been implicated not only in the survival and adaptation of the organisms to extreme environments but also to form the basis of the membrane composition of the last universal common ancestor (LUCA). In nature, a diverse range of archaeal lipids is found, the most common are the diether (or archaeol) and the tetraether (or caldarchaeol) lipids that form a monolayer. Variations in chain length, cyclization and other modifications lead to diversification of these lipids. The biosynthesis of these lipids is not yet well understood however progress in the last decade has led to a comprehensive understanding of the biosynthesis of archaeol. This review describes the current knowledge of the biosynthetic pathway of archaeal ether lipids; insights on the stability and robustness of archaeal lipid membranes; and evolutionary aspects of the lipid divide and the LUCA. It examines recent advances made in the field of pathway reconstruction in bacteria.

Keywords: archaea, ether lipids, isoprenoids, biosynthesis, lipid divide

INTRODUCTION

The “Woeseian Revolution” in 1977 defined the three domains of life as the Eukarya, the Bacteria and the Archaea (Woese and Fox, 1977). The archaeal membrane lipid composition is one of the most remarkable feature distinguishing Archaea from Bacteria and Eukarya where the hydrocarbon chain consists of isoprenoid moieties which are ether linked to the enantiomeric glycerol backbone, glycerol-1-phosphate (G1P) in comparison to glycerol-3-phosphate (G3P) of bacteria and eukarya that is ester linked to the fatty acid derived hydrocarbon chain. Polar head groups on the other hand are common in all three domains of life. Other than this core archaeal diether lipid structure, a bipolar tetraether lipid structure is also prevalent in many archaea that span the entire archaeal membrane forming a monolayer (Koga and Morii, 2007). It should be stressed that ether-linked lipids are not unique to archaea *per se*, but are also found in Bacteria and Eukarya, although not ubiquitously distributed and usually only a minor component of the lipid membrane.

The stereo specificity of archaeal lipids and their unique structure was hypothesized to be chemically more stable thereby rendering the organism with the ability to resist and thrive in extreme environmental conditions (Koga, 2012). However, archaea are also found in mesophilic and neutrophilic environment where such a structural role of ether lipids is still not postulated. At the same time, the distinguishing lipid structures have formed the basis to the evolutionary studies describing archaeal and bacterial differentiation. Several models hypothesizing the early evolution of archaeal and bacterial phospholipid biosynthesis were proposed

to answer intriguing questions about the nature of the ancestral membrane lipid composition (Lombard et al., 2012b). Understanding the archaeal lipid biosynthetic pathway is crucial to the above studies.

Decades of studies on the biosynthesis of archaeal lipids have advanced our knowledge on the major enzymatic processes but the pathway is, however still not completely understood. Several enzymes of the pathway have been studied and characterized biochemically but there are also gaps in our understanding of the archaeal lipid biosynthetic pathway and little is known about its regulation. With more genome sequences becoming available, advanced phylogenetic studies have been performed recently (Boucher et al., 2004; Daiyasu et al., 2005; Lombard et al., 2012a; Villanueva et al., 2014) and this helped to more precisely define its evolution. This review will focus on existing knowledge and recent studies on the enzymes of the pathway, the physicochemical properties of archaeal lipids, and the theories on the lipid divide.

BIOSYNTHESIS OF ARCHAEOAL MEMBRANE LIPIDS

ISOPRENOID BUILDING BLOCKS AND CHAIN ELONGATION

Isoprenoids are ubiquitous to all three domains of life. They are structurally diverse, forming more than 30,000 different compounds in nature ranging from steroids, quinones, carotenoids, and membrane lipids. The building blocks of isoprenoids are universal carbon five subunits called isopentenyl pyrophosphate (IPP) and dimethylallyl pyrophosphate (DMAPP) that are isomers. The biosynthetic pathway leading to the synthesis of IPP

and DMAPP vary in different organisms (reviewed in Lombard and Moreira, 2011; Matsumi et al., 2011). To date, three pathways have been reported – 2-C-methyl-D-erythritol 4-phosphate/1-deoxy-D-xylulose 5-phosphate pathway (MEP/DOXP pathway) and two mevalonate (MVA) pathways. The MEP pathway genes share no homology to genes of the MVA pathway where pyruvate and glyceraldehyde-3-phosphate molecules are condensed together to form 1-deoxy-D-xylulose-5-phosphate (DXP) which is subsequently converted to IPP and DMAPP by five enzymes (Xue and Ahring, 2011; **Figure 1**). The MEP pathway is most common in bacteria although some Firmicutes possess the MVA pathway. The MVA pathway consists of seven enzymatic reactions where two acetyl-CoA molecules are condensed to form acetoacetyl-CoA, which is further condensed to form 3-hydroxy-3-methylglutaryl CoA (HMG-CoA). HMG-CoA undergoes phosphorylation and decarboxylation to form IPP via the formation of MVA (**Figure 1**). The classical MVA pathway is common to eukaryotes while some plants and photosynthetic eukaryotes possessing the MEP pathway in addition (Lombard and Moreira, 2011).

Interestingly, homologs of the last three enzymes of the classical MVA pathway, i.e., phosphomevalonate kinase, diphosphomevalonate decarboxylase and isopentenyl diphosphate isomerase could not be found in the majority of archaea (except in Sulfolobales that have classical MVA pathway; Boucher et al., 2004). This search led to the discovery of the alternate MVA pathway that differs from the classical one in the last three steps (**Figure 1**). The enzyme isopentenyl kinase (IPK) was first discovered in *Methanocaldococcus jannaschii* and found to be conserved in archaea (Grochowski et al., 2006). Its structure was determined (Dellas and Noel, 2010) and IPK enzymes from *Methanothermobacter thermautotrophicus* and *Thermoplasma acidophilum* were characterized biochemically (Chen and Poulter, 2010). In the alternate MVA pathway, phosphomevalonate is decarboxylated to isopentenyl phosphate by a decarboxylase (enzyme yet to be identified in archaea), which is subsequently phosphorylated to IPP by IPK. Furthermore, instead of the typical IDI1 isomerase that performs the last step of the classical MVA pathway, archaea have IDI2 which is not homologous to IDI1 but that performs the same reaction (Matsumi et al., 2011). Interestingly, a decarboxylase enzyme that converts phosphomevalonate to isopentenyl phosphate was found in green non-sulfur bacteria *Roseiflexus castenholzii* along with the presence of IPK enzymes indicating the existence of alternate MVA pathway in organisms other than archaea (Dellas et al., 2013). In general, IPP and DMAPP are synthesized by the MEP pathway in most of the bacteria and by two MVA pathways in eukarya and archaea. The classical MVA pathway of eukaryotes and the alternate MVA pathway of archaea share four of their seven steps.

The isoprenoid building blocks IPP and DMAPP undergo sequential condensation reactions where DMAPP acts as the first allylic acceptor of IPP leading to the formation of a carbon 10 (C10) compound termed geranyl diphosphate (GPP). Further condensation reactions proceed with the addition of IPP molecules where the chain length increases each time by a C5 unit forming farnesyl (C15), geranylgeranyl (C20), farnesylgeranyl (C25) diphosphate etc. This reaction of chain elongation is catalyzed by enzymes belonging to the family of prenyl transferases that are common to all three domains of life (Wang and Ohnuma,

1999; Vandermoten et al., 2009). Depending on the length and geometry of the final molecule, prenyl transferases can have several members in its family. The geometry of the molecule could be *cis* or *trans* and the chain length of the *trans* form generally ranges from C10 (e.g., monoterpenes) to C50 (e.g., Coenzyme Q10) and even longer for the *cis* forms. The chain length found in archaeal membrane lipids is always in the *trans* form and composed mostly of C20 [geranylgeranyl diphosphate (GGPP)] or C25 (farnesylgeranyl diphosphate). Tetraethers are composed of a C40 chain length, the synthesis of which is still unknown (discussed below). The archaeal prenyltransferase enzymes GGPP synthase and farnesylgeranyl diphosphate synthase synthesize specifically C20 or C25 product chain lengths, respectively (Ohnuma et al., 1994; Tachibana et al., 2000; Hemmi et al., 2002; Lai et al., 2009). They belong to short chain *trans* prenyl transferases family (that catalyze reactions ranging from C10–C25). Interestingly, a bifunctional prenyltransferase that catalyzes the synthesis of both C15 and C20 isoprenoids has been characterized from *Thermococcus kodakaraensis* (Fujiwara et al., 2004) and *Methanobacterium thermoadautotrophicum* and is considered to be an ancient enzyme (Chen and Poulter, 1993). Multiple sequence alignment of homologues of the family display high sequence similarity with seven conserved regions where region two and seven contain the highly conserved aspartate rich sequences called first aspartate-rich motif (FARM) and second aspartate-rich motif (SARM) domains, respectively. Numerous mutagenesis and structural studies including several members of the family show that the region within the aspartate rich domains are involved in the binding and catalysis of the substrate while the regions flanking these domains are the major determinants of the chain length as they contribute to the size of the active site hydrophobic pocket. For example, GGPP synthase from *Sulfolobus acidocaldarius* could be mutated (at Phe-77 which is fifth amino acid upstream of FARM) to catalyze longer chain length (C30–C50) products (Ohnuma et al., 1996) and farnesyl pyrophosphate (FPP) synthase of *Escherichia coli* could be mutated (at Tyr-79, also 5th amino acid upstream of FARM) to create GGPP synthase (Lee et al., 2005). Physical factors have also been shown to influence the chain length of the product, e.g., the bifunctional enzyme farnesyl diphosphate/GGPP synthase of *Thermococcus kodakaraensis* shows an increase in the FPP/GGPP ratio with the reaction temperature (Fujiwara et al., 2004).

GLYCEROL-1-PHOSPHATE BACKBONE

The glycerophosphate backbone of archaea has an opposite stereoconfiguration than those of bacteria and eukarya. The archaeal enzyme responsible is G1P dehydrogenase that shares homology with alcohol and glycerol dehydrogenases but no homology to the bacterial/eukaryal G3P dehydrogenase. They belong to two separate families. However, both catalyze the reduction of dihydroxyacetone phosphate (DHAP) using nicotinamide adenine dinucleotide hydrogen (NADH) or nicotinamide adenine dinucleotide phosphate hydrogen (NADPH) as substrate (**Figure 2**; **Table 1**). G1P dehydrogenase uses Zn²⁺ for metal ion interaction in its active site (Han and Ishikawa, 2005) and transfers the pro-R hydrogen of NADH in contrast to G3P dehydrogenase that transfers the pro-S hydrogen (Koga et al., 2003); both enzymes bind the nicotinamide ring in an opposite orientation.

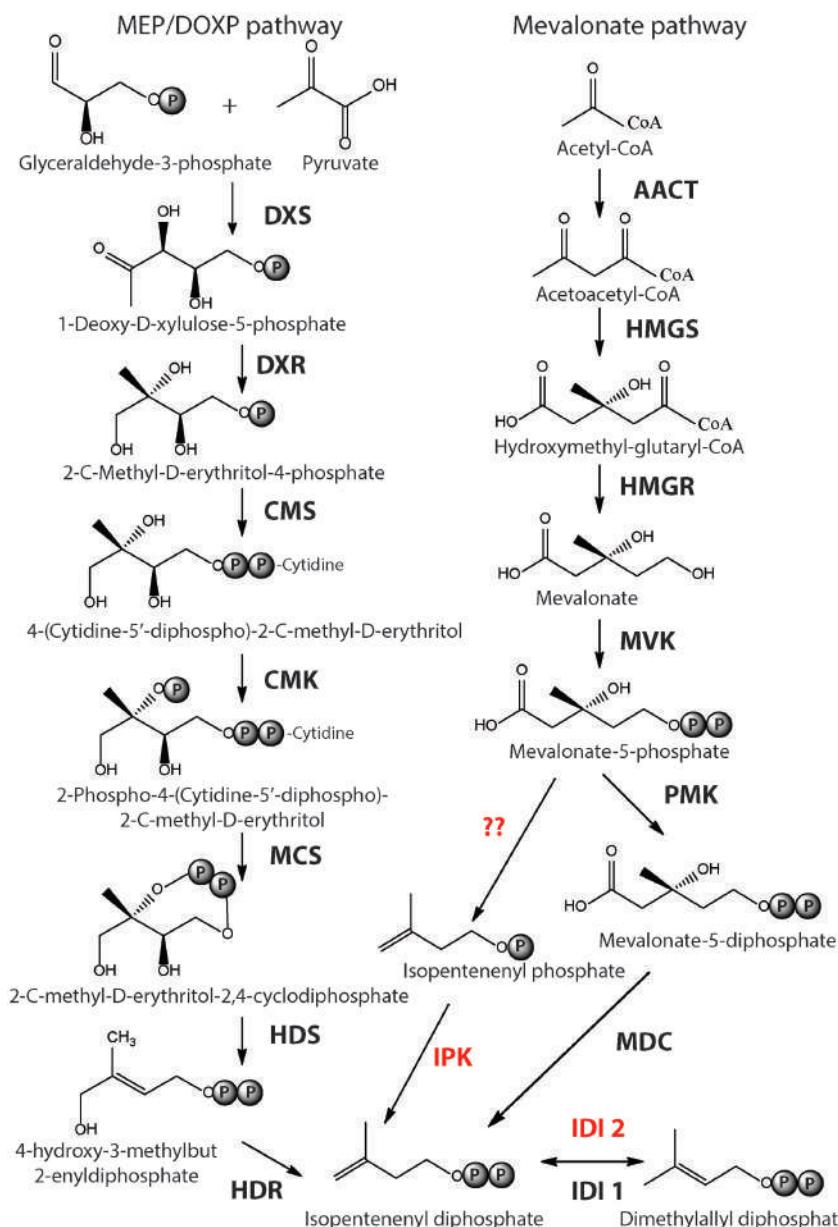


FIGURE 1 | Biosynthesis of Isoprenoid building blocks. The three pathways leading to the synthesis of IPP and DMAPP – MEP/DOXP pathway, Mevalonate and the alternate mevalonate pathway (in red), which is prevalent in archaea are shown. Mevalonate and the alternate mevalonate pathways share four of their seven enzymatic steps. The enzyme acronyms are written in bold. DXS, 1-Deoxy-d-xylulose 5-phosphate synthase; DXR, 1-Deoxy-d-xylulose 5-phosphate reductase; CMS, 2-C-methyl-d-erythritol 4-phosphate cytidyltransferase; CMK, 4-diphosphocytidyl-2-C-methyl-

d-erythritol kinase; MCS, 2-C-methyl-d-erythritol 2,4-cyclodiphosphate synthase; HDS, (E)-4-Hydroxy-3-methyl-but-2-enyl pyrophosphate synthase; HDR, (E)-4-Hydroxy-3-methyl-but-2-enyl pyrophosphate reductase; IDI, Isopentenyl diphosphate isomerase; AACT, acetoacetyl-CoA thiolase; HMGS, 3-hydroxy-3-methylglutaryl-CoA synthase; HMGR, 3-hydroxy-3-methylglutaryl-CoA reductase; MVK, mevalonate kinase; PMK, phosphomevalonate kinase; MDC, mevalonate-5-decarboxylase; IPK, isopentenyl phosphate kinase.

G1P dehydrogenase is conserved in archaea. The enzyme has been purified and characterized from *Methanothermobacter thermoautotrophicus* as an octamer (Nishihara and Koga, 1997), *Aeropyrum pernix* as a homodimer (Han et al., 2002) and from *Sulfolobus tokodaii* (Koga et al., 2006). Its activity has been accessed in cell free homogenates of several archaea (Koga and Morii, 2007; Lai et al., 2009).

Until recently it was thought that the stereo specificity is the hallmark of the ‘lipid divide’ where the G1P backbone is exclusively attributed to archaea. This was challenged by the discovery and characterization of the bacterial G1P dehydrogenase homolog of *Bacillus subtilis* which is annotated as ‘AraM’ (Guldan et al., 2008). It is also found in other related Gram positive and negative bacteria. Similar to the G1P dehydrogenase of *Aeropyrum*

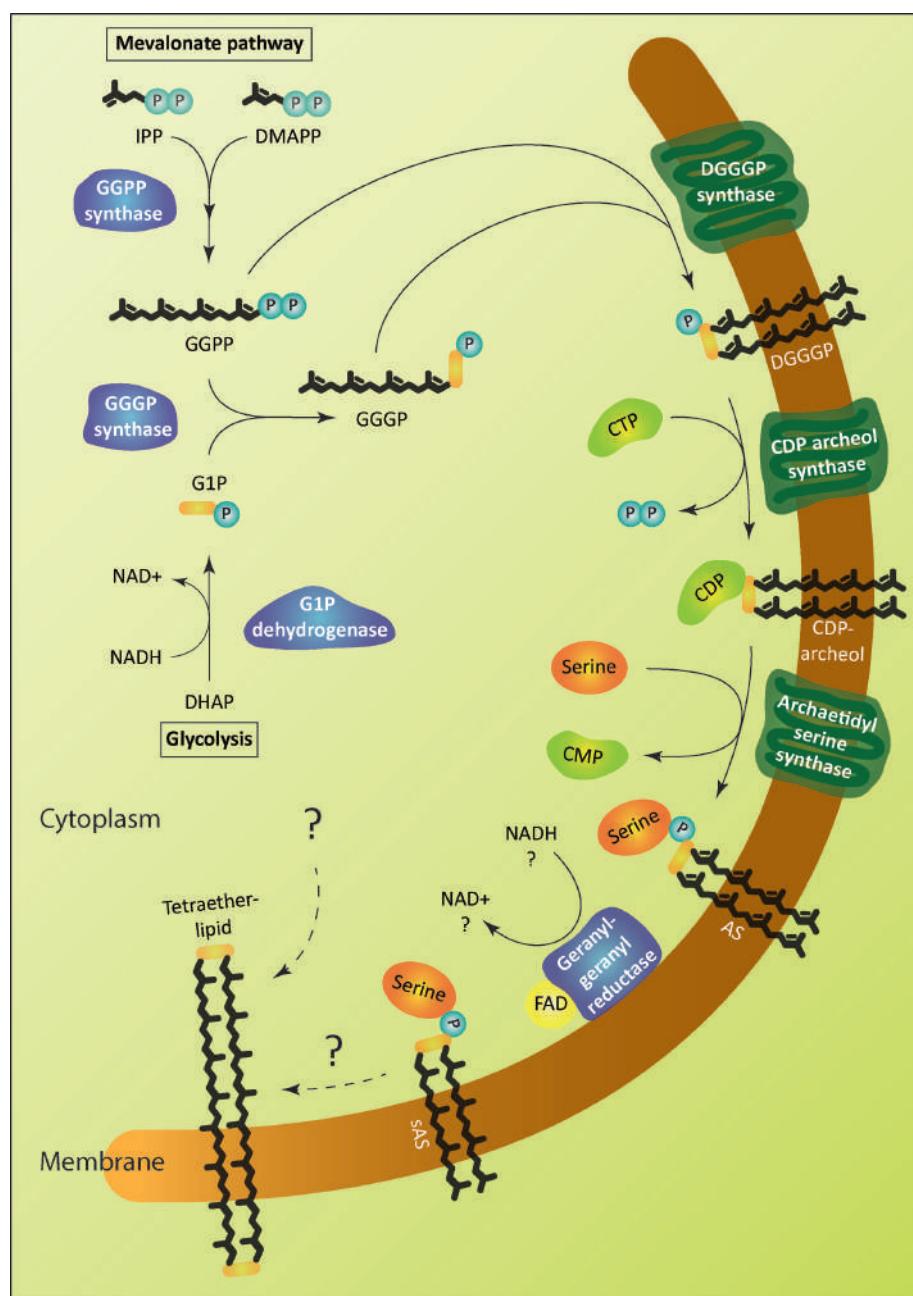


FIGURE 2 | Enzymatic pathway of archaeal lipid biosynthesis. The soluble enzymes of the pathway are colored in blue and the membrane proteins in green. The biosynthetic steps leading to the formation of tetraether lipids is unknown. Archaetidylserine (AS) and saturated AS (sAS) are depicted as an example of polar head

group modification. Isopentenyl pyrophosphate (IPP) and dimethylallyl pyrophosphate (DMAPP) are isomers. GGPP, geranylgeranyl diphosphate; G1P, *sn*-glycerol-1-phosphate; GGGP, 3-O-geranylgeranyl-*sn*-glycerol-1-phosphate; DGGGP, 2,3-bis-O- geranylgeranyl-*sn*-glycerol-1-phosphate.

pernix, AraM forms a homodimer and performs G1P dehydrogenase activity. However, the two enzymes have different catalytic efficiencies and AraM is Ni^{2+} ion dependent (Han et al., 2002; Guldan et al., 2008). Remarkably, the G1P molecule eventually becomes part of an archaea type ether lipid heptaprenylglycerol phosphate in *B. subtilis*, the function of which is still unknown (Guldan et al., 2011).

ETHER LINKAGES

The first and second ether bonds between G1P and GGPP is catalyzed by the enzyme geranylgeranylglycerol diphosphate (GGGP) synthase and di-0- geranylgeranylglycerol diphosphate (DGGGP) synthase respectively. GGGP synthase is a conserved enzyme found in all archaea except Nanoarchaeota, which is a symbiont and possesses no genes of the lipid biosynthesis pathway

Table 1 | Summary of kinetic parameters derived for enzymes of archaeal lipid biosynthetic pathway.

Name	Organism	Reaction	Kinetic parameters	Reference
GGPP synthase	<i>Methanobacterium thermoformicicum</i>	IPP + FPP \rightarrow GGPP (60°C)	K_m IPP = 51.4 μM K_m FPP = 28 μM V_{max} = 437.9 nmol/min/mg	Tachibana et al. (1993)
FPP/GGPP synthase	<i>Thermococcus kodakaraensis</i>	IPP + DMAPP \rightarrow GPP	K_m IPP = 23 μM , DMAPP = 9.5 μM	Fujiwara et al. (2004)
		IPP + GPP \rightarrow FPP	K_m IPP = 22 μM , GPP = 2.2 μM	
		IPP + FPP \rightarrow GGPP (70°C)	K_m IPP = 3 μM , FPP = 1.7 μM	
		IPP + DMAPP \rightarrow GPP	K_m IPP = 79 μM , DMAPP = 13 μM	
		IPP + GPP \rightarrow FPP	K_m IPP = 31 μM , GPP = 3 μM	
G1P dehydrogenase	<i>Aeropyrum pernix</i>	IPP + FPP \rightarrow GGPP (90°C)	K_m IPP = 16 μM , FPP = 0.81 μM	
		DHAP+NADH \rightarrow G1P (65°C)	K_m DHAP = 0.46 mM, k_{cat} (min^{-1}) = 154.25 K_m NADH = 0.032 mM, k_{cat} (min^{-1}) = 143.96	Han et al. (2002)
GGGP synthase	<i>Methanobacterium thermoautotrophicum</i>	G1P + GGPP \rightarrow GGGP (55°C)	K_m G1P = 13.5 μM K_m GGPP = 0.51 μM k_{cat} (s^{-1}) = 0.34	Soderberg et al. (2001)
	<i>Thermoplasma acidophilum</i>	G1P+GGPP \rightarrow GGGP (55°C)	K_m G1P = 21.2 μM K_m GGPP = 75 nM k_{cat} (s^{-1}) = 6.1, V_{max} = 13.5 $\mu\text{mol}/\text{min}/\text{mg}$	Nemoto (2003)
CDP-archaeol synthase	<i>Archaeoglobus fulgidus</i>	DGGGP + CTP \rightarrow CDP archaeol (65°C)	K_m DGGGP = 0.12 mM k_{cat} (s^{-1}) = 0.55	Jain et al. (2014)

(Podar et al., 2013). It is also found in some bacteria where the polypropenyl diphosphate substrate chain length could vary, e.g., PcrB of *Bacillus subtilis* which is a heptaprenyl diphosphate synthase (Ren et al., 2013). GGGP synthase is a crucial enzyme in the biosynthetic pathway of phospholipid metabolism in archaea as it brings together the three important characteristic features of the archaeal lipid structure – stereoisomeric G1P glycerol backbone and isoprenoid GGPP side chain linking them together via an ether bond (Figure 2; Table 1). Phylogenetic analysis of the GGGP synthase enzymes distinguishes it into two families, group I and group II, both comprising of archaeal and bacterial sequences. Several enzymes from both the groups have been characterized and a recent study performed the biochemical analysis of 17 members of GGGP synthase family (Peterhoff et al., 2014). The enzymes of group I form dimers (except the monomeric GGGP synthase of *Halobacterium salinarum*) and the group II enzymes are dimeric or hexameric in nature. Both the groups are further subdivided into Ia, Ib, IIa, and IIb with a and b corresponding archaea and bacteria, respectively. Crystal structures of enzymes from all the four groups have been solved. The first crystal structure from the group I GGGP synthase of *Archaeoglobus fulgidus* displays a modified triose phosphate isomerase (TIM)-barrel structure (Payandeh et al., 2006). It forms a dimer bound to the G1P substrate with a central eight-stranded parallel β -barrel and a hydrophobic core surrounded by α -helices (Figure 3A). Helix-3 is replaced by a ‘strand’ which is a novel TIM-barrel modification not observed

previously. The substrate GGPP binds to the deep cleft traversing the top of the β -barrel. There is a ‘plug’ at the bottom of the barrel and the active site lies at the C-terminal end. The G1P molecule sits near the top inner rim of the barrel and the phosphate group binds to the standard phosphate-binding motif of the TIM-barrel. G1P forms 14 hydrogen bonds within the active site. The ($\beta\alpha$)8-barrel fold is found in all the other structures of the GGGP synthases as well with the active site at the C-terminus. The crystal structure of group II archaeal hexameric GGGP synthase of *Methanothermobacter thermoautrophicus* displays a combination of three dimers that resemble the group I dimer (Figures 3B,C). In group II, however, the plug of the barrel is longer than in group I and there are ‘limiter residues’ that restrict the length of hydrophobic pocket to accommodate the polypropenyl diphosphates of a specific length. Interestingly, an aromatic anchor residue is responsible for the hexameric configuration of the enzyme, mutation of which causes it to dimerize without any loss of activity (Peterhoff et al., 2014).

The intrinsic membrane protein DGGGP synthase catalyzes the formation of the second ether bond between the substrate GGGP and GGPP to form DGGGP (Figure 2). It belongs to the family of ubiquinone-biosynthetic (UbiA) prenyltransferases, the members of which are responsible for the biosynthesis of respiratory quinones, chlorophyll, heme etc. by transferring a prenyl group to the acceptors that generally have hydrophobic ring structures. DGGGP synthase is divergent among archaea and could not be

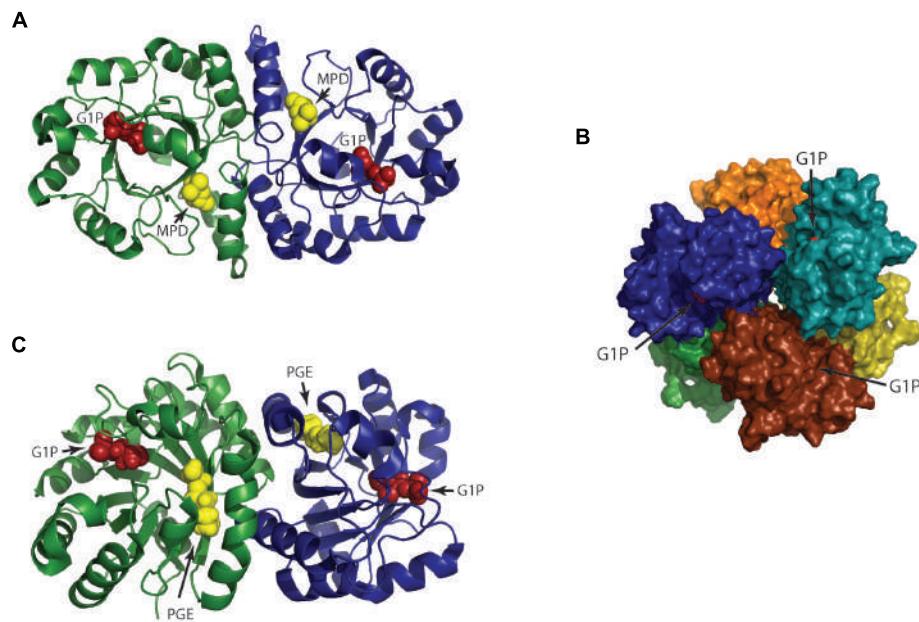


FIGURE 3 | Structure of 3-O-geranylgeranyl-sn-glyceryl-1-phosphate synthase (GGGP synthase). (A) Crystal structure of group I dimeric GGGP synthase from *Archaeoglobus fulgidus* solved at 2.0 Å resolution (PDB: 2F6X) has a novel TIM-barrel modification (Payandeh et al., 2006). Sn-glycerol-1-phosphate (G1P) and a cryoprotectant (MPD) used in the experiment are colored as red and yellow spheres respectively. The cavity for MPD likely represents the binding site for the second substrate geranylgeranyl diphosphate (GGPP). (B) Crystal structure of group II GGGP synthase from *Methanothermobacter*

thermoautotrophicus that forms a hexamer as a combination of three dimers was solved at 2.8 Å resolution (PDB: 4MM1; Peterhoff et al., 2014). G1P is displayed as red spheres. (C) A dimer from (B) is represented to compare with the dimeric group I GGGP synthase of (A). The G1P and triethylene glycol (PGE) are colored as red and yellow spheres respectively. The hexameric subunits are rotated to each other unlike the dimers in (A). The anchoring of G1P in group I and II takes place by the standard phosphate binding motif but their G1P binding pocket are different.

identified in the genomes of Thaumarchaeota (Villanueva et al., 2014).

Unlike other enzymes of the pathway, DGGGP synthase has not been well characterized probably due to technical limitations with overexpression of the membrane protein. The DGGGP synthase activity was first found in the membrane fraction of *Methanothermobacter marburgensis* (Zhang and Poulter, 1993). Later the gene was identified in the genome of *Sulfolobus solfataricus* as UbiA-2, cloned in *E. coli* and purified to study the Mg²⁺ dependent enzymatic activity using radiolabeled substrates and mass spectrometry (Hemmi et al., 2004). The ratio of the substrates utilized in the reaction was found to be 1:1.1 in a double labeling experiment using [³H]GGPP and [¹⁴C]GGGP, respectively. Specificity for GGPP and GGGP was also measured by substituting them with different prenyl substrates, of which none of them were used in the reaction by DGGGP synthase. In another study, DGGGP synthase was shown to accept both the S and R form of GGGP showing that unlike GGPP synthase, it is enantio unselective (Zhang et al., 2006). DGGGP synthase activity of *Archaeoglobus fulgidus* (Lai et al., 2009) and *Methanoscincus acetivorans* (Yokoi et al., 2012) was also observed in *E. coli* when the corresponding genes were expressed along with four previous enzymes of the pathway. However, the expression level of the enzyme was either too low to detect (Lai et al., 2009) or not investigated (Yokoi et al., 2012). In a later study, a higher expression level of DGGGP synthase of *Archaeoglobus fulgidus* was obtained in *E. coli* by changing

the ribosome-binding site and the activity of purified DGGGP synthase was monitored (Jain et al., 2014).

CDP ARCHAEOOL FORMATION

The next step in the archaeal lipid biosynthetic pathway is the activation of DGGGP by cytidine triphosphate (CTP) to form the substrate for polar head group attachment called cytidine diphosphate (CDP)-archaeol (**Figure 2; Table 1**). The reaction is brought about by the enzyme CDP-archaeol synthase (CarS), the activity of which was first studied in the membrane fraction of *Methanothermobacter thermoautotrophicus* (Morii et al., 2000). Using various synthetic substrate analog, the activity was found to be specific for unsaturated archaeodic acid with geranylgeranyl chains and did not depend on the stereo specificity or ether/ester bond of the substrate. Minute amount of CDP-archaeol were also detected in growing cells labeled with inorganic ³²P. The gene responsible for this activity was only identified in a recent study (Jain et al., 2014). The enzyme CarS is conserved among archaea (except Nanoarchaeota). However, like the enzyme DGGGP synthase, it could not be identified in the families of Thaumarchaeota.

Interestingly, an analogous reaction is found in the bacterial phospholipid biosynthetic pathway where phosphatidic acid is activated by CTP to form CDP diacylglycerol by the enzyme CDP diacylglycerol synthase (CdsA). Although the sequence similarity between CdsA and CarS is very low, hydropathy profile alignment of the two families shows similarity in their secondary structure

with overlapping transmembrane segments and cytoplasmic loop regions residing in the C-terminus half. CarS from *Archaeoglobus fulgidus* was expressed and purified from *E. coli*. Similar to CdsA, CarS activity was found to be dependent on Mg²⁺, both accepts CTP and deoxycytidine triphosphate (dCTP) as substrates and does not utilize adenosine triphosphate (ATP), guanosine triphosphate (GTP), or thymidine triphosphate (TTP) nucleotides in the reaction using substrate DGGGP. However, the two enzymes displayed distinct activity with respect to the lipid substrate specificity where CarS only accepts unsaturated archaetidic acid with geranylgeranyl chains, while CdsA takes phosphatidic acid (Jain et al., 2014).

POLAR HEAD GROUP ATTACHMENT

The polar head groups serine, ethanolamine, glycerol and *myo*-inositol are found in the phospholipids in all three domains of life. The enzymes that catalyze the replacement of the cytidine monophosphate (CMP) entity of CDP-archaeol or CDP-diacylglycerol with a polar head group are homologous and belong to CDP-alcohol phosphatidyltransferase family (Koga, 2011). Archaetidylserine (AS) synthase catalyzes the formation of AS from CDP-archaeol and L-serine (Figure 2) and is homologous to bacterial phosphatidylserine (PS) synthase. The enzyme can be classified into two subclasses. Subclass I includes enzymes distributed in Gram-negative bacteria, such as *E. coli* while subclass II enzymes are widespread among Gram-positive bacteria (*B. subtilis*), yeast and archaea. Studies using cell free extracts of *Methanothermobacter thermautotrophicus*, *B. subtilis*, and *E. coli* showed that both the AS and PS synthase from *Methanothermobacter thermautotrophicus* and *B. subtilis* have a broad substrate specificity and can accept lipid derivatives from archaea or bacteria. On the other hand, the *E. coli* PS synthase was specific for bacterial lipid derivatives only (Morii and Koga, 2003).

Archaetidylinositol phosphate (AI) synthase catalyzes the reaction where precursors L-*myo*-inositol-1-phosphate and CDP-archaeol are converted to AI phosphate as an intermediate which is further dephosphorylated to AI (Morii et al., 2009). This reaction is similar to the bacterial phosphatidylinositol phosphate (PI) synthase. Similar to AS and PS synthase, the AI and PI synthase show a broad substrate specificity accepting both, archaeal and bacterial lipid derivatives as substrates (Morii et al., 2014). Enzymes homologous to PS decarboxylase and phosphatidylglycerol (PG) synthase have been identified in archaea as AS decarboxylase and archaetidylglycerol (AG) synthase but not yet characterized biochemically (Daiyasu et al., 2005).

SATURATION OF DOUBLE BONDS

The mature phospholipids of archaea exist in their fully saturated form. The archaeal enzyme digeranylgeranylglycerophospholipid reductase catalyzes the hydrogenation or saturation of the geranylgeranyl chains of unsaturated archaetidic acid (DGGGP) in a stereospecific manner (Xu et al., 2010). It belongs to the geranylgeranyl reductase (GGR) family that includes GGR from plant and prokaryotes that are mainly involved in photosynthesis. Prenyl reductases other than GGRs are also found in all three domains of life and these enzymes catalyze the complete or partial reduction

of isoprenoid compounds like respiratory quinones, tocopherol, dolichol, and other polyprenols (Ogawa et al., 2014).

The structures of the archaeal GGR monomer from *Thermoplasma acidophilum* (Xu et al., 2010) and *Sulfolobus acidocaldarius* (Sasaki et al., 2011) show that they belong to p-hydroxybenzoate hydroxylase (PHBH) superfamily of flavoproteins (Figure 4A). The GGR from the thermophilic archaea *Thermoplasma acidophilum* was crystallized in complex with flavin adenine dinucleotide (FAD) where FAD adopts the close confirmation that possibly changes with the binding of the substrate, like in other members of the PHBH family. The reduction of FAD is brought about by either NADH or other reducing agents. Since the protein was overexpressed in *E. coli*, a surrogate lipid-like ligand assigned as phosphatidylglycerol (PGX) was found in the active site forming an imperfect fit to the substrate binding pocket. The lipid binding cavity of GGR is R shaped having two tunnels where the larger tunnel B is more permissive than the smaller tunnel A which is restricted in shape (Figure 4B). The *S. acidocaldarius* GGR is structurally similar to GGR from *Thermoplasma acidophilum* in FAD binding and the catalytic region (Figures 4C,D) but not in the C-terminal domain which is longer in *S. acidocaldarius* GGR. The conserved sequence motif (YxWxFPx7-8GxG) lies in the large cavity of the catalytic domain and is thought to keep the substrate in position for the reduction reaction as also indicated by mutational studies. Although the enzymes reduce GGGP, they also reduce the double bonds of related compounds like GGGP and GGPP (Sasaki et al., 2011). Another study where the *Methanoscincus acetivorans* GGR was expressed in *E. coli* along with four previous genes of the archaeal lipid biosynthetic pathway, the DGGGP derivative with a fully saturated isoprenoid chain could be obtained (Isobe et al., 2014). Interestingly, the saturation only took place when GGR was coexpressed with a ferredoxin gene found upstream of GGR in the genome of *Methanoscincus acetivorans*, the ferredoxin possibly functioning as a specific electron donor. However, no ferredoxin coexpression was required when the *Methanoscincus acetivorans* GGR was replaced by *S. acidocaldarius* GGR in the same study. Also, the conservation of the ferredoxin gene upstream of GGR in other archaea was not analyzed.

It is not known at what step of the biosynthetic pathway, hydrogenation takes place. However, since that CarS is specific for the unsaturated substrate, saturation probably takes place after the formation of CDP-archaeol. Although the enzyme AS synthase can accept both saturated and unsaturated substrates for catalysis, the detection of unsaturated AS in the cells of *Methanothermobacter thermautotrophicus* suggests that hydrogenation may already take place after the polar head groups are attached (Koga and Morii, 2007).

TETRAETHER FORMATION

Tetraether (caldarchaeol) lipid structure with varying number (0–8) of cyclopentane moieties are widespread among archaea and a dominating membrane lipid structure in Crenarchaeota and Thaumarchaeota. Euryarchaeota synthesize archaeol or both archaeol and caldarchaeols. On the other hand, Thaumarchaeota have characteristic tetraether lipids with four cyclopentane moieties and a cyclohexane moiety (Villanueva et al., 2014). One of the most intriguing steps in the archaeal biosynthetic pathway is

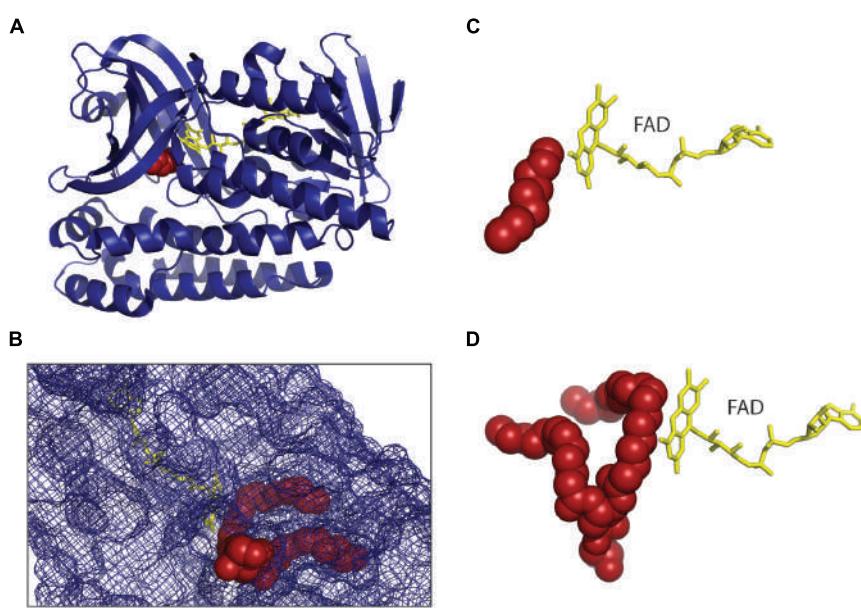


FIGURE 4 | Structure of geranylgeranyl reductase (GGR). (A) The crystal structure of *Sulfolobus acidocaldarius* GGR at 1.85 Å resolution (PDB: 3ATQ) is shown (Sasaki et al., 2011) in complex with FAD (yellow). It is similar to the structure of GGR from *Thermoplasma acidophilum* (PDB: 3OZ2; Xu et al., 2010) sharing the FAD binding and the catalytic domain. Both the enzymes were crystallized where a lipid molecule (red) originating from the host

organism was found in the substrate binding site. (B) The lipid binding cavity of GGR is R shaped having two tunnels where the larger tunnel B is more permissive and the smaller tunnel A is restricted in shape. One edge is located at the branching point of the tunnel. (C,D) Comparison of the ‘near perfect’ placement of the hydrocarbons of lipid moieties on the *re* face of the FAD from *S. acidocaldarius* GGR (C) and *Thermoplasma acidophilum* GGR (D).

the tetraether formation. *In vivo* studies suggested that tetraethers are formed from saturated diethers via head to head condensation reaction. Pulse chase and labeling experiment of *Thermoplasma acidophilum* cells with [¹⁴C]-MVA showed that the label first incorporates into the archaeol until saturation and only then into caldarchaeol. When an inhibitor of tetraether lipid synthesis (terbinafine) is used, pulse labeling leads to the accumulation of diethers and this phenomenon can be reversed by removal of the inhibitor (Nemoto et al., 2003). However, in another study, radiolabelled archaeol was not incorporated into the tetraethers of *Methanospirillum hungatei* cells and the presence of double bonds was necessary for the incorporation of labeled DGGGP into the tetraethers of *Methanobacterium thermoautotrophicus* cells (Poulter et al., 1988; Eguchi et al., 2000). The enzyme responsible for the formation of the presumed and unusual C–C bond for tetraethers has not been identified and there is no *in vitro* data to support this hypothesis (Koga and Morii, 2007). Recently, an alternative pathway for tetraether and cyclopentane ring formation was hypothesized (Villanueva et al., 2014). A multiple lock and key mechanism was proposed owing to the ‘greater functional plasticity’ of the enzymes IPP synthase, GGGP synthase, and DGGGP synthase so that they can accommodate prenyl substrates with a ring structure and chain length longer than C20. The cyclopentane rings would be formed early in the pathway before attachment of the glycerol moiety and the C20 geranyl molecules would couple together via a tail-to-tail mechanism to form the C40 phytoene chain by phytoene synthase, an enzyme that is wide spread in archaea. However, both possibilities still need to be experimentally demonstrated.

PHYSICOCHEMICAL PROPERTIES OF ARCHAEOAL LIPIDS

ARCHAEAL MEMBRANE LIPID COMPOSITION – RESPONSE TO STRESS

Within the archaeal lipids there is a great diversity varying in length, composition and configuration of the side chains (Figure 5). The most common archaeal core lipid is *sn*-2,3-diphytanoylglycerol diether, generally called archaeol, which can undergo several modifications including hydroxylation and condensation. Pioneering studies on archaeal membrane lipid composition and biosynthesis have been performed in early eighties and nineties especially in halophiles using *in vivo* labeling experiments by Kates and colleagues (Kates, 1992, 1993; Kamekura and Kates, 1999) and reviewed in detail (Koga and Morii, 2007). Among others, the studies showed that halophiles are mainly characterized by the phospholipid known as phosphatidylglycerol methyl phosphate (PGP-Me) along with sulfated and desulfated archaeols (Kates, 1993). Archaeol bearing elongated hydrocarbon chains (C20–C25) are found in some methanogens and halobacteriales (Figure 5A; Koga et al., 1993; Kamekura and Kates, 1999). A seemingly head-to-head condensation of two diether lipid molecules is one of the most frequent and functionally important structural variations that leads to a glycerol-dialkylglycerol tetraether lipid, known as caldarchaeol. It should be stressed, however, that the enzymatic mechanism resulting in this lipid species is entirely unresolved. This core lipid is the most widespread in Archaea, characterized by different modifications depending on the archaeal species. It is in particular abundant among the phyla Euryarchaeota, Crenarchaeota, and Thaumarchaeota. Up to 8 cyclopentane moieties can be found in lipids of the Thermoplasmatales (Figure 5B) and in the Euryarchaeota phylum,

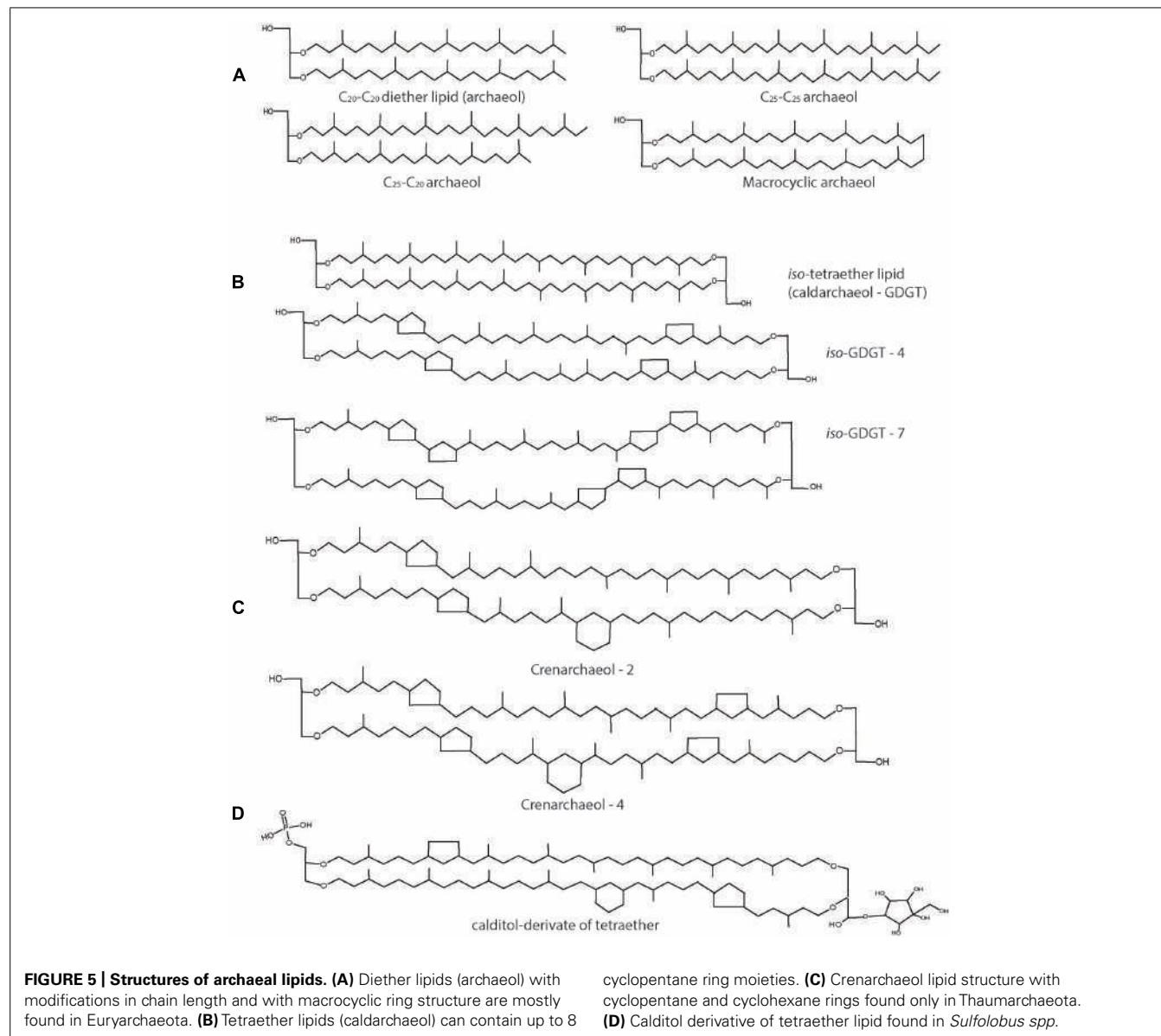


FIGURE 5 | Structures of archaeal lipids. (A) Diether lipids (archaeol) with modifications in chain length and with macrocyclic ring structure are mostly found in Euryarchaeota. (B) Tetraether lipids (caldarchaeol) can contain up to 8

cyclopentane ring moieties. (C) Crenarchaeol lipid structure with cyclopentane and cyclohexane rings found only in Thaumarchaeota. (D) Calditol derivative of tetraether lipid found in *Sulfolobus* spp.

in general (Macalady et al., 2004; UDA et al., 2004; Schouten et al., 2013). Interestingly, the presence of cyclopentane and cyclohexane ring is a distinct feature of Thaumarchaeota leading to a structure known as creanarchaeol (Figure 5C; Pitcher et al., 2010; Damsté et al., 2012). In some thermoacidophiles and methanogens, a polar head group called nonitol is found which is composed of nine-carbon chain. Recent studies have revealed that the 9-carbon nonitol structure is often found as a polyhydroxylated cyclopentanic form called calditol (Koga and Morii, 2005). Therefore, these structure are now known as glycerol-dialkyl-calditol-tetraether (Figure 5D) and are the major component of the membrane of *Sulfolobus* species (Sugai et al., 1995; Untersteller et al., 1999; Gambacorta et al., 2002). The presence of tetraether lipids and the ratio of diether/tetraether lipids vary depending on the archaeal species and also upon growth conditions (Gambacorta et al., 1995). Likewise, there

is also a wide diversity of polar lipids in archaea including phospholipids, glycolipids, phosphoglycolipids, sulpholipids, and aminolipids (Koga and Morii, 2005). The occurrence of different polar head groups depends on the archael family and can be used as unique taxonomic marker (Ulrich et al., 2009). Aminolipids, for instance, are prevalent in methanogens and are completely absent in halophiles and thermophiles (Gambacorta et al., 1995).

Bacteria and Eukarya use several mechanisms to maintain the membrane fluidity over a range of temperatures such the regulation of fatty acid composition adapting the degree of branching, saturation and chain length (Zhang and Rock, 2008). The homeoviscous adaption theory states that the lipid compositions in the membrane varies in response to environmental stresses in order to preserve a proper membrane fluidity (Oger and Cario, 2013). However, the exact changes in fatty acid composition in

the membrane upon, for instance, a temperature shift differs from species to species. In *E. coli* the degree of fatty acid unsaturation increases along with a lowering of the temperature while some *Bacillus* species increase the amount of iso-fatty acids with the growth temperatures. The membrane also has to maintain its permeability barrier and in general it is believed that at the growth temperature, the lipid bilayer is a liquid crystalline state. The phase transition temperature at which the membrane is transferred from the crystalline into the liquid state is considered as a very important characteristic, and depends on the length of the hydrocarbon chains, the degree of saturation and the position of methyl groups; in Bacteria, the transition temperature ranges between -20 to up to 65°C compared to archaea where the range is much wider between even up to 100°C , a temperature where some archaea grow (Koga, 2012).

Archaea use different mechanisms to maintain the liquid crystalline phase over the entire growth temperature range. One control mechanism has been reported in the archaeon *Methanococcus jannaschii*, in which case the membrane properties are finely tuned by varying the ratio of diether, macrocyclic diether, and tetraether lipid (Sprott et al., 1991). In contrast, at higher temperatures, hyperthermophilic archaea may incorporate a higher degree of cyclopentane rings in their isoprenoid chain that increase the transition temperature. The presence of such lipid structures in archaea is an indication of a need to preserve the membrane function at the hostile environmental conditions. In particular, the presence of tetraether lipids and the chemically stable ether bonds are major adaptions (Jacquemet et al., 2009). The latter confers resistance to phospholipases attack from other organisms. Despite the general thought that the isoprenoid chains of the ether lipid are involved in thermal resistance, it has been shown that they are not absolutely required for tolerance to high temperature. Archaeol and caldarcheol were also found in the thermophilic methanarchaeota (65°C) and in mesophilic species (37°C ; Koga and Nakano, 2008). This suggests that the remarkable properties of the membranes of hyperthermophiles depend not exclusively on the tetraether composition of their lipids but that other aspects are involved as well. When fully stretched, the tetraether lipids span the entire membrane thickness leading to a monolayer, which is believed to stiffen the membrane in the presence of high growth temperatures (Kates, 1992; Jacquemet et al., 2009). This may also protect the membrane from possible lysis at high temperatures (Jacquemet et al., 2009). The presence of cyclic structures, in particular cyclopentene rings (De Rosa et al., 1980), is a hallmark for high growth temperatures and causes an increased membrane packing and thus a reduction in membrane fluidity (Benvegnu et al., 2008). However, also the characteristics of the polar head groups may influence membrane fluidity since the proper balance of negative and positive charges at the membrane surface is essential for its functioning. Therefore, varying the proportion of different polar head groups might be another way to response to external stresses (Oger and Cario, 2013). Furthermore, modification of the polar headgroups with carbohydrates increases hydrogen bonding between the lipids and thus will influence the stability of the membranes. In halophiles, the presence of glycerol methylphosphate attached to the archaeal moiety, contributes to the low membrane permeability under

high salt concentration (Tenchov et al., 2006). A further example is cold adaption in the psychrophile archaeon *Methanococcoides burtonii*, where an increase in the degree of unsaturation in the isoprenoid chains allow growth at low temperatures as they exist in glacial environments (Russell and Nichols, 1999; Oger and Cario, 2013).

IN VITRO BASED STUDIES ON THE STABILITY OF ARCHAEOAL LIPIDS

Temperature impacts the membrane properties, influencing the ion permeability and phase behavior. It is believed that the unique membrane organization of archaeal tetraether lipid in a monolayer structure along with the presence of ether bonds renders such membranes, thermal resistance. Therefore, many studies have been conducted to understand the higher thermal stability of archaeal membrane lipids compared to the bacterial phospholipids. By comparing liposomes made of a polar lipid fraction from *S. acidocaldarius* and liposomes prepared from a bacterial lipid (POPC) or a synthetic lipid with a phytanyl chain (DPhPC), the importance of the methyl branched isoprenoid chain in membrane stability has been examined (Chang, 1994). When incubated at 100°C for approximately half hour, archaeal liposomes showed a very low ion leakage compared to POPC and DPhPC liposomes that collapse after a few minutes incubation at that temperature. Likewise, only very slow release (8–10%) of the fluorophore carboxyfluorescin (CF) was observed with liposomes composed of *S. acidocaldarius* lipids compared to *E. coli* liposomes (50%) over a time interval of 62 days, while liposomes composed of a lipid extract from the thermophile bacterium *B. stearothermophilus* showed an intermediate stability (Elferink et al., 1994).

The extremely high heat tolerance of archaeal liposomes open opportunities for biotechnological applications with the ability to be stable even after several autoclaving cycles, which might be exploited for biomedical uses. Autoclaving is a common and effective method for decontamination and the possibility to autoclave the archaeosomes vesicle for the drug delivery gives new prospects for the liposomes formulation. Several studies have been performed regarding the ability of archaeosomes to enhance the immune response when used as novel vaccine and drug delivery system (Sprott et al., 1997; Patel and Sprott, 1999; Patel et al., 2000). The superior adjuvant activity of archaeosomes compared to the liposomes evokes sufficient immunostimulation and sustained immune response against cancer or specific infectious diseases. Archaeosomes confer higher stability to the vesicles enhancing the fusion with the immune cells (Krishnan and Sprott, 2008). These properties are dependent on the type of polar head groups and degrees of glycosylation on archaeal lipids (Dicaire et al., 2010; Whitfield et al., 2010). An enhanced cytotoxic cell response was in fact observed in archaeosomes enriched with archaeatidylserine and archaeatidylethanolamine due to their fusogenic nature (Dicaire et al., 2010; Sprott et al., 2012). The effective archaeosomes stability against autoclaving was tested showing a remarkable strength against 2–3 cycles of autoclaving at pH 4.0–10.0 (Brown et al., 2009). Besides the high heat tolerance, archaeal lipids are known to be resistant to conditions of extremely low pH and their low proton permeability contributes to maintaining a constant intracellular pH. Monolayer liposomes reconstituted from the lipid fraction of

S. acidocaldarius indeed exhibited very low proton permeability even at higher temperatures (Elferink et al., 1994; Komatsu and Chong, 1998). However, at acidic conditions the archaeosomes appeared less resistant to autoclaving possibly because of a higher protonation of the polar head groups, influencing hydrogen-bonds formation among the lipids. Liposomes containing long sugar chains linked to the phospholipids show much lower proton permeability than liposomes composed of lipids with only one sugar unit. The glycolipids amount in the polar lipid fraction of *Thermoplasma acidophilum* increases at lower pH and this seems a general mechanism for acidophiles against the chemically unstable conditions (Shimada et al., 2008). The low permeability of such liposomes can be exploited for drug-delivery with the added value of a high resistance against phospholipase A2, B and pancreatic lipase (Choquet et al., 1994). Due to their high pH tolerance they can easily pass the gastro-intestinal tract without damage. Overall, these studies confirm that the presence of tetraethers in archaeal membranes confers these membranes with a remarkable stability against high temperatures, low pH, and high salt. The degree of hydrocarbon chain saturation, the features of the polar head groups and the presence of cyclopentane ring (Dannenmuller et al., 2000) appear of secondary importance in providing stability.

THE LIPID DIVIDE

PROSPECT OF LUCA WITH MIXED MEMBRANE LIPID COMPOSITION

During the last decades many theories have been proposed about the origin of life and how the differentiation between the three domains of life occurred. All of these theories accept the existence of the last universal common ancestor (LUCA), also known as LUCAS or cencestor from which organisms have diverged. Particular attention has been given on the membrane composition of LUCA. Isoprenoids are involved in a wide range of functions, and found both in archaea and bacteria. Whereas fatty acid metabolism is also widely distributed, fatty acid biosynthesis seems underdeveloped in archaea and in some organisms even absent. Therefore, one of the hypothesis is that early life forms were dependent on the presence of membrane lipids with isoprenoid hydrocarbon core. However, the most divergent feature that is at the base of the Lipid Divide is the glycerophosphate backbone. Archaea contains G1P as glycerophosphate moiety while bacteria depend on G3P. These two compounds are synthetized by two different enzymes that are not evolutionary correlated (Koga et al., 1998).

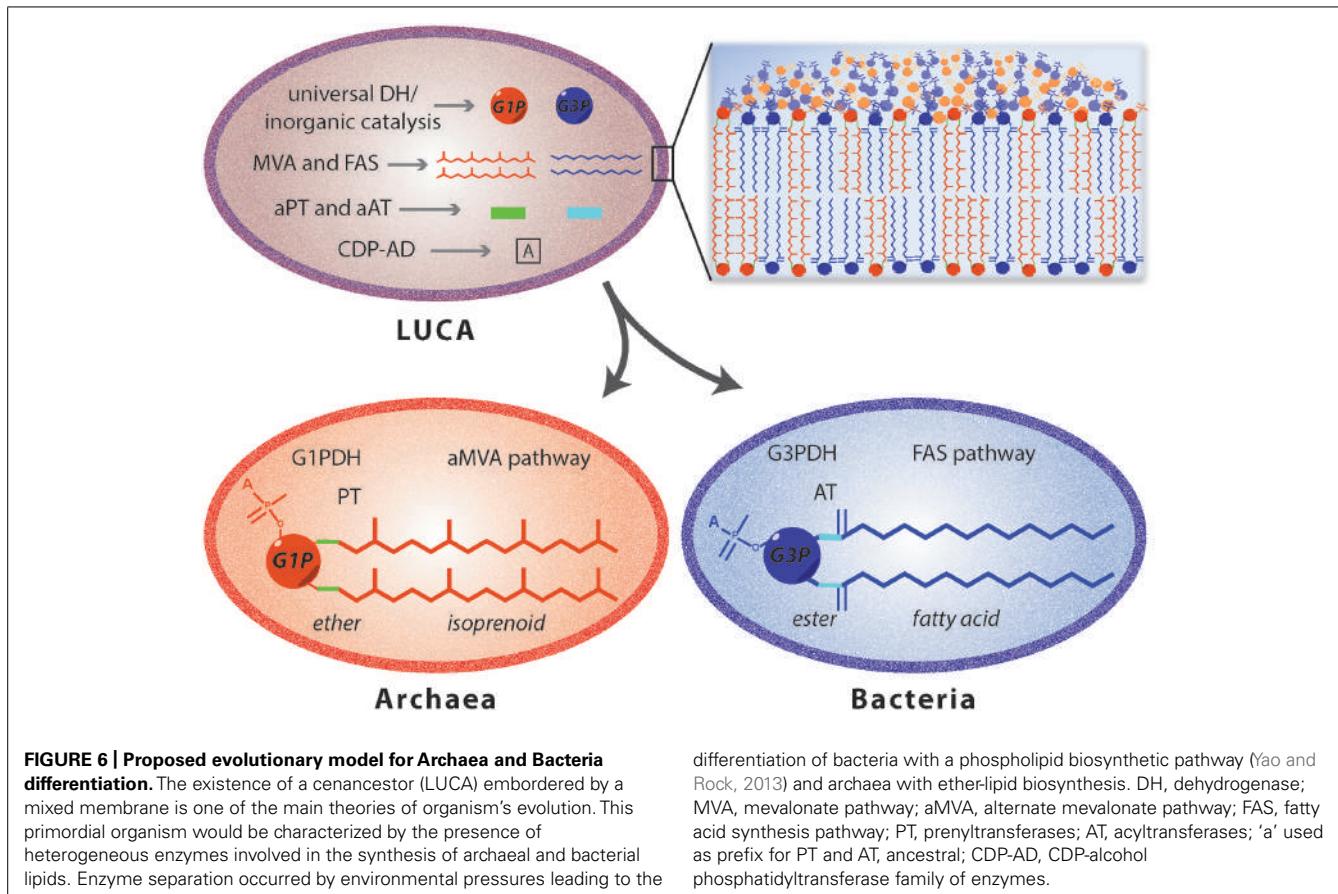
Thus, different hypothesis were suggested to understand the process that brought the ancestor to archaeal, bacterial and eukaryotic organisms. According to Koga et al. (1998) the evolution of the two phospholipid pathways occurred independently leading to the simultaneous appearance of bacterial and archaeal enzymes. This contrast the theory of Martin and Russell (2003), according to whom the two different organisms evolved from an ancestor characterized by iron monosulphide compartment. Another hypothesis (Wächtershäuser, 2003) proposed a three stage process from the pre-cell to the eukaryotic cell. It was suggested that there was a cencestor with a chemically derived heterochiral membrane containing both the enantiomeric forms of the glycerophosphate backbone, which slowly diverged into a more stable homochiral membrane organism leading to the differentiation

between Bacteria and Archaea. This idea implies that such heterochiral mixed membranes are intrinsically unstable leading to the emergence of chiral selective enzymes and a divide between bacteria and archaea. However, hybrid membranes were tested for their stability using a mixture of egg phosphatidylcholine and extracted lipid from *Sulfolobus solfataricus* (Fan et al., 1995) and a higher stability compared to liposomes reconstituted of only archaeal lipid was observed. Moreover a heterochiral membrane consisting of bacterial G3P lipids and archaeal G1P lipids was analyzed (Shimada and Yamagishi, 2011) and the heterochiral membranes were found to be more stable at higher temperatures compared to liposomes prepared from only the bacterial lipids. Thus, based on these studies, the hypothesis of the existence of an ancestor with both the G1PDH and G3PDH would be possible but there must have been other factors that have driven the segregation into the two different domains. The theory of the coexistence of both the enzymes for the glycerophosphate enantiomers productions can be also extended to the other enzymes involved either in the synthesis of isoprenoid and fatty acid based phospholipids (Figure 6). In LUCA, four different membrane lipids may be obtained by the combination of the two glycerophosphate backbones with either isoprenoid or fatty acid hydrocarbon chains (Koga, 2011). Environmental pressure and the need to adapt to extreme conditions may have induced archaea to evolve their membranes (Valentine, 2007; Lombard et al., 2012b). In particular, the use of different hydrocarbon chains in response to the growth environment may have induced the segregation in Archaea and Bacteria and resulted in homochiral membrane formation (Koga, 2014).

DIFFERENTIATION OF MEMBRANE LIPIDS IN ARCHAEA AND BACTERIA

The structural variability found in the membrane lipid composition of archaea and bacteria reflects differences and similarities in the respective biosynthetic pathways. When compared with the well-characterized bacterial ester-lipid biosynthetic pathway, several similarities with the archaeal ether-lipid biosynthesis are evident.

First, the sequence of reactions that yield the final membrane lipid from the building blocks is basically the same even though some enzymes involved in these reactions are equipped with specific features to the lipid of the two different domains. Second, the glycerophosphate backbone in both cases is synthetized by a reduction of DHAP at the 2-OH using NADH as cofactor while the two hydrocarbon chains are linked to the same position on the glycerol moiety. Third, the polar head attachment takes place via a CDP-activated intermediate (Koga, 2014). In particular the peculiar features that distinguish archaeal lipids from bacterial ones occur in the first half of the biosynthetic pathway while the last stages, which involve the replacement of the CDP GROUP with one of the polar groups, is essentially similar in these two domains of life. For the latter reactions, the enzymes belong to the same superfamily and share sequence similarity (Daiyasu et al., 2005; Koga, 2011). On the other hand, the isoprenoid building blocks synthesis differs in Bacteria and in Archaea since it takes place via two different pathway, the DOXP and the alternate MVA pathway, respectively (Koga and Morii, 2007). The other striking difference involved in the lipid divide is the enantiomeric



configuration of the glycerophosphate backbone along with the saturation of the double bonds present on the isoprenoid chains and further modification of the hydrocarbon chains. The remarkable similarities of both biosynthetic pathways are an indication of the existence of a common ancestor with promiscuous enzymatic composition that sufficed for the synthesis of heterochiral membrane lipids.

REVERSING EVOLUTION – SYNTHESIS OF ARCHAEOAL LIPIDS IN BACTERIA?

Did a complex heterochiral membrane ever exist? There are several speculations about the membrane lipid composition of LUCA (Lombard et al., 2012b) and one evolutionary way to approach this is to design a cell which has a heterochiral membrane. The prospect of engineering *E. coli* with membranes harboring archaeal lipids has been initiated in several studies. In the first study, five genes were overexpressed in *E. coli*, four from a hyperthermophilic archaeon *Archaeoglobus fulgidus* (G1P dehydrogenase, GGPP synthase, GGGP synthase and DGGGP synthase) and one from *E. coli* (IPP isomerase; Lai et al., 2009). The enzyme IPP isomerase catalyzes the isomerization of IPP and DMAPP. Together with GGPP synthase, the carbon flux could be increased toward the synthesis of GGPP building blocks which had been demonstrated previously for carotenoid production in *E. coli* (Wang et al., 1999). The activity of each enzyme was monitored by different methods. IPP isomerase and GGPP synthase were analyzed by lycopene based

calorimetric assay, G1P dehydrogenase activity was measured spectrophotometrically to detect DHAP dependent NADH oxidation reaction and using radiolabeled substrates, GGGP synthase activity was measured by thin layer chromatography (TLC) and high performance liquid chromatography (HPLC). To detect the formation of DGGGP in *E. coli*, lipids were extracted and dephosphorylated from a 24 h growing culture and analyzed by liquid chromatography mass spectrometry (LC-MS/MS) using electron spray ionisation (ESI-MS). This study showed that indeed the archaeal lipid biosynthesis pathway is functional in *E. coli*. Unlike G1P dehydrogenase and GGGP synthase, DGGGP synthase protein could not be detected and the amount of DGGGP formed was not quantitated.

In another study, the same four genes but now from the methanogen *Methanosarcina acetivorans* were chosen as it grows at 37°C and are readily overexpressed in *E. coli* (Yokoi et al., 2012). Activity of all the enzymes was monitored in a TLC based assay using radiolabeled substrates. The lipids were extracted from the recombinant *E. coli* cells and unlike in previous study, lipids were not dephosphorylated but analyzed directly using LC/ESI-MS. Interestingly, the result showed that DGGGP synthesized by the archaeal enzymes was further metabolized by *E. coli* endogenous enzymes to form a PG-type derivative of DGGGP which was named DGGGP-Gro. Two speculations were made regarding the endogenous enzymes that might have brought about the reaction. If it emerged from the phospholipid biosynthetic enzymes of

E. coli, it would require three enzymes to recognize and accept the archaeal substrate, namely CdsA, *sn*-G3P transferase and the phosphatase. However, addition of CTP to their *in vitro* TLC based assay did not increase the amount of product formation and no other polar head group attachment was observed. The other possibility is that the *sn*-1-phosphoglycerol group from the osmoregulated periplasmic glucans was transferred to digeranylgeranylglycerol (DGGGOH) directly by the phosphoglyceroltransferase system. The estimated amount of total archaeal membrane lipids extracted from *E. coli* cells in this study was only 60 µg/g wet cells and at these levels it is not possible to study the influence of archaeal lipids on the physical properties of the cytoplasmic membrane. In a follow up study, GGR and ferredoxin (see section saturation of double bonds) were introduced along with the other genes from *Methanoscincus acetivorans* and expressed in *E. coli*. The formation of DGGGP-Gro was reduced yielding mostly saturated archaetidic acid in *E. coli* (Isobe et al., 2014).

The archaeal lipids have also been synthesized *in vitro* using a set of five purified enzymes, two from bacteria and three from archaea. A mutant of FPP synthase of *E. coli* was used that was shown previously to synthesize GGPP, G1P dehydrogenase was from *B. subtilis*, GGGP synthase from methanogen *Methanococcus maripaludis*, DGGGP synthase and CarS from the hyperthermophilic *Archaeoglobus fulgidus*. All enzymes were purified, and by using substrates IPP, FPP, DHAP, and NADH, the enzymes were shown to be able to synthesize CDP-archaeol in the presence of CTP, Mg²⁺ and detergent at 37°C (Jain et al., 2014). The feasibility of synthesizing archaeal lipids in *E. coli* and *in vitro* are promising first steps toward deciphering the biosynthetic pathway further and eventually understanding the properties of a cell with a heterochiral membrane lipid composition.

FUTURE CHALLENGES

Although during the last decade, many of the intimate features of the archaeal lipid biosynthesis pathway have been resolved, there are still several important questions that need to be answered. Understanding the mechanism of tetraether formation and identifying the enzyme(s) involved in the reaction requires thorough investigations. *In vitro* analysis of such a reaction(s) would be great advancement in the field. Various other derivatives of diether and tetraether lipids like cyclopentane and macrocyclic ring formation, glycosylation and formation of crenarchaeol are also not well understood. The pathway refractory of archaeal lipid biosynthesis in *E. coli* is currently incomplete where the amount of archaeal lipids formed in comparison to the *E. coli* lipids is very low. Challenging aspect is to modulate and suppress the endogenous pathway and integrate the archaeal lipid biosynthesis pathway in the genome of the host yielding the exclusive formation of these lipid species. Further structural and biochemical analysis of the enzymes of the pathway from different families of archaea would progress the field and bring it in par with the understanding of bacterial phospholipid biosynthesis. Also, regulation of the phospholipid metabolism is poorly understood and could be enhanced through the use of genetic studies, which now became feasible because of the rapid developments in genetic toolbox in archaea. However, to study essential genes using these techniques is still a challenge.

Recent studies have shown that in spite of the uniqueness of the archaeal membrane lipid structure, they are not as distinct as previously thought. The presence of fatty acids and isoprenoids in the three domains of life and the common mode of polar head group attachment in bacteria and archaea, and the presence of homologues of archaeal G1P dehydrogenase and GGGP synthase in bacteria are a few of the similarities. Other question that still remains unanswered is the exact reason why G1P is ether linked to the isoprenoid hydrocarbon chains and G3P is linked via ester bonds to the fatty acid chains; are there still organisms with such mixed membranes? In order to answer such questions, more biochemical and functional investigation are needed on the archaeal lipid biosynthetic pathway along with a deep phylogenetic analysis.

ACKNOWLEDGMENTS

Samta Jain and Antonella Caforio are supported by the research program of the biobased ecologically balanced sustainable industrial chemistry (BE-BASIC). We thank Ilja Kusters, John van der Oost, Melvin Siliakus and Servé Kengen for fruitful discussions.

REFERENCES

- Benvegnu, T., Lemièvre, L., and Cammas-Marion, S. (2008). Archaeal lipids: innovative materials for biotechnological applications. *Eur. J. Org. Chem.* 2008, 4725–4744. doi: 10.1002/ejoc.200800452
- Boucher, Y., Kamekura, M., and Doolittle, W. F. (2004). Origins and evolution of isoprenoid lipid biosynthesis in archaea. *Mol. Microbiol.* 52, 515–527. doi: 10.1111/j.1365-2958.2004.03992.x
- Brown, D. A., Venegas, B., Cooke, P. H., English, V., and Chong, P. L.-G. (2009). Bipolar tetraether archaeosomes exhibit unusual stability against autoclaving as studied by dynamic light scattering and electron microscopy. *Chem. Phys. Lipids* 159, 95–103. doi: 10.1016/j.chph.2009.03.004
- Chang, E. L. (1994). Unusual thermal stability of liposomes made from bipolar tetraether lipids. *Biochem. Biophys. Res. Commun.* 202, 673–679. doi: 10.1006/bbrc.1994.1983
- Chen, A., and Poulter, C. D. (1993). Purification and characterization of farnesyl diphosphate/geranylgeranyl diphosphate synthase. A thermostable bifunctional enzyme from *Methanobacterium thermoautotrophicum*. *J. Biol. Chem.* 268, 11002–11007.
- Chen, M., and Poulter, C. D. (2010). Characterization of thermophilic archaeal isopentenyl phosphate kinases. *Biochemistry* 49, 207–217. doi: 10.1021/bi9017957
- Choquet, C. G., Patel, G. B., Sprott, G. D., and Beveridge, T. J. (1994). Stability of pressure-extruded liposomes made from archaeobacterial ether lipids. *Appl. Microbiol. Biotechnol.* 42, 375–384. doi: 10.1007/BF00902745
- Daiyasu, H., Kuma, K.-I., Yokoi, T., Morii, H., Koga, Y., and Toh, H. (2005). A study of archaeal enzymes involved in polar lipid synthesis linking amino acid sequence information, genomic contexts and lipid composition. *Archaea* 1, 399–410. doi: 10.1155/2005/452563
- Damsté, J. S. S., Rijpstra, W. I. C., Hopmans, E. C., Jung, M.-Y., Kim, J.-G., Rhee, S.-K., et al. (2012). Intact polar and core glycerol dibiphytanyl glycerol tetraether lipids of group I.1a and I.1b thaumarchaeota in soil. *Appl. Environ. Microbiol.* 78, 6866–6874. doi: 10.1128/AEM.01681-12
- Dannenmuller, O., Arakawa, K., Eguchi, T., Kakinuma, K., Blanc, S., Albrecht, A. M., et al. (2000). Membrane properties of archaeal macrocyclic diether phospholipids. *Chemistry* 6, 645–654. doi: 10.1002/(SICI)1521-3765(20000218)6:4 < 645::AID-CHEM645 > 3.0.CO;2-A
- Dellas, N., and Noel, J. P. (2010). Mutation of archaeal isopentenyl phosphate kinase highlights mechanism and guides phosphorylation of additional isoprenoid monophosphates. *ACS Chem. Biol.* 5, 589–601. doi: 10.1021/cb1000313
- Dellas, N., Thomas, S. T., Manning, G., and Noel, J. P. (2013). Discovery of a metabolic alternative to the classical mevalonate pathway. *Elife* 2, e00672. doi: 10.7554/elife.00672
- De Rosa, M., Esposito, E., Gambacorta, A., Nicolaus, B., and Bu'Lock, J. D. (1980). Effects of temperature on ether lipid composition of *Caldariella acidiphila*. *Phytochemistry* 19, 827–831. doi: 10.1016/0031-9422(80)85120-X

- Dicaire, C. J., Yu, S. H., Whitfield, D. M., and Sprott, G. D. (2010). Isopranoïd- and dipalmityl-aminophospholipid adjuvants impact differently on longevity of CTL immune responses. *J. Liposome Res.* 20, 304–314. doi: 10.3109/08982100903544151
- Eguchi, T., Takeo, H., Morita, M., Kakinuma, K., and Koga, Y. (2000). Unusual double-bond migration as a plausible key reaction in the biosynthesis of the isoprenoidal membrane lipids of methanogenic archaea. *Chem. Commun.* 2000, 1545–1546. doi: 10.1039/b003948i
- Elferink, M. G. L., de Wit, J. G., Driessens, A. J. M., and Konings, W. N. (1994). Stability and proton-permeability of liposomes composed of archaeal tetraether lipids. *Biochim. Biophys. Acta Biomembr.* 1193, 247–254. doi: 10.1016/0005-2736(94)90160-0
- Fan, Q., Relini, A., Cassinadri, D., Gambacorta, A., and Gliozi, A. (1995). Stability against temperature and external agents of vesicles composed of archaeal bolaform lipids and egg PC. *Biochim. Biophys. Acta* 1240, 83–88. doi: 10.1016/0005-2736(95)00157-X
- Fujiwara, S., Yamanaka, A., Hirooka, K., Kobayashi, A., Imanaka, T., and Fukusaki, E.-I. (2004). Temperature-dependent modulation of farnesyl diphosphate/geranylgeranyl diphosphate synthase from hyperthermophilic archaea. *Biochem. Biophys. Res. Commun.* 325, 1066–1074. doi: 10.1016/j.bbrc.2004.10.129
- Gambacorta, A., Caracciolo, G., Trabasso, D., Izzo, I., Spinella, A., and Sodano, G. (2002). Biosynthesis of calditol, the cyclopentanoid containing moiety of the membrane lipids of the archaeon *Sulfolobus solfataricus*. *Tetrahedron Lett.* 43, 451–453. doi: 10.1016/S0040-4039(01)02187-6
- Gambacorta, A., Gliozi, A., and De Rosa, M. (1995). Archaeal lipids and their biotechnological applications. *World J. Microbiol. Biotechnol.* 11, 115–131. doi: 10.1007/BF00339140
- Grochowski, L. L., Xu, H., and White, R. H. (2006). *Methanocoldococcus jannaschii* uses a modified mevalonate pathway for biosynthesis of isopentenyl diphosphate. *J. Bacteriol.* 188, 3192–3198. doi: 10.1128/JB.188.9.3192-3198.2006
- Guldan, H., Matysik, F.-M., Bocola, M., Sterner, R., and Babinger, P. (2011). Functional assignment of an enzyme that catalyzes the synthesis of an archaea-type ether lipid in bacteria. *Angew. Chem. Int. Ed. Engl.* 50, 8188–8191. doi: 10.1002/anie.201101832
- Guldan, H., Sterner, R., and Babinger, P. (2008). Identification and characterization of a bacterial glycerol-1-phosphate dehydrogenase: Ni²⁺-dependent AraM from *Bacillus subtilis*. *Biochemistry* 47, 7376–7384. doi: 10.1021/bi8005779
- Han, J.-S., and Ishikawa, K. (2005). Active site of Zn(2+)-dependent sn-glycerol-1-phosphate dehydrogenase from *Aeropyrum pernix* K1. *Archaea* 1, 311–317. doi: 10.1155/2005/257264
- Han, J.-S., Kosugi, Y., Ishida, H., and Ishikawa, K. (2002). Kinetic study of sn-glycerol-1-phosphate dehydrogenase from the aerobic hyperthermophilic archaeon, *Aeropyrum pernix* K1. *Eur. J. Biochem.* 269, 969–976. doi: 10.1046/j.0014-2956.2001.02731.x
- Hemmi, H., Ikekiri, S., Yamashita, S., and Nishino, T. (2002). Novel medium-chain prenyl diphosphate synthase from the thermoacidophilic archaeon *Sulfolobus solfataricus*. *J. Bacteriol.* 184, 615–620. doi: 10.1128/JB.184.3.615-620.2002
- Hemmi, H., Shibuya, K., Takahashi, Y., Nakayama, T., and Nishino, T. (2004). (S)-2,3-Di-O-geranylgeranylglyceryl phosphate synthase from the thermoacidophilic archaeon *Sulfolobus solfataricus*. Molecular cloning and characterization of a membrane-intrinsic prenyltransferase involved in the biosynthesis of archaeal ether-linked memb. *J. Biol. Chem.* 279, 50197–50203. doi: 10.1074/jbc.M409207200
- Isobe, K., Ogawa, T., Hirose, K., Yokoi, T., Yoshimura, T., and Hemmi, H. (2014). Geranylgeranyl reductase and ferredoxin from *Methanoscincus acetivorans* are required for the synthesis of fully reduced archaeal membrane lipid in *Escherichia coli* cells. *J. Bacteriol.* 196, 417–423. doi: 10.1128/JB.00927-13
- Jacquemet, A., Barbeau, J., Lemiègre, L., and Benvegnu, T. (2009). Archaeal tetraether bipolar lipids: structures, functions and applications. *Biochimie* 91, 711–717. doi: 10.1016/j.biichi.2009.01.006
- Jain, S., Caforio, A., Fodran, P., Lolkema, J. S., Minnaard, A. J., and Driessens, A. J. M. (2014). Identification of CDP-archaeol synthase, a missing link of ether lipid biosynthesis in archaea. *Chem. Biol.* 21, 1392–1401. doi: 10.1016/j.chembiol.2014.07.022
- Kamekura, M., and Kates, M. (1999). Structural diversity of membrane lipids in members of Halobacteriaceae. *Biosci. Biotechnol. Biochem.* 63, 969–972. doi: 10.1271/bbb.63.969
- Kates, M. (1992). Archaeabacterial lipids: structure, biosynthesis and function. *Biochem. Soc. Symp.* 58, 51–72.
- Kates, M. (1993). Biology of halophilic bacteria, Part II. Membrane lipids of extreme halophiles: biosynthesis, function and evolutionary significance. *Experientia* 49, 1027–1036. doi: 10.1007/BF01929909
- Koga, Y. (2011). Early evolution of membrane lipids: how did the lipid divide occur? *J. Mol. Evol.* 72, 274–282. doi: 10.1007/s00239-011-9428-5
- Koga, Y. (2012). Thermal adaptation of the archaeal and bacterial lipid membranes. *Archaea* 2012, 789652. doi: 10.1155/2012/789652
- Koga, Y. (2014). From promiscuity to the lipid divide: on the evolution of distinct membranes in archaea and bacteria. *J. Mol. Evol.* 78, 234–242. doi: 10.1007/s00239-014-9613-4
- Koga, Y., Kyuragi, T., Nishihara, M., and Sone, N. (1998). Did archaeal and bacterial cells arise independently from noncellular Precursors? A hypothesis stating that the advent of membrane phospholipid with enantiomeric glycerophosphate backbones caused the separation of the two lines of descent. *J. Mol. Evol.* 47, 631. doi: 10.1007/PL00006419
- Koga, Y., and Morii, H. (2005). Recent advances in structural research on ether lipids from archaea including comparative and physiological aspects. *Biosci. Biotechnol. Biochem.* 69, 2019–2034. doi: 10.1271/bbb.69.2019
- Koga, Y., and Morii, H. (2007). Biosynthesis of ether-type polar lipids in archaea and evolutionary considerations. *Microbiol. Mol. Biol. Rev.* 71, 97–120. doi: 10.1128/MMBR.00033-06
- Koga, Y., and Nakano, M. (2008). A dendrogram of archaea based on lipid component parts composition and its relationship to rRNA phylogeny. *Syst. Appl. Microbiol.* 31, 169–182. doi: 10.1016/j.syapm.2008.02.005
- Koga, Y., Nishihara, M., Morii, H., and Akagawa-Matsushita, M. (1993). Ether polar lipids of methanogenic bacteria: structures, comparative aspects, and biosyntheses. *Microbiol. Rev.* 57, 164–182.
- Koga, Y., Ohga, M., Tsujimura, M., Morii, H., and Kawarabayasi, Y. (2006). Identification of sn-glycerol-1-phosphate dehydrogenase activity from genomic information on a hyperthermophilic archaeon, *Sulfolobus tokodaii* strain 7. *Biosci. Biotechnol. Biochem.* 70, 282–285. doi: 10.1271/bbb.70.282
- Koga, Y., Sone, N., Noguchi, S., and Morii, H. (2003). Transfer of pro-R hydrogen from NADH to dihydroxyacetonephosphate by sn-glycerol-1-phosphate dehydrogenase from the archaeon *Methanothermobacter thermautotrophicus*. *Biosci. Biotechnol. Biochem.* 67, 1605–1608. doi: 10.1271/bbb.67.1605
- Komatsu, H., and Chong, P. L. (1998). Low permeability of liposomal membranes composed of bipolar tetraether lipids from thermoacidophilic archaeabacterium *Sulfolobus acidocaldarius*. *Biochemistry* 37, 107–115. doi: 10.1021/bi972163e
- Krishnan, L., and Sprott, G. D. (2008). Archaeosome adjuvants: immunological capabilities and mechanism(s) of action. *Vaccine* 26, 2043–2055. doi: 10.1016/j.vaccine.2008.02.026
- Lai, D., Lluncor, B., Schröder, I., Gunsalus, R. P., Liao, J. C., and Monbouquette, H. G. (2009). Reconstruction of the archaeal isoprenoid ether lipid biosynthesis pathway in *Escherichia coli* through digeranylgeranylgluceryl phosphate. *Metab. Eng.* 11, 184–191. doi: 10.1016/j.ymben.2009.01.008
- Lee, P. C., Petri, R., Mijts, B. N., Watts, K. T., and Schmidt-Dannert, C. (2005). Directed evolution of *Escherichia coli* farnesyldiphosphate synthase (IspA) reveals novel structural determinants of chain length specificity. *Metab. Eng.* 7, 18–26. doi: 10.1016/j.ymben.2004.05.003
- Lombard, J., López-García, P., and Moreira, D. (2012a). Phylogenomic investigation of phospholipid synthesis in archaea. *Archaea* 2012, 630910. doi: 10.1155/2012/630910
- Lombard, J., López-García, P., and Moreira, D. (2012b). The early evolution of lipid membranes and the three domains of life. *Nat. Rev. Microbiol.* 10, 507–515. doi: 10.1038/nrmicro2815
- Lombard, J., and Moreira, D. (2011). Origins and early evolution of the mevalonate pathway of isoprenoid biosynthesis in the three domains of life. *Mol. Biol. Evol.* 28, 87–99. doi: 10.1093/molbev/msq177
- Macalady, J. L., Vestling, M. M., Baumler, D., Boekelheide, N., Kaspar, C. W., and Banfield, J. F. (2004). Tetraether-linked membrane monolayers in *Ferroplasma* spp: a key to survival in acid. *Extremophiles* 8, 411–419. doi: 10.1007/s00792-004-0404-5
- Martin, W., and Russell, M. J. (2003). On the origins of cells: a hypothesis for the evolutionary transitions from abiotic geochemistry to chemoautotrophic prokaryotes, and from prokaryotes to nucleated cells. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 358, 59–83; discussion 83–85. doi: 10.1098/rstb.2002.1183

- Matsumi, R., Atomi, H., Driessens, A. J. M., and van der Oost, J. (2011). Isoprenoid biosynthesis in archaea—biochemical and evolutionary implications. *Res. Microbiol.* 162, 39–52. doi: 10.1016/j.resmic.2010.10.003
- Morii, H., Kiyonari, S., Ishino, Y., and Koga, Y. (2009). A novel biosynthetic pathway of archaetidyl-myo-inositol via archaetidyl-myo-inositol phosphate from CDP-archaeol and D-glucose 6-phosphate in methanogenic archaeon *Methanothermobacter thermoautotrophicus* cells. *J. Biol. Chem.* 284, 30766–30774. doi: 10.1074/jbc.M109.034652
- Morii, H., and Koga, Y. (2003). CDP-2,3-Di-O-geranylgeranyl-sn-glycerol:L-serine O-archaetidyltransferase (archaetidylserine synthase) in the methanogenic archaeon *Methanothermobacter thermoautotrophicus*. *J. Bacteriol.* 185, 1181–1189. doi: 10.1128/JB.185.4.1181-1189.2003
- Morii, H., Nishihara, M., and Koga, Y. (2000). CTP-2,3-di-O-geranylgeranyl-sn-glycero-1-phosphate cytidyltransferase in the methanogenic archaeon *Methanothermobacter thermoautotrophicus*. *J. Biol. Chem.* 275, 36568–36574. doi: 10.1074/jbc.M005925200
- Morii, H., Ogawa, M., Fukuda, K., and Taniguchi, H. (2014). Ubiquitous distribution of phosphatidylinositol phosphate synthase and archaetidylinositol phosphate synthase in bacteria and archaea, which contain inositol phospholipid. *Biochem. Biophys. Res. Commun.* 443, 86–90. doi: 10.1016/j.bbrc.2013.11.054
- Nemoto, N. (2003). Purification and characterization of geranylgeranylglycerol phosphate synthase from a thermoacidophilic archaeon, *Thermoplasma acidophilum*. *J. Biochem.* 133, 651–657. doi: 10.1093/jb/mvg083
- Nemoto, N., Shida, Y., Shimada, H., Oshima, T., and Yamagishi, A. (2003). Characterization of the precursor of tetraether lipid biosynthesis in the thermoacidophilic archaeon *Thermoplasma acidophilum*. *Extremophiles* 7, 235–243. doi: 10.1007/s00792-003-0315-x
- Nishihara, M., and Koga, Y. (1997). Purification and properties of sn-glycerol-1-phosphate dehydrogenase from *Methanobacterium thermoautotrophicum*: characterization of the biosynthetic enzyme for the enantiomeric glycerophosphate backbone of ether polar lipids of archaea. *J. Biochem.* 122, 572–576. doi: 10.1093/oxfordjournals.jbchem.a021791
- Ogawa, T., Isobe, K., Mori, T., Asakawa, S., Yoshimura, T., and Hemmi, H. (2014). A novel geranylgeranyl reductase from the methanogenic archaeon *Methanoscarcina acetivorans* displays unique regiospecificity. *FEBS J.* 281, 3165–3176. doi: 10.1111/febs.12851
- Oger, P. M., and Cario, A. (2013). Adaptation of the membrane in archaea. *Biophys. Chem.* 183, 42–56. doi: 10.1016/j.bpc.2013.06.020
- Ohnuma, S., Hirooka, K., Hemmi, H., Ishida, C., Ohto, C., and Nishino, T. (1996). Conversion of product specificity of archaeabacterial geranylgeranyl-diphosphate synthase. Identification of essential amino acid residues for chain length determination of prenyltransferase reaction. *J. Biol. Chem.* 271, 18831–18837. doi: 10.1074/jbc.271.31.18831
- Ohnuma, S., Suzuki, M., and Nishino, T. (1994). Archaeabacterial ether-linked lipid biosynthetic gene. Expression cloning, sequencing, and characterization of geranylgeranyl-diphosphate synthase. *J. Biol. Chem.* 269, 14792–14797.
- Patel, G. B., Agnew, B. J., Deschatelets, L., Fleming, L. P., and Sprott, G. D. (2000). In vitro assessment of archaeosome stability for developing oral delivery systems. *Int. J. Pharm.* 194, 39–49. doi: 10.1016/S0378-5173(99)00331-2
- Patel, G. B., and Sprott, G. D. (1999). Archaeabacterial ether lipid liposomes (archaeosomes) as novel vaccine and drug delivery systems. *Crit. Rev. Biotechnol.* 19, 317–357. doi: 10.1080/0738-859991229170
- Payandeh, J., Fujihashi, M., Gillon, W., and Pai, E. F. (2006). The crystal structure of (S)-3-O-geranylgeranylgluceryl phosphate synthase reveals an ancient fold for an ancient enzyme. *J. Biol. Chem.* 281, 6070–6078. doi: 10.1074/jbc.M509377200
- Peterhoff, D., Beer, B., Rajendran, C., Kumpula, E.-P., Kapetanou, E., Guldan, H., et al. (2014). A comprehensive analysis of the geranylgeranylgluceryl phosphate synthase enzyme family identifies novel members and reveals mechanisms of substrate specificity and quaternary structure organization. *Mol. Microbiol.* 92, 885–899. doi: 10.1111/mmi.12596
- Pitcher, A., Rychlik, N., Hopmans, E. C., Spieck, E., Rijpstra, W. I. C., Ossebaar, J., et al. (2010). Crenarchaeol dominates the membrane lipids of *Candidatus Nitrosphaera gargensis*, a thermophilic group I.1b archaeon. *ISME J.* 4, 542–552. doi: 10.1038/ismej.2009.138
- Podar, M., Makarova, K. S., Graham, D. E., Wolf, Y. I., Koonin, E. V., and Reysenbach, A.-L. (2013). Insights into archaeal evolution and symbiosis from the genomes of a nanoarchaeon and its inferred crenarchaeal host from Obsidian Pool, Yellowstone National Park. *Biol. Dir.* 8, 9. doi: 10.1186/1745-6150-8-9
- Poulter, C. D., Aoki, T., and Daniels, L. (1988). Biosynthesis of isoprenoid membranes in the methanogenic archaeabacterium *Methanospirillum hungatei*. *J. Am. Chem. Soc.* 110, 2620–2624. doi: 10.1021/ja00216a041
- Ren, F., Feng, X., Ko, T.-P., Huang, C.-H., Hu, Y., Chan, H.-C., et al. (2013). Insights into TIM-barrel prenyl transferase mechanisms: crystal structures of PcrB from *Bacillus subtilis* and *Staphylococcus aureus*. *ChemBioChem* 14, 195–199. doi: 10.1002/cbic.201200748
- Russell, N. J., and Nichols, D. S. (1999). Polyunsaturated fatty acids in marine bacteria—a dogma rewritten. *Microbiology* 145 (Pt 4), 767–779. doi: 10.1099/13500872-145-4-767
- Sasaki, D., Fujihashi, M., Iwata, Y., Murakami, M., Yoshimura, T., Hemmi, H., et al. (2011). Structure and mutation analysis of archaeal geranylgeranyl reductase. *J. Mol. Biol.* 409, 543–557. doi: 10.1016/j.jmb.2011.04.002
- Schouten, S., Hopmans, E. C., and Sinninghe Damsté, J. S. (2013). The organic geochemistry of glycerol dialkyl glycerol tetraether lipids: a review. *Org. Geochem.* 54, 19–61. doi: 10.1016/j.orggeochem.2012.09.006
- Shimada, H., Nemoto, N., Shida, Y., Oshima, T., and Yamagishi, A. (2008). Effects of pH and temperature on the composition of polar lipids in *Thermoplasma acidophilum* HO-62. *J. Bacteriol.* 190, 5404–5411. doi: 10.1128/JB.00415-08
- Shimada, H., and Yamagishi, A. (2011). Stability of heterochiral hybrid membrane made of bacterial sn-G3P lipids and archaeal sn-G1P lipids. *Biochemistry* 50, 4114–4120. doi: 10.1021/bi200172d
- Soderberg, T., Chen, A., and Poulter, C. D. (2001). Geranylgeranylgluceryl phosphate synthase. Characterization of the recombinant enzyme from *Methanobacterium thermoautotrophicum*. *Biochemistry* 40, 14847–14854. doi: 10.1021/bi0111799
- Sprott, G. D., Meloche, M., and Richards, J. C. (1991). Proportions of diether, macrocyclic diether, and tetraether lipids in *Methanococcus jannaschii* grown at different temperatures. *J. Bacteriol.* 173, 3907–3910.
- Sprott, G. D., Tolson, D. L., and Patel, G. B. (1997). Archaeosomes as novel antigen delivery systems. *FEMS Microbiol. Lett.* 154, 17–22. doi: 10.1016/S0378-1097(97)00294-2
- Sprott, G. D., Yeung, A., Dicaire, C. J., Yu, S. H., and Whitfield, D. M. (2012). Synthetic archaeosome vaccines containing triglycosylarchaeols can provide additive and long-lasting immune responses that are enhanced by archaetidylserine. *Archaea* 2012, 513231. doi: 10.1155/2012/513231
- Sugai, A., Sakuma, R., Fukuda, I., Kurosawa, N., Itoh, Y. H., Kon, K., et al. (1995). The structure of the core polyol of the ether lipids from *Sulfolobus acidocaldarius*. *Lipids* 30, 339–344. doi: 10.1007/BF02536042
- Tachibana, A., Tanaka, T., Taniguchi, M., and Oi, S. (1993). Potassium-stimulating mechanism of geranylgeranyl diphosphate synthase of *Methanobacterium thermoformicum* SF-4. *J. Biochem.* 114, 389–392.
- Tachibana, A., Yano, Y., Otani, S., Nomura, N., Sako, Y., and Taniguchi, M. (2000). Novel prenyltransferase gene encoding farnesylgeranyl diphosphate synthase from a hyperthermophilic archaeon, *Aeropyrum pernix*. Molecular evolution with alteration in product specificity. *Eur. J. Biochem.* 267, 321–328. doi: 10.1046/j.1432-1327.2000.00967.x
- Tenchov, B., Vescio, E. M., Sprott, G. D., Zeidel, M. L., and Mathai, J. C. (2006). Salt tolerance of archaeal extremely halophilic lipid membranes. *J. Biol. Chem.* 281, 10016–10023. doi: 10.1074/jbc.M600369200
- UDA, I., SUGAI, A., ITOH, Y. H., and ITOH, T. (2004). Variation in molecular species of core lipids from the order *Thermoplasmatales* depends on the growth temperature. *J. Oleo Sci.* 53, 399–404. doi: 10.5650/jos.53.399
- Ulrich, N. P., Gmajner, D., and Raspor, P. (2009). Structural and physicochemical properties of polar lipids from thermophilic archaea. *Appl. Microbiol. Biotechnol.* 84, 249–260. doi: 10.1007/s00253-009-2102-9
- Untersteller, É., Fritz, B., Blériot, Y., and Sinay, P. (1999). The structure of calditol isolated from the thermoacidophilic archaeabacterium *Sulfolobus acidocaldarius*. *C. R. Acad. Sci. Ser. IIIC Chem.* 2, 429–433. doi: 10.1016/S1387-1609(00)88556-4
- Valentine, D. L. (2007). Adaptations to energy stress dictate the ecology and evolution of the Archaea. *Nat. Rev. Microbiol.* 5, 316–323. doi: 10.1038/nrmicro1619
- Vandermoten, S., Haubruege, E., and Cusson, M. (2009). New insights into short-chain prenyltransferases: structural features, evolutionary history and potential for selective inhibition. *Cell. Mol. Life Sci.* 66, 3685–3695. doi: 10.1007/s00018-009-0100-9
- Villanueva, L., Damsté, J. S. S., and Schouten, S. (2014). A re-evaluation of the archaeal membrane lipid biosynthetic pathway. *Nat. Rev. Microbiol.* 12, 438–448. doi: 10.1038/nrmicro3260

- Wächtershäuser, G. (2003). From pre-cells to Eukarya—a tale of two lipids. *Mol. Microbiol.* 47, 13–22. doi: 10.1046/j.1365-2958.2003.03267.x
- Wang, C. W., Oh, M. K., and Liao, J. C. (1999). Engineered isoprenoid pathway enhances astaxanthin production in *Escherichia coli*. *Biotechnol. Bioeng.* 62, 235–241. doi: 10.1002/(SICI)1097-0290(19990120)62:2<235::AID-BIT14>3.0.CO;2-U
- Wang, K., and Ohnuma, S. (1999). Chain-length determination mechanism of isoprenyl diphosphate synthases and implications for molecular evolution. *Trends Biochem. Sci.* 24, 445–451. doi: 10.1016/S0968-0004(99)1464-4
- Whitfield, D. M., Yu, S. H., Dicaire, C. J., and Sprott, G. D. (2010). Development of new glycosylation methodologies for the synthesis of archaeal-derived glycolipid adjuvants. *Carbohydr. Res.* 345, 214–229. doi: 10.1016/j.carres.2009.10.011
- Woese, C. R., and Fox, G. E. (1977). Phylogenetic structure of the prokaryotic domain: the primary kingdoms. *Proc. Natl. Acad. Sci. U.S.A.* 74, 5088–5090. doi: 10.1073/pnas.74.11.5088
- Xu, Q., Eguchi, T., Mathews, I. I., Rife, C. L., Chiu, H.-J., Farr, C. L., et al. (2010). Insights into substrate specificity of geranylgeranyl reductases revealed by the structure of digeranylgeranylglycerophospholipid reductase, an essential enzyme in the biosynthesis of archaeal membrane lipids. *J. Mol. Biol.* 404, 403–417. doi: 10.1016/j.jmb.2010.09.032
- Xue, J., and Ahring, B. K. (2011). Enhancing isoprene production by genetic modification of the 1-deoxy-d-xylulose-5-phosphate pathway in *Bacillus subtilis*. *Appl. Environ. Microbiol.* 77, 2399–2405. doi: 10.1128/AEM.02341-10
- Yao, J., and Rock, C. O. (2013). Phosphatidic acid synthesis in bacteria. *Biochim. Biophys. Acta* 1831, 495–502. doi: 10.1016/j.bbapap.2012.08.018
- Yokoi, T., Isobe, K., Yoshimura, T., and Hemmi, H. (2012). Archaeal phospholipid biosynthetic pathway reconstructed in *Escherichia coli*. *Archaea* 2012, 438931. doi: 10.1155/2012/438931
- Zhang, D., and Poulter, C. D. (1993). Biosynthesis of archaeabacterial ether lipids. Formation of ether linkages by prenyltransferases. *J. Am. Chem. Soc.* 2, 1270–1277. doi: 10.1021/ja00057a008
- Zhang, H., Shibuya, K., Hemmi, H., Nishino, T., and Prestwich, G. D. (2006). Total synthesis of geranylgeranylgluceryl phosphate enantiomers: substrates for characterization of 2,3-O-digeranylgeranylgluceryl phosphate synthase. *Org. Lett.* 8, 943–946. doi: 10.1021/o10530878
- Zhang, Y.-M., and Rock, C. O. (2008). Membrane lipid homeostasis in bacteria. *Nat. Rev. Microbiol.* 6, 222–233. doi: 10.1038/nrmicro1839

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

Received: 26 September 2014; accepted: 06 November 2014; published online: 26 November 2014.

Citation: Jain S, Caforio A and Driessens AJM (2014) Biosynthesis of archaeal membrane ether lipids. *Front. Microbiol.* 5:641. doi: 10.3389/fmicb.2014.00641

This article was submitted to Microbial Physiology and Metabolism, a section of the journal *Frontiers in Microbiology*.

Copyright © 2014 Jain, Caforio and Driessens. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Archaeal S-layer glycoproteins: post-translational modification in the face of extremes

Lina Kandiba and Jerry Eichler *

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

Edited by:

Sonja-Verena Albers, University of Freiburg, Germany

Reviewed by:

Kelly Bidle, Rider University, USA

Reinhard Rachel, University of Regensburg, Germany

***Correspondence:**

Jerry Eichler, Department of Life Sciences, Ben Gurion University of the Negev, PO Box 653, Beersheva 84105, Israel

e-mail: jeichler@bgu.ac.il

Corresponding to the sole or basic component of the surface (S)-layer surrounding the archaeal cell in most known cases, S-layer glycoproteins are in direct contact with the harsh environments that characterize niches where Archaea can thrive. Accordingly, early work examining archaeal S-layer glycoproteins focused on identifying those properties that allow members of this group of proteins to maintain their structural integrity in the face of extremes of temperature, pH, and salinity, as well as other physical challenges. However, with expansion of the list of archaeal strains serving as model systems, as well as growth in the number of molecular tools available for the manipulation of these strains, studies on archaeal S-layer glycoproteins are currently more likely to consider the various post-translational modifications these polypeptides undergo. For instance, archaeal S-layer glycoproteins can undergo proteolytic cleavage, both *N*- and *O*-glycosylation, lipid-modification and oligomerization. In this mini-review, recent findings related to the post-translational modification of archaeal S-layer glycoproteins are considered.

Keywords: Archaea, lipid modification, post-translational modification, protein glycosylation, S-layer glycoprotein

Although Archaea are now recognized as denizens of an enormous range of environments, they remain best known in their capacities as extremophiles, namely organisms able to thrive in some of the most physically challenging settings on the planet. In direct contact with these often hostile surroundings, the archaeal cell surface must not only maintain its integrity but also must carry out a variety of normal physiological functions. In Bacteria, the cell boundary consists of membranes and a peptidoglycan-based cell wall together with other polysaccharide-based molecules (e.g., lipopolysaccharide, teichoic acid) and proteins (Braun, 1975; Lugtenberg and Van Alphen, 1983; Raetz et al., 2007), in many cases comprising a surface (S)-layer (Fagan and Fairweather, 2014). By contrast, the cell wall in Archaea tends to be much simpler. Apart from a number of documented examples (König, 2001), the S-layer, in many cases comprising a single protein species but not always (Peters et al., 1995; Grogan, 1996; Veith et al., 2009), corresponds to the sole cell wall structure (Eichler, 2003; Albers and Meyer, 2011). Studies from several groups studying different Archaea have shown that the S-layer glycoprotein is not just a standardized building block used to generate the two-dimensional lattice of the S-layer but rather that S-layer glycoproteins undergo a variety of post-translational modifications. In this mini-review, recent findings concerning such processing of archaeal S-layer glycoproteins are considered.

DIFFERENCES IN THE SUGAR COATING

The S-layer glycoprotein of the haloarchaeon *Halobacterium salinarum* offered the first example of *N*-glycosylation in a domain other than the Eukarya (Mescher and Strominger, 1976a). This

observation led to a flurry of biochemical activity aimed at describing the composition of *N*-linked glycans decorating the *Hbt. salinarum* S-layer glycoprotein and their biosynthesis (cf. Lechner and Wieland, 1989). However, the lack of sufficient genetic tools for manipulating this and other archaeal species shown to contain glycosylated S-layer proteins (Sumper et al., 1990; Brockl et al., 1991; Karcher et al., 1993) stood in the way of gaining detailed information into such post-translational modification of this protein. Since then, the sequencing of a growing list of archaeal genomes, the development of techniques for manipulating the genetic content of numerous strains and the analytical power of mass spectrometry have been combined to help clear obstacles encountered by earlier studies of S-layer glycoprotein *N*-glycosylation.

Genomic analyses point both to the presence of S-layer glycoproteins and *N*-glycosylation machineries in almost all sequenced Archaea (Magidovich and Eichler, 2009; Albers and Meyer, 2011; Kaminski et al., 2013a). Still, the majority of research on archaeal S-layer glycoprotein *N*-glycosylation to date has focused on *Methanococcus voltae*, *Methanococcus maripaludis*, *Sulfolobus acidocaldarius*, and *Haloflexax volcanii* (for recent review, see Jarrell et al., 2014). In each of these species, genes involved in the assembly and attachment of *N*-linked glycans and often their protein products have been studied. Yet, apart from *S. acidocaldarius*, where *N*-glycosylation is essential for cell survival (Meyer and Albers, 2014), the elimination of such protein processing seemingly has limited impact on the organism (Abu-Qarn and Eichler, 2006; Chaban et al., 2006; VanDyke et al., 2009). As such, one can ask why Archaea devote such a significant number of genes to this post-translational modification. Recent studies on

Hfx. volcanii S-layer glycoprotein have begun to shed light on this point.

The *Hfx. volcanii* S-layer glycoprotein contains seven putative N-glycosylation sites (Sumper et al., 1990). Of these, Asn-13 and Asn-83 are modified by a pentasaccharide comprising a hexose, two hexuronic acids, a methyl ester of hexuronic acid and a mannose (Abu-Qarn et al., 2007; Guan et al., 2010; Magidovich et al., 2010). However, when *Hfx. volcanii* cells are grown in medium containing 1.75 M NaCl (“low salt” conditions) rather than 3.4 M NaCl (“high salt” conditions), Asn-498 is modified by a distinct glycan comprising a sulfated hexose, two hexoses and a rhamnose (Guan et al., 2012). Indeed, the same glycan had been reported earlier as bound to dolichol phosphate in *Hfx. volcanii* grown in the presence of 1.25 M NaCl (Kuntz et al., 1997), the lipid carrier that serves as the platform for N-glycan assembly in this and other Archaea (Lechner et al., 1985; Guan et al., 2010; Calo et al., 2011). As such, it would appear that the *Hfx. volcanii* S-layer glycoprotein undergoes differential N-glycosylation as a function of environmental salinity. While it remains to be defined how such differential S-layer glycoprotein N-glycosylation translates into an appropriate response to changes in surrounding salt levels, the path involved in the biogenesis of the so-called “low-salt” tetrasaccharide has been revealed (Kaminski et al., 2013b). Unexpectedly, the cluster of genes involved does not include an obvious oligosaccharyltransferase, namely that enzyme responsible for transferring a glycan from its lipid carrier to select Asn residues of target proteins (Mohorko et al., 2011). The observation that AglB, the only known archaeal oligosaccharyltransferase (Abu-Qarn and Eichler, 2006; Chaban et al., 2006), is not involved in “low-salt” tetrasaccharide attachment implies the existence of a novel yet undefined enzyme as serving this role (Kaminski et al., 2013b).

It is possible that N-glycosylation of the *Hfx. volcanii* S-layer glycoprotein is even more complicated still. It was recently reported that the Asn-732 position is modified by a sulfoquinovose-hexose-based glycan, N-linked via a chitobiose core (Parente et al., 2014). Moreover, the composition of this glycan was modified in response to the absence or presence of a membrane-localized rhomboid protease. The presence of such a glycan in *Hfx. volcanii* is surprising, given this organism does not contain a homolog of *S. acidocaldarius* Agl3 (Meyer et al., 2011), a UDP-sulfoquinovose synthase responsible for converting UDP-glucose and sodium sulfite into UDP-sulfoquinovose, the activated form of this sugar that is presumably used in *S. acidocaldarius* and presumably *Hfx. volcanii* N-glycosylation. It should also be noted that Asn-732 is found in the same C-terminal region as a cluster of O-glycosylated threonine residues (Sumper et al., 1990) and a lipid anchor (see below). This suggests that post-translational modification of the *Hfx. volcanii* S-layer glycoprotein C-terminal region is a complex event that requires the orchestrated involvement of numerous protein processing pathways.

Unlike *Hfx. volcanii*, which must cope with an environment characterized by molar concentrations of salt, *S. acidocaldarius* is a thermophile that grows optimally at 75–80°C and pH 2–3 (Brock et al., 1972). Possibly due to the challenges presented by its surroundings, not only is N-glycosylation essential in

S. acidocaldarius (Meyer and Albers, 2014) but at least one of the glycoproteins comprising the S-layer in this species (SlaA) presents an extremely high number of N-glycosylation sites. The 1,395 amino acid-long protein contains 31 potential sites for N-glycosylation scattered throughout the polypeptide, translating to an N-glycosylation site every 45 residues on average, with the highest density of such sites being seen in the C-terminal quarter of the protein (Peyfoon et al., 2010). In the region spanning Lys-1004 to Gln-1395, nine of the 11 potential N-glycosylation sites were experimentally verified as being charged with a tri-branched hexasaccharide comprising a glucose, a mannose, two N-acetylglucosamines and a sulfoquinovose, an unusual sugar routinely found in chloroplasts and photosynthetic bacteria (Zahringer et al., 2000; Peyfoon et al., 2010). Indeed, a tallying of the number of putative N-glycosylation sites in 20 different archaeal S-layer glycoproteins reveals that the S-layer glycoproteins of thermo(acido)philes can contain up to 20-fold more such sites than do S-layer glycoproteins in species isolated from other growth conditions (Jarrell et al., 2014). Based on this comparison, it was proposed that such high densities of N-glycosylation sites reflect the need for a rigid and stable cell wall to cope with the challenges of elevated temperatures and acidity encountered by thermo(acido)philic Archaea.

The importance of S-layer glycoprotein glycosylation was also demonstrated in recent work linking the activity of a transcription factor controlling the expression of genes involved in sugar metabolism with S-layer glycoprotein glycosylation and hence, with the maintenance of cell shape in *Hbt. salinarum* (Todor et al., 2014). TrmB binds to the promoters of over 110 genes encoding products involved in various metabolic processes in response to glucose concentrations. Yet, *Hbt. salinarum* does not catabolize glucose, cannot use glucose as the sole carbon or energy source and does not actively transport glucose from the media (Gochnauer and Kushner, 1969; Severina et al., 1990). As such, it was proposed that TrmB activity ensures that sufficient amounts of glucose and other monosaccharides are available for S-layer glycoprotein glycosylation. S-layer glycoprotein glycosylation is directly related to *Hbt. salinarum* maintaining its rod-like shape, with a loss of N-glycosylation leading to the appearance of round cells (Mescher and Strominger, 1976b). Hence, TrmB activity is linked to *Hbt. salinarum* shape and by extension to cell growth, since this process requires the presence of sufficient fully processed S-layer glycoprotein.

Finally, O-glycosylation, where the glycan is linked to the hydroxyl group of Ser or Thr residues, has been reported for both the *Hbt. salinarum* and the *Hfx. volcanii* S-layer glycoproteins (Mescher and Strominger, 1976a; Sumper et al., 1990). In both proteins, Thr-rich regions adjacent to the predicted membrane-spanning domain of the protein are modified with galactose-glucose disaccharides. Still today, nothing is known of the pathways responsible for O-glycosylation in Archaea.

HANGING ON BY A LIPID

Just as S-layer glycoproteins have served as tractable reporters of archaeal protein glycosylation, they have also been central to our understanding of lipid modification in Archaea, namely the covalent linkage of lipid-based groups to a polypeptide chain.

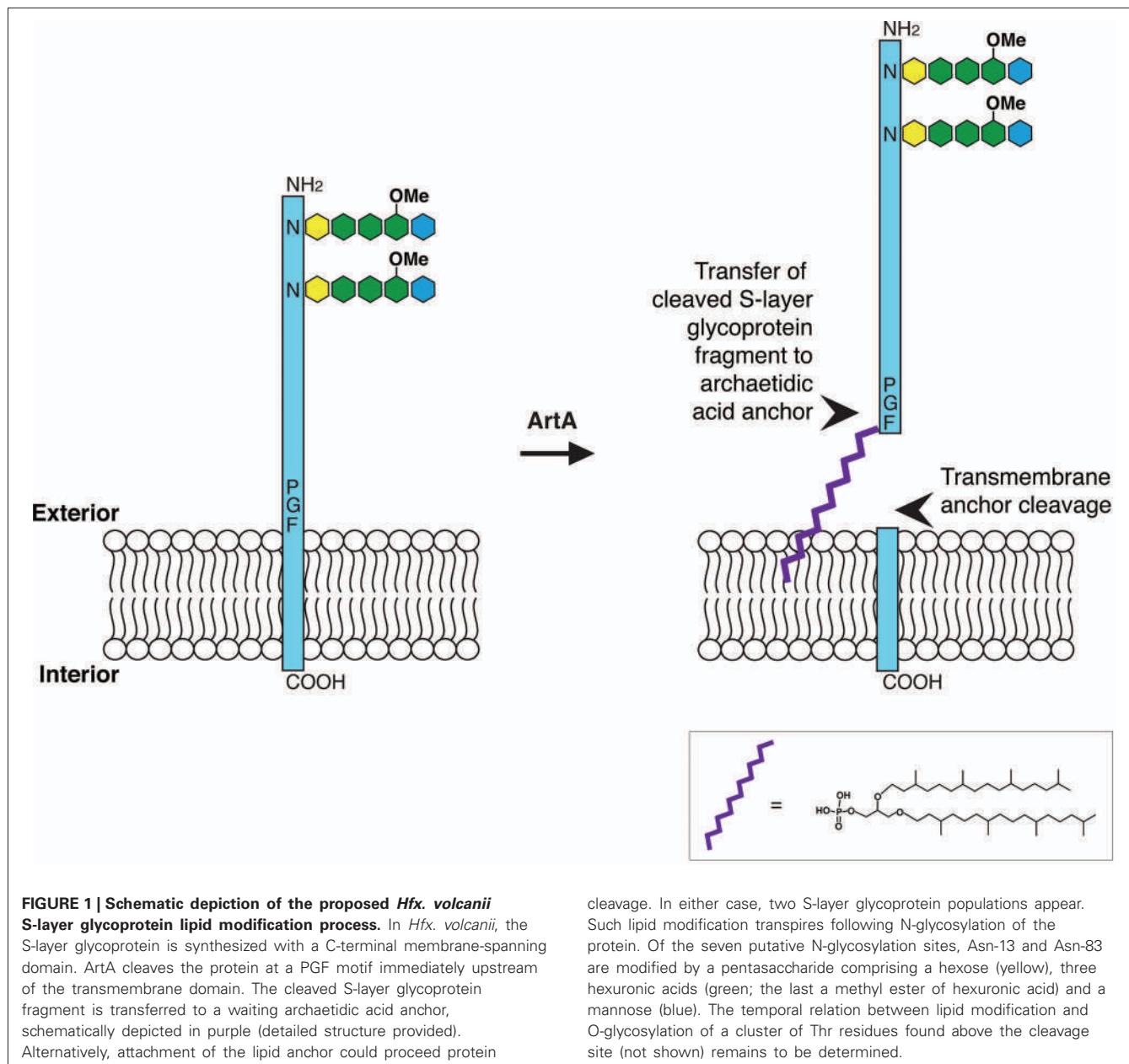


FIGURE 1 | Schematic depiction of the proposed *Hfx. volcanii* S-layer glycoprotein lipid modification process. In *Hfx. volcanii*, the S-layer glycoprotein is synthesized with a C-terminal membrane-spanning domain. ArtA cleaves the protein at a PGF motif immediately upstream of the transmembrane domain. The cleaved S-layer glycoprotein fragment is transferred to a waiting archaetidic acid anchor, schematically depicted in purple (detailed structure provided). Alternatively, attachment of the lipid anchor could proceed protein

Relying on various biochemical approaches, it was shown that the S-layer glycoproteins of *Hbt. salinarum* and *Hfx. volcanii* undergo lipid modification (Kikuchi et al., 1999; Konrad and Eichler, 2002). However, it is only of late that insight into the process of such lipid modification has been provided.

Analysis of the deduced amino acid sequence of the *Hfx. volcanii* S-layer glycoprotein (Sumper et al., 1990) predicts the existence of a 20-residue-long C-terminal membrane-spanning domain, thought to anchor the protein within the membrane. At the same time, it was shown that EDTA treatment leads to the release of the S-layer glycoprotein into the surrounding growth medium (Cline et al., 1989). Solving the paradox of how an apparently integral membrane protein could be solubilized

cleavage. In either case, two S-layer glycoprotein populations appear. Such lipid modification transpires following N-glycosylation of the protein. Of the seven putative N-glycosylation sites, Asn-13 and Asn-83 are modified by a pentasaccharide comprising a hexose (yellow), three hexuronic acids (green; the last a methyl ester of hexuronic acid) and a mannose (blue). The temporal relation between lipid modification and O-glycosylation of a cluster of Thr residues found above the cleavage site (not shown) remains to be determined.

by divalent cation chelation began with studies showing incorporation of radiolabeled prenyl precursors into the *Hfx. volcanii* S-layer glycoprotein. This observation led to the conclusion that the protein is subjected to magnesium-dependent processing associated with lipid modification (Eichler, 2001; Konrad and Eichler, 2002). A decade later, a combination of sequential solubilization steps, native gel electrophoresis and mass spectrometry pointed to the existence of two distinct subpopulations of the S-layer glycoprotein, the first corresponding to an EDTA-solubilized pool anchored to the membrane via a covalently linked archaetidic acid lipid anchor and the second representing detergent-solubilized pool anchored to the membrane likely via the C-terminal membrane-spanning domain

(Kandiba et al., 2013). Both S-layer glycoproteins were shown to be *N*-glycosylated.

In the same period, it was proposed that the Pro-Gly-Phe motif found just upstream of the presumed C-terminal membrane-spanning domain of the *Hfx. volcanii* and *Hbt. salinarum* S-layer glycoproteins is processed similarly as a comparable motif found in certain membrane-linked Gram-positive bacterial proteins (Haft et al., 2012). In Bacteria, this motif is cleaved by a transpeptidase called an exosortase and the released protein is linked to the cell wall via a waiting lipid anchor. Accordingly, genome sequence analysis predicted the existence of an archaeal version of exosortase, termed archaeosortase A (ArtA). Subsequent genetic and biochemical work confirmed not only the existence of ArtA but also its ability to cleave the *Hfx. volcanii* S-layer glycoprotein at the C-terminal Pro-Gly-Phe motif described above (Abdul Halim et al., 2013).

Together, the results of these recent studies argue that in *Hfx. volcanii* (and likely in *Hbt. salinarum* as well), the S-layer glycoprotein is initially synthesized with a C-terminal membrane-spanning domain. This precursor is cleaved by ArtA and the processed S-layer glycoprotein is transferred to a waiting archaeatidic acid anchor lipid anchor in a magnesium-dependent manner. Still, as only selected aspects of this hypothesized pathway (**Figure 1**) have been demonstrated, further experiments await.

CONCLUSION

Corresponding to the building block of the S-layer, the outermost limit of the archaeal cell surface, S-layer glycoproteins are not only in direct contact with the harsh environments Archaea can inhabit but are also amongst the first archaeal proteins to encounter any changes in those environments. Post-translational modification of S-layer glycoproteins offer a rapid and reversible response to such changes. Soon, ongoing efforts in laboratories around the world will not only provide further insight into the pathways recruited for these protein processing events but will also hopefully reveal how such modifications affect S-layer structure and stability. Indeed, with the availability of high resolution structures of archaeal S-glycoproteins (Arbing et al., 2012), it will be possible to obtain detailed understanding of the contributions of post-translational modification to S-layer architecture not only as a function of environment but also of growth stage and other physiological conditions.

AUTHOR CONTRIBUTIONS

All authors made substantial contributions to the acquisition, analysis, and interpretation of data described in this report. All authors critically reviewed the report and approved the final version. All authors agree to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

ACKNOWLEDGMENTS

Research in the Eichler laboratory is supported by the Israel Science Foundation (grant 8/11) and the US Army Research Office (W911NF-11-1-520).

REFERENCES

- Abdul Halim, M. F., Pfeiffer, F., Zou, J., Frisch, A., Haft, D., Wu, S., et al. (2013). *Haloférax volcanii* archaeosortase is required for motility, mating, and C-terminal processing of the S-layer glycoprotein. *Mol. Microbiol.* 88, 1164–1175. doi: 10.1111/mmi.12248
- Abu-Qarn, M., and Eichler, J. (2006). Protein *N*-glycosylation in Archaea: defining *Haloférax volcanii* genes involved in S-layer glycoprotein glycosylation. *Mol. Microbiol.* 61, 511–525. doi: 10.1111/j.1365-2958.2006.05252.x
- Abu-Qarn, M., Yurist-Doutsch, S., Giordano, A., Trauner, A., Morris, H. R., Hitchin, P., et al. (2007). *Haloférax volcanii* AglB and AglD are involved in *N*-glycosylation of the S-layer glycoprotein and proper assembly of the surface layer. *J. Mol. Biol.* 374, 1224–1236. doi: 10.1016/j.jmb.2007.10.042
- Albers, S. V., and Meyer, B. H. (2011). The archaeal cell envelope. *Nat. Rev. Microbiol.* 9, 414–426. doi: 10.1038/nrmicro2576
- Arbing, M. A., Chan, S., Shin, A., Phan, T., Ahn, C. J., Rohlin, L., et al. (2012). Structure of the surface layer of the methanogenic archaean *Methanoscincus acetivorans*. *Proc. Natl. Acad. Sci. U.S.A.* 109, 18112–18117. doi: 10.1073/pnas.1120595109
- Braun, V. (1975). Covalent lipoprotein from the outer membrane of *Escherichia coli*. *Biochim. Biophys. Acta* 415, 335–77. doi: 10.1016/0304-4157(75)90013-1
- Brock, T. D., Brock, K. M., Belly, R. T., and Weiss, R. L. (1972). *Sulfolobus*: a new genus of sulfur-oxidizing bacteria living at low pH and high temperature. *Arch. Microbiol.* 84, 54–68.
- Brockl, G., Behr, M., Fabry, S., Hensel, R., Kaudewitz, H., Biendl, E., et al. (1991). Analysis and nucleotide sequence of the genes encoding the surface-layer glycoproteins of the hyperthermophilic methanogens *Methanothermus fervidus* and *Methanothermus sociabilis*. *Eur. J. Biochem.* 199, 147–152. doi: 10.1111/j.1432-1033.1991.tb16102.x
- Calo, D., Guan, Z., Naparstek, S., and Eichler, J. (2011). Different routes to the same ending: comparing the *N*-glycosylation processes of *Haloférax volcanii* and *Halococcus marismortui*, two halophilic Archaea from the Dead Sea. *Mol. Microbiol.* 81, 1166–1177. doi: 10.1111/j.1365-2958.2011.07781.x
- Chaban, B., Voisin, S., Kelly, J., Logan, S. M., and Jarrell, K. F. (2006). Identification of genes involved in the biosynthesis and attachment of *Methanococcus voltae* *N*-linked glycans: insight into *N*-linked glycosylation pathways in Archaea. *Mol. Microbiol.* 61, 259–268. doi: 10.1111/j.1365-2958.2006.05226.x
- Cline, S. W., Lam, W. L., Charlebois, R. L., Schalkwyk, L. C., and Doolittle, W. F. (1989). Transformation methods for halophilic archaeabacterial. *Can. J. Microbiol.* 35, 148–152. doi: 10.1139/m89-022
- Eichler, J. (2001). Post-translational modification unrelated to protein glycosylation follows translocation of the S-layer glycoprotein across the plasma membrane of the haloarchaeon *Haloférax volcanii*. *Eur. J. Biochem.* 268, 4366–4373. doi: 10.1046/j.1432-1327.2001.02361.x
- Eichler, J. (2003). Facing extremes: archaeal surface-layer (glyco-) proteins. *Microbiology* 149, 3347–3351. doi: 10.1099/mic.0.26591-0
- Fagan, R. P., and Fairweather, N. F. (2014). Biogenesis and functions of bacterial S-layers. *Nat. Rev. Microbiol.* 12, 211–222. doi: 10.1038/nrmicro3213
- Gochnauer, M. B., and Kushner, D. J. (1969). Growth and nutrition of extremely halophilic bacteria. *Can. J. Microbiol.* 15, 1157–1165. doi: 10.1139/m69-211
- Grogan, D. W. (1996). Organization and interactions of cell envelope proteins of the extreme thermoacidophile *Sulfolobus acidocaldarius*. *Can. J. Microbiol.* 42, 1163–1171. doi: 10.1139/m96-148
- Guan, Z., Naparstek, S., Calo, D., and Eichler, J. (2012). Protein glycosylation as an adaptive response in Archaea: growth at different salt concentrations leads to alterations in *Haloférax volcanii* S-layer glycoprotein *N*-glycosylation. *Environ. Microbiol.* 14, 743–753. doi: 10.1111/j.1462-2920.2011.02625.x
- Guan, Z., Naparstek, S., Kaminski, L., Konrad, Z., and Eichler, J. (2010). Distinct glycan-charged phosphodolichol carriers are required for the assembly of the pentasaccharide *N*-linked to the *Haloférax volcanii* S-layer glycoprotein. *Mol. Microbiol.* 78, 1294–1303. doi: 10.1111/j.1365-2958.2010.07405.x
- Haft, D. H., Payne, S. H., and Selengut, J. D. (2012). Archaeosortases and exosortases are widely distributed systems linking membrane transit with posttranslational modification. *J. Bacteriol.* 194, 36–48. doi: 10.1128/JB.06026-11
- Jarrell, K. F., Ding, Y., Meyer, B. H., Albers, S. V., Kaminski, L., and Eichler, J. (2014). *N*-linked glycosylation in Archaea: a structural, functional, and genetic analysis. *Microbiol. Mol. Biol. Rev.* 78, 304–341. doi: 10.1128/MMBR.00052-13
- Kaminski, L., Lurie-Weinberger, M. N., Allers, T., Gophna, U., and Eichler, J. (2013a). Phylogenetic- and genome-derived insight into the evolutionary

- history of *N*-glycosylation in Archaea. *Mol. Phylogenet. Evol.* 68, 327–339. doi: 10.1016/j.ympev.2013.03.024
- Kaminski, L., Guan, Z., Yurist-Doutsch, S., and Eichler, J. (2013b). Two distinct *N*-glycosylation pathways process the *Haloferax volcanii* S-layer glycoprotein upon changes in environmental salinity. *mBio* 4:e00716-13. doi: 10.1128/mBio.00716-13
- Kandiba, L., Guan, Z., and Eichler, J. (2013). Lipid modification gives rise to two distinct *Haloferax volcanii* S-layer glycoprotein populations. *Biochim. Biophys. Acta* 1828, 938–943. doi: 10.1016/j.bbamem.2012.11.023
- Karcher, U., Schroder, H., Haslinger, E., Allmaier, G., Schreiner, R., Wieland, F., et al. (1993). Primary structure of the heterosaccharide of the surface glycoprotein of *Methanothermus fervidus*. *J. Biol. Chem.* 268, 26821–26826.
- Kikuchi, A., Sagami, H., and Ogura, K. (1999). Evidence for covalent attachment of diphentanylglycerol phosphate to the cell-surface glycoprotein of *Halobacterium halobium*. *J. Biol. Chem.* 274, 18011–18016. doi: 10.1074/jbc.274.25.18011
- König, H. (2001). Archaeal cell walls. *eLS*. doi: 10.1038/npg.els.0000384
- Konrad, Z., and Eichler, J. (2002). Lipid modification of proteins in Archaea: attachment of a mevalonic acid-based lipid moiety to the surface-layer glycoprotein of *Haloferax volcanii* follows protein translocation. *Biochem. J.* 366, 959–964. doi: 10.1042/BJ20020757
- Kuntz, C., Sonnenbichler, J., Sonnenbichler, I., Sumper, M., and Zeitler, R. (1997). Isolation and characterization of dolichol-linked oligosaccharides from *Haloferax volcanii*. *Glycobiology* 7, 897–904. doi: 10.1093/glycob/7.7.897
- Lechner, J., and Wieland, F. (1989). Structure and biosynthesis of prokaryotic glycoproteins. *Annu. Rev. Biochem.* 58, 173–194. doi: 10.1146/annurev.bi.58.070189.001133
- Lechner, J., Wieland, F., and Sumper, M. (1985). Biosynthesis of sulfated saccharides *N*-glycosidically linked to the protein via glucose. Purification and identification of sulfated dolichyl monophosphoryl tetrasaccharides from halobacteria. *J. Biol. Chem.* 260, 860–866.
- Lugtenberg, B., and Van Alphen, L. (1983). Molecular architecture and functioning of the outer membrane of *Escherichia coli* and other gram-negative bacteria. *Biochim. Biophys. Acta* 737, 51–115. doi: 10.1016/0304-4157(83)90014-X
- Magidovich, H., and Eichler, J. (2009). Glycosyltransferases and oligosaccharyl transferases in Archaea: putative components of the *N*-glycosylation pathway in the third domain of life. *FEMS Microbiol. Lett.* 300, 122–130. doi: 10.1111/j.1574-6968.2009.01775.x
- Magidovich, H., Yurist-Doutsch, S., Konrad, Z., Ventura, V. V., Dell, A., Hitchen, P. G., et al. (2010). AgIP is a *S*-adenosyl-L-methionine-dependent methyltransferase that participates in the *N*-glycosylation pathway of *Haloferax volcanii*. *Mol. Microbiol.* 76, 190–199. doi: 10.1111/j.1365-2958.2010.07090.x
- Mescher, M. F., and Strominger, J. L. (1976a). Purification and characterization of a prokaryotic glucoprotein from the cell envelope of *Halobacterium salinarum*. *J. Biol. Chem.* 251, 2005–2014.
- Mescher, M. F., and Strominger, J. L. (1976b). Structural (shape-maintaining) role of the cell surface glycoprotein of *Halobacterium salinarum*. *Proc. Natl. Acad. Sci. U.S.A.* 73, 2687–2691. doi: 10.1073/pnas.73.8.2687
- Meyer, B. H., and Albers, S. V. (2014). AglB, catalyzing the oligosaccharyl transferase step of the archaeal *N*-glycosylation process, is essential in the thermoacidophilic crenarchaeon *Sulfolobus acidocaldarius*. *Microbiologyopen* 3, 531–543. doi: 10.1002/mbo3.185
- Meyer, B. H., Zolghadr, B., Peyfoon, E., Pabst, M., Panico, M., Morris, H. R., et al. (2011). Sulfoquinovose synthase—an important enzyme in the *N*-glycosylation pathway of *Sulfolobus acidocaldarius*. *Mol. Microbiol.* 82, 1150–1163. doi: 10.1111/j.1365-2958.2011.07875.x
- Mohorko, E., Glockshuber, R., and Aeby, M. (2011). Oligosaccharyltransferase: the central enzyme of *N*-linked protein glycosylation. *J. Inherit. Metab. Dis.* 34, 869–878. doi: 10.1007/s10545-011-9337-1
- Parente, J., Casabuono, A., Ferrari, M. C., Paggi, R. A., De Castro, R. E., Couto, A. S., et al. (2014). A rhomboid protease gene deletion affects a novel oligosaccharide *N*-linked to the S-layer glycoprotein of *Haloferax volcanii*. *J. Biol. Chem.* 289, 11304–11317. doi: 10.1074/jbc.M113.546531
- Peters, J., Nitsch, M., Kuhlmorgen, B., Golbik, R., Lupas, A., Kellermann, J., et al. (1995). Tetrabrachion: a filamentous archaeabacterial surface protein assembly of unusual structure and extreme stability. *J. Mol. Biol.* 245, 385–401. doi: 10.1006/jmbi.1994.0032
- Peyfoon, E., Meyer, B., Hitchen, P. G., Panico, M., Morris, H. R., Haslam, S. M., et al. (2010). The S-layer glycoprotein of the crenarchaeote *Sulfolobus acidocaldarius* is glycosylated at multiple sites with chitobiose-linked *N*-glycans. *Archaea* pii: 754101. doi: 10.1155/2010/754101
- Raetz, C. R., Reynolds, C. M., Trent, M. S., and Bishop, R. E. (2007). Lipid A modification systems in gram-negative bacteria. *Annu. Rev. Biochem.* 76, 295–329. doi: 10.1146/annurev.biochem.76.010307.145803
- Severina, L. O., Pimenov, N. V., and Plakunov, V. K. (1990). Glucose transport into the extremely halophilic archaeabacteria. *Arch. Microbiol.* 155, 131–136. doi: 10.1007/BF00248606
- Sumper, M., Berg, E., Mengele, R., and Strobel, I. (1990). Primary structure and glycosylation of the S-layer protein of *Haloferax volcanii*. *J. Bacteriol.* 172, 7111–7118.
- Todor, H., Dulmage, K., Gillum, N., Bain, J. R., Muehlbauer, M. J., and Schmid, A. K. (2014). A transcription factor links growth rate and metabolism in the hypersaline adapted archaeon *Halobacterium salinarum*. *Mol. Microbiol.* 93, 1172–1182. doi: 10.1111/mmi.12726
- VanDyke, D. J., Wu, J., Logan, S. M., Kelly, J. F., Mizuno, S., Aizawa, S., et al. (2009). Identification of genes involved in the assembly and attachment of a novel flagellin *N*-linked tetrasaccharide important for motility in the archaeon *Methanococcus maripaludis*. *Mol. Microbiol.* 72, 633–644. doi: 10.1111/j.1365-2958.2009.06671.x
- Weith, A., Klingl, A., Zolghadr, B., Lauber, K., Mentele, R., Lottspeich, F., et al. (2009). *Acidianus*, *Sulfolobus* and *Metallosphaera* surface layers: structure, composition and gene expression. *Mol. Microbiol.* 73, 58–72. doi: 10.1111/j.1365-2958.2009.06746.x
- Zahringer, U., Moll, H., Hettmann, T., Knirel, Y. A., and Schafer, G. (2000). Cytochrome b558/566 from the archaeon *Sulfolobus acidocaldarius* has a unique Asn-linked highly branched hexasaccharide chain containing 6-sulfoquinovose. *Eur. J. Biochem.* 267, 4144–4149. doi: 10.1046/j.1432-1327.2000.01446.x

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

Received: 15 September 2014; accepted: 14 November 2014; published online: 26 November 2014.

Citation: Kandiba L and Eichler J (2014) Archaeal S-layer glycoproteins: post-translational modification in the face of extremes. *Front. Microbiol.* 5:661. doi: 10.3389/fmicb.2014.00661

This article was submitted to Microbial Physiology and Metabolism, a section of the journal Frontiers in Microbiology.

Copyright © 2014 Kandiba and Eichler. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Cytochromes c in Archaea: distribution, maturation, cell architecture, and the special case of *Ignicoccus hospitalis*

Arnulf Kletzin^{1*}, Thomas Heimerl^{2†}, Jennifer Flechsler², Laura van Niftrik³, Reinhard Rachel² and Andreas Klingl⁴

OPEN ACCESS

Edited by:

Sonja-Verena Albers,
University of Freiburg, Germany

Reviewed by:

Ulrike Kappler,

University of Queensland, Australia
Christiane Dahl,
Rheinische
Friedrich-Wilhelms-Universität Bonn,
Germany

***Correspondence:**

Arnulf Kletzin,
Department of Biology, Sulfur
Biochemistry and Microbial
Bioenergetics, Technische Universität
Darmstadt, Schnittspahnstraße 10,
64287 Darmstadt, Germany
Kletzin@bio.tu-darmstadt.de

†Present Address:

Thomas Heimerl,
LOEWE Research Center for Synthetic
Microbiology (SYNMIKRO), Philipps
University of Marburg, Marburg,
Germany

Specialty section:

This article was submitted to
Microbial Physiology and Metabolism,
a section of the journal
Frontiers in Microbiology

Received: 12 January 2015

Accepted: 23 April 2015

Published: 12 May 2015

Citation:

Kletzin A, Heimerl T, Flechsler J, van Niftrik L, Rachel R and Klingl A (2015)
Cytochromes c in Archaea:
distribution, maturation, cell
architecture, and the special case of
Ignicoccus hospitalis.
Front. Microbiol. 6:439.
doi: 10.3389/fmicb.2015.00439

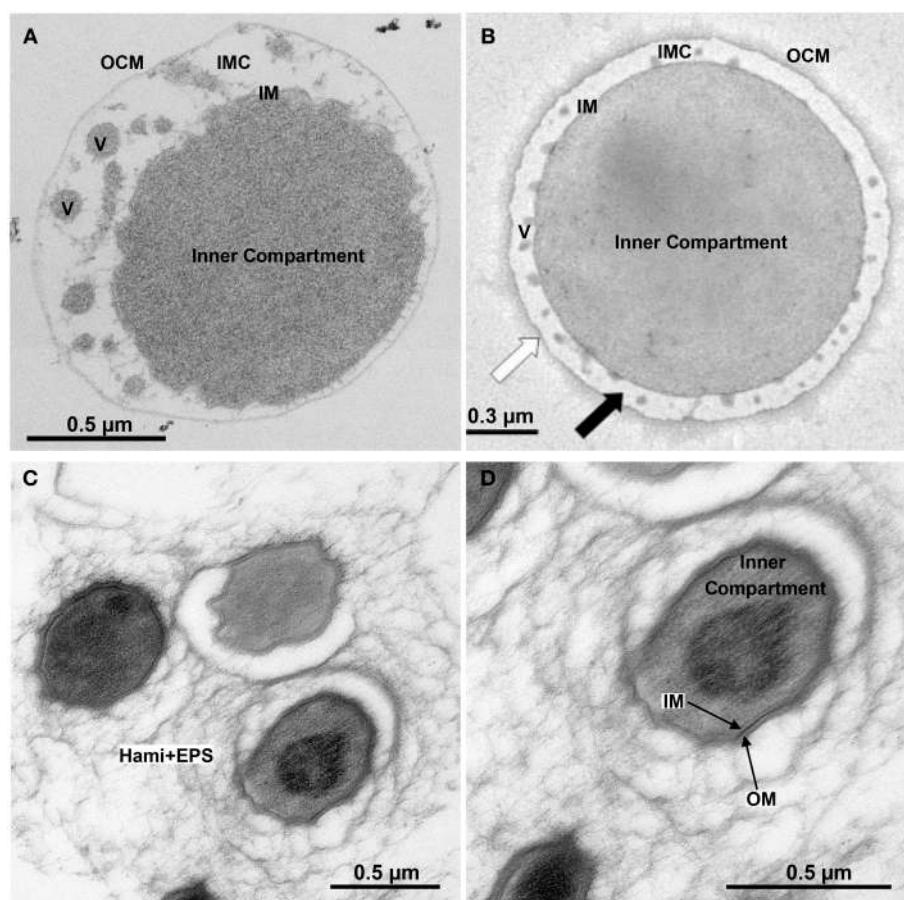
¹ Department of Biology, Sulfur Biochemistry and Microbial Bioenergetics, Technische Universität Darmstadt, Darmstadt, Germany, ² Fakultät für Biologie und Vorklinische Medizin, Zentrum für Elektronenmikroskopie, Universität Regensburg, Regensburg, Germany, ³ Department of Microbiology, Institute for Water and Wetland Research, Radboud University Nijmegen, Nijmegen, Netherlands, ⁴ Department of Biology I, Plant Development, Biocenter LMU Munich, Planegg-Martinsried, Germany

Cytochromes c (Cyc) are widespread electron transfer proteins and important enzymes in the global nitrogen and sulfur cycles. The distribution of Cyc in more than 300 archaeal proteomes deduced from sequence was analyzed with computational methods including pattern and similarity searches, secondary and tertiary structure prediction. Two hundred and fifty-eight predicted Cyc (with single, double, or multiple heme c attachment sites) were found in some but not all species of the *Desulfurococcales*, *Thermoproteales*, *Archaeoglobales*, *Methanococcinales*, *Halobacteriales*, and in two single-cell genome sequences of the *Thermoplasmatales*, all of them Cren- or Euryarchaeota. Other archaeal phyla including the *Thaumarchaeota* are so far free of these proteins. The archaeal Cyc sequences were bundled into 54 clusters of mutual similarity, some of which were specific for Archaea while others had homologs in the Bacteria. The cytochrome c maturation system I (CCM) was the only one found. The highest number and variability of Cyc were present in those species with known or predicted metal oxidation and/or reduction capabilities. Paradoxical findings were made in the haloarchaea: several Cyc had been purified biochemically but corresponding proteins were not found in the proteomes. The results are discussed with emphasis on cell morphologies and envelopes and especially for double-membraned Archaea-like *Ignicoccus hospitalis*. A comparison is made with compartmentalized bacteria such as the *Planctomycetes* of the Anammox group with a focus on the putative localization and roles of the Cyc and other electron transport proteins.

Keywords: cytochrome c, Archaea, *Ignicoccus hospitalis*, ANME, anammox planctomycetes, bioinformatics, molecular modeling

Introduction

The chemolithotrophic, hyperthermophilic Archaeon *Ignicoccus hospitalis* is unusual in several aspects (Huber et al., 2012). First, it is the only host of the symbiotic and/or parasitic Archaeon *Nanoarchaeum equitans*. Second, *I. hospitalis* cells do not possess a cell wall. Instead they comprise

**FIGURE 1 | Ultrastructure of double-membraned Archaea.**

Transmission electron micrographs of ultrathin sections of high-pressure frozen, Epon-embedded cells; (**A**) *Ignicoccus hospitalis*; (**B**) *Methanomassiliicoccus luminyensis* taken and modified from Dridi et al. (2012), with permission from the *International Journal of Systematic and*

Evolutionary Microbiology, confirmation number 11261287; (**C,D**)

Candidatus “Altarchaeum hamiconexum.” Abbreviations and explanation: IMC, intermembrane compartment; IM, inner membrane; OCM, outer cellular membrane; V, vesicles; EPS, extracellular polymeric substances; Hami, extracellular long hooked pili (Moissl et al., 2005).

two membrane systems: an inner membrane (IM) encompassing the densely contrasted inner compartment, which contains DNA, ribosomes, and presumably many biosynthetic enzymes (**Figure 1**; Huber et al., 2012). The outer cellular membrane (OCM) surrounds the cell and contains regularly arrayed small hydrophobic proteins (Burghardt et al., 2007; Huber et al., 2012). A lightly contrasted intermembrane compartment separates both membranes (IMC, 50–1000 nm in width). The IMC contains densely contrasted tubes and vesicles directly involved in the interplay between both membranes (Huber et al., 2012; Meyer et al., 2014). The energy-converting enzymes ATP synthase, hydrogenase, sulfur reductase, and acetyl-CoA synthase are located in the OCM representing the cellular and bioenergetic boundary of the cell from the non-living environment (Küper et al., 2010; Meyer et al., 2012). Therefore, the OCM of *I. hospitalis* is not equivalent to the outer membrane of Gram-negative bacteria (Huber et al., 2012).

The coloration of *I. hospitalis* cells is a third unusual aspect: soluble extracts and membrane fractions are brightly red

resulting from a high content of soluble and membrane-bound cytochromes *c* (Cyt*c*). We had purified three different Cyt*c* from *I. hospitalis* cells, however, we can so far only speculate about their *in vivo* function (Naß et al., 2014). Two of these proteins, named Igni_0955 and Igni_1359 after their GenBank locus tag numbers, were present in soluble and membrane extracts, while the third one (Igni_0530) was present only in the membrane fractions.

Cytochromes *c* are widely distributed in the living world. For example, *Pseudomonas*, *Paracoccus*, and *Thermus* species possess the genes for the canonical mitochondrial-type respiratory chain including the *bc₁* complex (complex III) and the soluble monoheme Cyt*c* as electron carrier between complexes III and IV (reviewed, for example, in Mooser et al., 2006; Noor and Soulimane, 2013). Among Archaea, the bioenergetics and the composition of electron transport chains was most thoroughly studied in *Pyrococcus*/*Thermococcus* spp., methanogens, haloarchaea, and *Sulfobolales* (Schäfer et al., 1999; Schäfer, 2004; Thauer et al., 2008; Mayer and Müller, 2014). Among these, only those methanogens of the *Methanosarcinales* order

possess Cytc, whereas they were not detectable—biochemically or by sequence comparisons—in the other taxa or in other methanogens (Thauer et al., 2008).

The hallmark of Cytc is a covalent ligation of a heme *b* moiety to the protein backbone. In most cases, two cysteine side chains—usually present in a sequence motif CxxCH—form thioether linkages to the heme backbone. The histidine provides the proximal axial ligand of the octahedral coordination sphere of the iron in the center of the heme. The distal axial ligand comes from a distant His, Met or, less frequently, other residues. Variations of this theme may involve penta- instead of hexa-coordinated hemes as for example in Cytc', a CxxCK heme-binding motif (e.g., in nitrite reductases; Lockwood et al., 2011) or different spacing of the cysteine residues (Kern et al., 2011). Motif variations usually occur in multiheme cytochromes *c* (MCC) acting as enzymes and not as electron transfer proteins.

The double thioether linkage is formed by maturation proteins, which are grouped by phylogenetic and functional relationship into five systems (Allen et al., 2006; Allen, 2011; de Vitry, 2011; Simon and Hederstedt, 2011; Stevens et al., 2011). In most bacteria, Cytc maturation (CCM) takes place on the positive (p) side of the cytoplasmic membrane (maturation systems I and II; Simon and Hederstedt, 2011; Stevens et al., 2011). The apoproteins are transported across the cytoplasmic membrane by the General Secretory Pathway (GSP) (Sec-System) so that they carry a recognizable signal sequence at their N-termini, which is—apart from the CxxCH motif—the second feature important for bioinformatic prediction of these proteins. System I or CCM (Cytc maturation) consists of up to nine different proteins including a heme ligase, chaperones, ATP-transporters, and protein disulfide isomerases (Stevens et al., 2011; Verissimo and Daldal, 2014). It occurs in Alpha- and other Gammaproteobacteria and it was identified in Archaea during a previous study (Allen et al., 2006). System II consists of less and mostly unrelated proteins compared to System I (Simon and Hederstedt, 2011).

The number of studies conducted about occurrence and function of Cytc in Archaea is limited and no systematic survey was so far performed. Apart from *I. hospitalis* (Naß et al., 2014), Cytc were found biochemically in the hydrogen-oxidizing and sulfur-reducing complex of the related Archaea *Pyrodictium abyssi* and *P. brockii* (Pihl et al., 1992; Dirmeier et al., 1998), in a *bc₁* complex from the likewise related microaerophilic *Aeropyrum pernix* (Kabashima and Sakamoto, 2011), in the nitrate reducer *Pyrobaculum aerophilum* (Feinberg and Holden, 2006; all of them hyperthermophilic *Crenarchaeota*), in cultured (*Methanosarcina* spp.) and uncultured species (ANME-1 and ANME-2) of the *Methanosarcinales* and in several haloarchaea (all Euryarchaeota; Kamlage and Blaut, 1992; Scharf et al., 1997; Sreeramulu et al., 1998; Sreeramulu, 2003; Meyer-Dierks et al., 2010; Wang et al., 2011, 2014). Surprisingly, experimental gene identification was accomplished only for a few of these species including the three multiheme Cytc from *I. hospitalis* and of the *bc₁* complex of the related *A. pernix* (Kabashima and Sakamoto, 2011; Naß et al., 2014).

When looking at *I. hospitalis* and trying to put the pieces of this puzzle together, questions arise about the distribution of Cytc in different types of archaeal cells, about their targeting and

about the nature and location of the biogenesis system. Since occurrence and distribution of Cytc in Archaea was not recently analyzed in detail, we present here the results of a systematic computational survey. The results are discussed with respect to cell ultrastructure and the physiology of the different archaeal with a special focus on the comparison of *I. hospitalis* with other single, double, and triple-membraned Archaea and Bacteria.

Materials and Methods

Bioinformatic Procedures

The complete non-redundant set of archaeal proteins was downloaded July 23rd, 2014 from Uniprot database in FASTA format (<http://www.uniprot.org/>). In addition, archaeal sequences deposited at GenBank in 2014 were downloaded January 6th, 2015, from the non-redundant protein database (NR). Both sets of sequences were curated for duplicate species and combined. The total set of 883,607 proteins (Table 1) were analyzed in installments of up to 30,000 sequences for the amino acid pattern CxxCH using the 3of5 algorithm (Seiler et al., 2006) installed locally at the HUSAR Sequence Analysis Facility at the German Cancer Research Center, Heidelberg (<http://genius.emblnet.dkfz-heidelberg.de/menu/w2h/w2hdkfz/>; 3of5 web server available at <http://www.dkfz.de/mga2/3of5/3of5.html>). The hits (Table 1) were converted into a tab-delimited list of accession numbers and corresponding hit motifs using the advanced “find and replace” features of Microsoft Word and finally inserted into a Microsoft Excel work sheet (Table S1). A list of database accession numbers (Uniprot identifiers and GenBank GI numbers) was generated from the appropriate Excel column and the full FASTA-formatted sequences were retrieved from the respective databases. They were also converted into a tab-delimited format and incorporated into the Excel table. Delimiters (\$, \$, #) were placed into additional columns for re-formatting purposes. For addition of the locus tags, the same set of sequences was retrieved in GenBank format, reformatted as above and copied into a separate work sheet. The column with the locus tags or gene designations was copied into the main table as appropriate.

The set of 4795 hit sequences was analyzed for transmembrane helices (TMH) using the TMHMM (one line per protein; <http://www.cbs.dtu.dk/services/TMHMM-2.0/>; Krogh et al., 2001) and SOSUI batch servers (<http://harrier.nagahama-i-bio.ac.jp/sosui/>; Hirokawa et al., 1998). The results were reformatted and again copied to the main table (Table S1). Signal sequences were predicted using SignalP (<http://www.cbs.dtu.dk/services/SignalP/>, model for Gram-negative bacteria; Petersen et al., 2011) and TatP (<http://www.cbs.dtu.dk/services/TatP/>; Bendtsen et al., 2005) for GSP and twin-arginine protein translocation (TAT) signal peptides, respectively. Proteins were also analyzed using OCTOPUS in cases of manually identified Cytc candidates with no result in the N-terminal TMH prediction. Sequences with three or more CxxCH motifs were defined as multiheme Cytc (MCC) unless shown not to be—by a high similarity to known non-cytochrome proteins in BLASTP searches (e.g., RecJ homologs). Additionally, various known Cytc and MCCs were used to query the Archaea subsection of the GenBank protein database. Sequences with two or one CxxCH motif were

TABLE 1 | Statistics of cytochrome c prediction in Archaea.

Total No. of archaeal proteins	888,023
Uniprot non-redundant proteins July 2014	816,158
Genbank additional archaeal proteins	71,865
No. of defined species/genomes strains	312
Total hits with CxxCH search	4795
No. of duplicated sequence hits	563
No. of unique sequences among duplicates	222
Multiheme cytochrome c candidate ($No_{CxxCH} \geq 3$ per sequence)	179
Same with N-term.TMH and/or predicted signal seq.	159
No. of predicted multiheme cytochromes c	167
No. of species/strains (cultured or uncultured) with multiheme cytochromes c	29
False positives (e.g., RecJ; 3x CxxCH each)	12
No. of proteins with 2 CxxCH/sequence	206
Same with N-term.TMH and/or predicted signal/TAT seq.	24
No. of predicted diheme cytochromes c	28
No. of species (cultured or uncultured) with diheme cytochromes c	20
False positives	178
No. of proteins with 1 CxxCH/sequence	4410
Same with N-term.TMH and/or predicted signal seq.	157
No. of predicted monoheme cytochromes c candidates	64
No. of species (cultured or uncultured) with monoheme cytochromes c	39
False positives	4347
Total No. of proteins subjected to structure prediction	1754
No. of cytochrome c candidates clustered for detailed analysis	350
No. of predicted archaeal cytochrome c proteins	258
No. of sequence similarity clusters	54
No. of predicted archaeal Cytc in species with 3 or more ccm genes ^a	241
No. of species/environmental samples	47
No. of predicted archaeal Cytc in species with 0–1 ccm genes ^b	17
No. of species	17

^aSee **Figure 2a**;^bSee **Figure 2b**.

considered Cytc candidates if they contained an N-terminal TMH and/or a signal sequence. Candidates were subjected to three-dimensional modeling using the batch processing mode of the Phyre² server (<http://www.sbg.bio.ic.ac.uk/phyre2/html/page.cgi?id=index>; Kelley and Sternberg, 2009). The results were purged from non-significant models (i.e., low confidence and/or alignment coverage percentage) and significant hits were used to evaluate the previously defined Cytc candidate clusters for completeness and correct identification. The *I. hospitalis* Cytc were also modeled using the I-Tasser server with omission of the respective signal sequences (<http://zhanglab.ccmb.med.umich.edu/I-TASSER/>; Roy et al., 2010). The resulting IgnI_0759

model was further adjusted by taking the predicted heme ligand out of the I-Tasser results files. The pdb coordinates including the heme were imported into UCSF Chimera (Pettersen et al., 2004) and the heme position was adjusted manually in order to build the thioether bonds between the heme and the two cysteine side chains followed by energy minimization. In the next round, a bond between the heme iron and the Ne atom of the proximal ligand His₃₂ was created and the energy minimization step repeated. The figure was prepared in Pymol (Delano, 2002).

The set of 4795 primary hit sequences was converted into a BLAST database using the standalone BLAST+ program downloaded from NCBI (http://blast.ncbi.nlm.nih.gov/Blast.cgi?CMD=Web&PAGE_TYPE=BlastDocs&DOC_TYPE=Download). Cytc candidates were compared against this database in order to find missing homologs and to identify clusters of mutually similar Cytc candidates. Clusters were aligned separately (**Supplementary Alignment File Archaea_Cytc.zip**). The multiheme cytochromes identified by Sharma et al. (2010) were also downloaded in FASTA format, converted into a separate BLAST database; they were used for the determination of cluster similarity and to relate clusters defined in this study to those from Sharma et al. (2010; **Table S1**). The primary hit sequences were also compared using BLASTP against the conserved domain database (CDD) installed locally.

Our methods differed from previous computational studies presented by Bertini et al. (2006) and Sharma et al. (2010). Both used HMMs (Sharma et al. for diheme and multiheme Cytc prediction only) and both used comparison against the protein family database for curation (PFAM; <http://pfam.xfam.org/>). Many of the cytochromes predicted here are not even clustered in PFAM or in NCBI's CDD for lack of 3D structures and/or biochemical description (**Table S1**, CDD search) so that we used clustering combined with structure prediction in order to identify Cytc folds in proteins. The main advantages of the methods used here are simplicity and no need for specialized software. They can be repeated from almost any standard PC or Mac using internet-available tools and free software (except for Microsoft Office products). The likewise freely accessible structure prediction part helped in assessing the previous conclusions.

The search for Cytc biosynthesis proteins was performed essentially as described (Allen et al., 2006). For system I, the CcmB, CcmC, CcmE, CcmF proteins from *Methanosarcina acetivorans*, *A. pernix*, *Haloarcula marismortui* (for GI numbers, see Allen et al., 2006), and *E. coli* were used in BLASTP searches against the archaeal proteins. BLASTP searches were repeated with archaeal hit sequences because sequence similarities were often low between unrelated Archaea. The *Leptospira interrogans* CcmH (GI:45656703) was used in addition, as homologs had so far not previously been found in Archaea (Allen et al., 2006). For System II, *Wolinella succinogenes* ResB was used (GI:34484157); for System III, the two heme lyases from *Saccharomyces cerevisiae* were used and for System IV the *Chlamydomonas reinhardtii* CCB1-4 proteins (de Vitry, 2011).

Electron Microscopy

For electron microscopy analysis, fresh *I. hospitalis* cells were cultivated, high-pressure frozen and freeze-substituted in 95%

acetone, 0.5% glutaraldehyde, 0.5% uranyl acetate, and 5% water as described (Rachel et al., 2010; Naß et al., 2014). After freeze-substitution fixation, samples were embedded in Epon. For localization of proteins on ultrathin sections, the primary antiserum directed against IgNI_0955 was used without further purification. For detection, secondary antibodies coupled to ultra-small gold particles were made visible by silver enhancement. Images were recorded as described (Naß et al., 2014). The *Kuenenia stuttgartiensis* cell had also been high-pressure frozen, freeze-substituted in acetone containing 2% OsO₄, 0.2% uranyl acetate, and 1% water and Epon-embedded as described (Wu et al., 2012). *Candidatus “Altiaarchaeum hamiconexum”* cells were sampled and prepared for electron microscopy as described elsewhere (Perras et al., 2014; Probst et al., 2014a).

Results

Prediction of Cytochromes c and their Maturation Proteins in Archaea

Motif and similarity searches and homology modeling were applied to the prediction of Cytc and their distribution in Archaea. 4795 archaeal proteins (**Table 1**) were found to contain at least one CxxCH amino acid pattern (**Table S1**). One hundred and seventy nine proteins contained at least three CxxCH motifs (defined here as MCCs), among those, 159 had a recognizable signal sequence and/or a predicted transmembrane helix (TMH) at their N-termini (**Table 1**). 12 sequences with three CxxCH motifs each were identified with BLASTP searches as RecJ exonuclease homologs and were considered as false positives. RecJ family proteins with 1–3 CxxCH motifs were among the most common random hits in the motif searches. The remaining 167 proteins from 29 archaeal species/strains were considered as multiheme cytochromes c (MCC; **Table 1** and **Table S1**).

The prediction of di- and mono-heme Cytc from the motif search resulted in a higher proportion of non-specific hits. Twenty eight out of 206 proteins from 20 species were identified as diheme Cytc candidates (**Tables S1, S2**). The majority of 4410 proteins with a single CxxCH motif (**Table 1**) were random hits with no recognizable similarity to Cytc or any feature suggestive of them being one. Among the 229 proteins with an N-terminal TMH and/or signal sequence, only those were considered as Cytc candidates if they were either similar to known Cytc sequences (e.g., cluster 30, homologs of the *A. pernix bc₁* complex), or if the CxxCH motif was conserved in a significant percentage of the homologs found in BLAST searches, and if the proteins were not *bona fide* members of other known protein families. Thioredoxin family proteins (including protein disulfide isomerases) were frequently occurring false positives with an N-terminal TMH; subunits of RNA and DNA polymerases, molybdopterin biosynthesis proteins, endonucleases, Zn²⁺-binding domains, and iron-sulfur proteins were among the most frequent false positives without a TMH.

One thousand seven hundred and fifty four proteins annotated as “hypotheticals” were subjected to batch structure prediction. The fold recognition often gave necessary hints for

the decision whether a protein or a cluster represents Cytc. No further Cytc candidates were spotted in this subset of the data. After reducing the score to 154 monoheme Cytc candidates falling into 30 similarity clusters (**Tables S1, S2**), 3D structure prediction was performed showing that 9 clusters all gave $\geq 96\%$ confidence predictions with various Cytc, the prediction results of cluster 47 were considered of intermediate quality (90% confidence). This and cluster 38 were included in the Cytc group. Seventeen sequence clusters were excluded from the Cytc group mostly because they gave significant modeling results with known non-Cytc proteins.

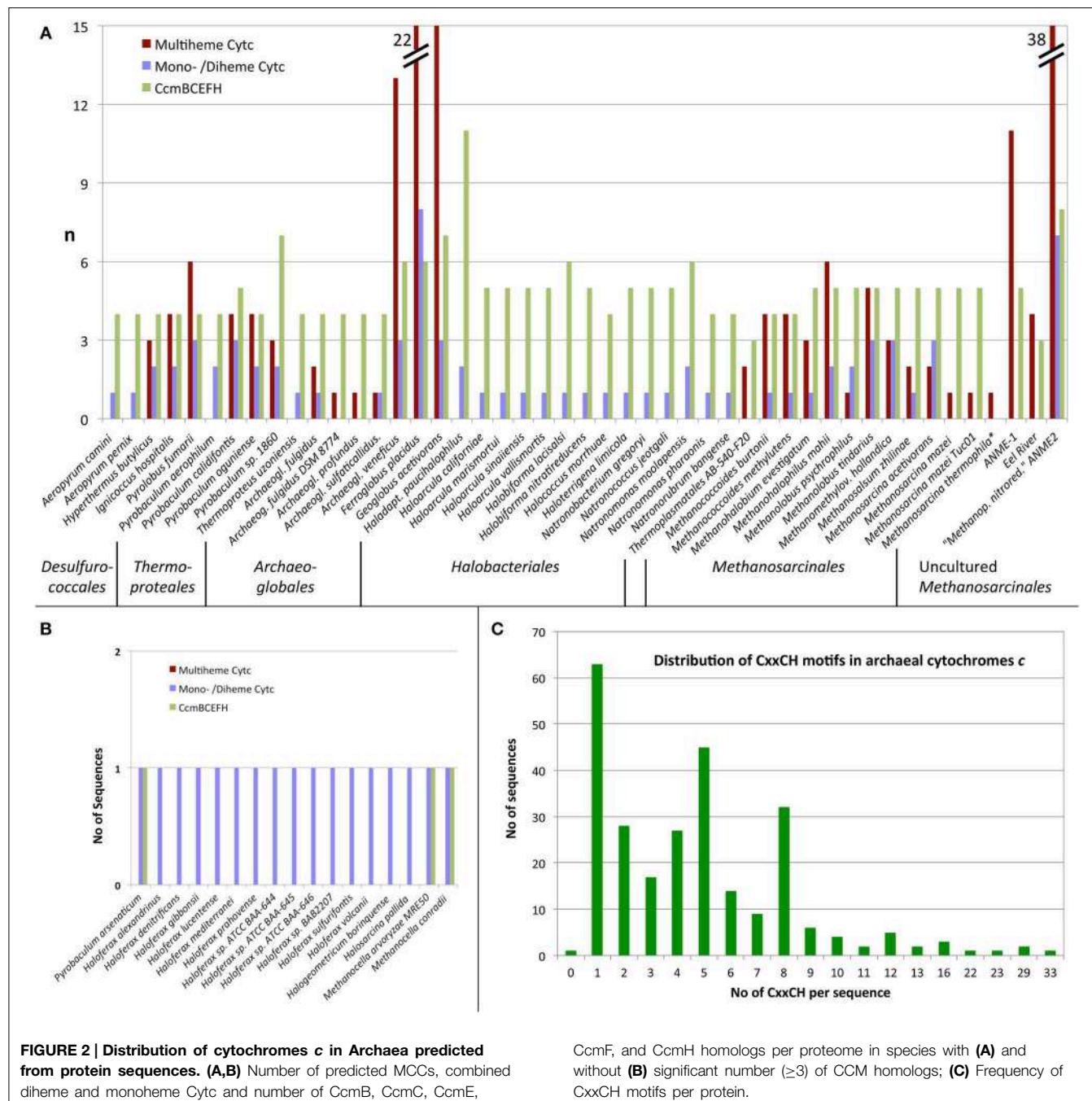
Multiheme Cytochromes c in Archaea

With one exception (**Figure 2**), the presence of MCCs-encoding genes was restricted to four of the major archaeal orders: the *Desulfurococcales*, the genus *Pyrobaculum* within the order of the *Thermoproteales* (both *Crenarchaeota*), the *Archaeoglobales*, and the *Methanosaecinales* including the methane-oxidizing environmental candidate species of the ANME-1 and ANME-2 groups (**Figures 2, 3**). The highest numbers of predicted MCC were found encoded in those species known or suspected to thrive anaerobically by iron respiration like *F. placidus* and in the uncultured methane-oxidizing Archaea of the ANME-1 and ANME-2 groups. The maximal number of CxxCH motifs in a single sequence was 33 in a large protein from the euryarchaeote *Ferroglobus placidus* (**Figure 2**).

The predicted MCCs were grouped in 34 clusters according to sequence similarity (**Tables S1, S2**, multiple alignments in the compressed supplemental sequence file Archaea_Cytc.zip). Some of the archaeal MCCs belong to well-known families like the hydroxylamine oxidoreductases (sequence cluster No. 4; 11 hits), octaheme tetrathionate reductases (cluster No. 5; 7 hits), or the periplasmic nitrite reductases (No. 69; 3 hits). In contrast, the protein function of most of the MCCs from Archaea is not known; many do not even have bacterial counterparts (e.g., clusters 1, 2, 11, 12 etc.; **Table S2**). Sometimes, structure prediction of MCC candidate proteins gave high-confidence (100%), full-length predictions. For example, protein models of cluster 1 matched with *Thioalkalivibrio nitratireducens* octaheme nitrite reductase (PDB accession 3f29) despite undetectably low sequence similarity, so that their function might nevertheless be inferred. Proteins of cluster 2 matched structurally octaheme tetrathionate reductases (PDB 1sp3; cluster 5). Other clusters gave more ambiguous results, which must be handled with care (**Table S2**), especially, when the number of CxxCH motifs in models and templates differed (e.g., clusters 8 and 9; not shown).

Di- and Mono-heme Cytochromes c

Among the predicted diheme Cytc (seven similarity clusters; **Table S2** and (Supplementary File Archaea_Cytc.zip) were peroxidases of the MauG type (cluster 29), thiosulfate dehydrogenases (TDH; cluster 28), a *bc₁* complex homolog from hyperthermophile *Pyrolobus fumarii* (cluster 30; the homologs from three other *Desulfurococcales* species have only one CxxCH motif including the biochemically characterized APE_1719 from *A. pernix*; Kabashima and Sakamoto, 2011),



and Split-Soret Cytc (cluster 21), the latter with predicted Twin-Arginine signal peptides. Six of these proteins were from *Archaeoglobales*, two from the crenarchaeote *Pyrolobus fumarii*, four from *Methanosarcinales* while the remaining 16 proteins were from various haloarchaea, which do not harbor MCCs as far as we know (Figure 2).

Structure prediction of the MauG peroxidases, the *bc*₁ complex homologs and the Split-Soret Cytc were consistent with the templates and they covered $\geq 70\%$ of the respective proteins with 100% confidence (Table S2). More interesting was

the case of the TDH homologs (cluster 28): modeling suggested structural similarity to SoxA proteins, which catalyze, together with SoxX, the oxidative transfer of thiosulfate to a cysteine side chain of SoxYZ. Sequence similarity between these two sulfur cycle enzymes is low but modeling showed structural similarity. These archaeal TDH homologs are encoded in genomes of five haloarchaeal species in operon-like arrangements with genes for CCM proteins.

Sharma et al. (2010) had predicted MCCs (with 2 hemes/protein or more instead of at least 3 hemes used here) in 8

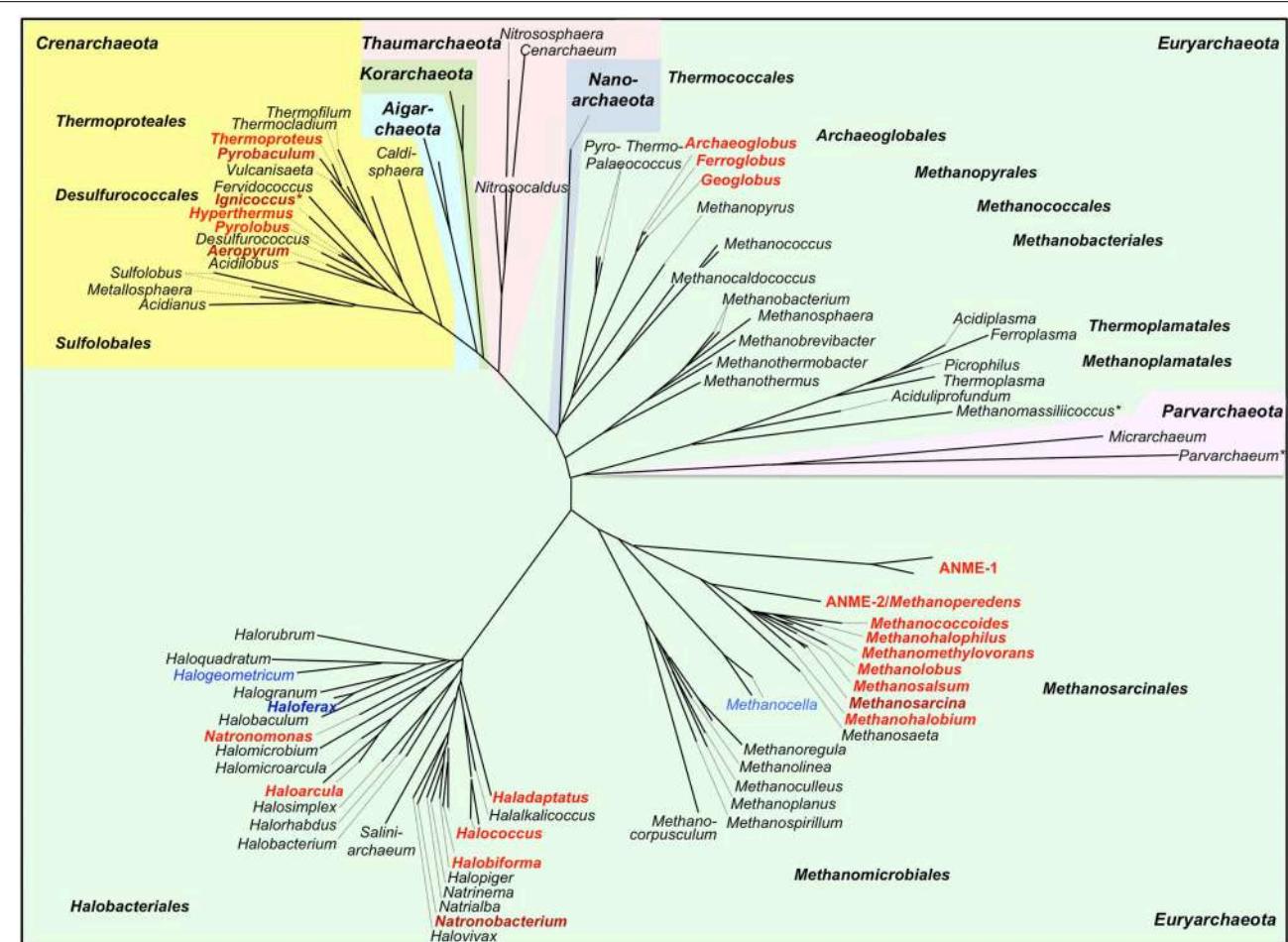


FIGURE 3 | Phylogenetic 16S rDNA dendrogram of the Archaea and distribution of predicted cytochrome c genes. The dendrogram was made from a 16S alignment both calculated with MAFFT (Katoh and Standley, 2014; <http://mafft.cbrc.jp/alignment/server/>). Dark red/dark blue, Cytc biochemically found; red, Cytc-encoding and ccm genes found (Figure 2A); blue, Cytc candidate genes but no or single ccm genes found

(Figure 2B); *Archaea with double membranes. Note: This dendrogram was created to depict the distribution of Cytc in Archaea and is based on 16S sequences only. The branching order of archaeal phyla does not correlate with more advanced dendograms based on concatenated protein sets. For a discussion of large-scale archaeal phylogeny, see for example Guy and Ettema (2011); Forterre (2013), and Petitjean et al. (2014).

out of 47 then-available archaeal genome sequences. We found all of those with the methods used in this study, however our interpretation was sometimes different. For example, they had identified *Methanospirillum hungatei* Mhun_1396 and its paralog Mhun_1882 as putative diheme Cytc. The proteins are highly conserved in methanogens but the CxxCH motifs are not so that we disregarded these two candidates. We also identified many previously unrecognized MCCs annotated as hypotheticals in genome sequences.

Sixty four proteins were assigned as monoheme Cytc candidates from 12 sequence clusters (Table 1 and Table S1). The modeling approach gave results with templates like cytochrome *c*₍₂₎, cytochrome P460, SoxX, and Cytc subunits of NO reductase (NorC) or ethylbenzene dehydrogenase (Table S2). A special case is the nitrite reductase subunit Pars_0592 from *Pyrobaculum arsenaticum*, which was identified with BLASTP searches and which is similar to its heme-*c* containing homologs (68 and

52% identity to the two *P. aerophilum* proteins PAE3598 and PAE1347, respectively) but which has a tyrosine residue instead of the first cysteine in the classical CxxCH motif. We suspect that there might be single or no covalent heme ligation in an otherwise functional protein.

Cytochrome c Maturation Proteins

Cytochromes *c* require maturation by heme ligases and, in most cases, transport proteins for the transfer of the heme moiety across the membrane to the electrochemically positive side. Cytochrome *c* maturation system I (CCM) originally described from *E. coli* is one of the two most common and the most complex CCM machinery of five known systems. The search for CCM proteins encoded in archaeal genomes was mainly done with sequence comparisons using BLAST and the CcmB, C, E, F, and H proteins as described by Allen et al. (2006) and in the Materials and Methods Section. The *I. hospitalis* genome encodes

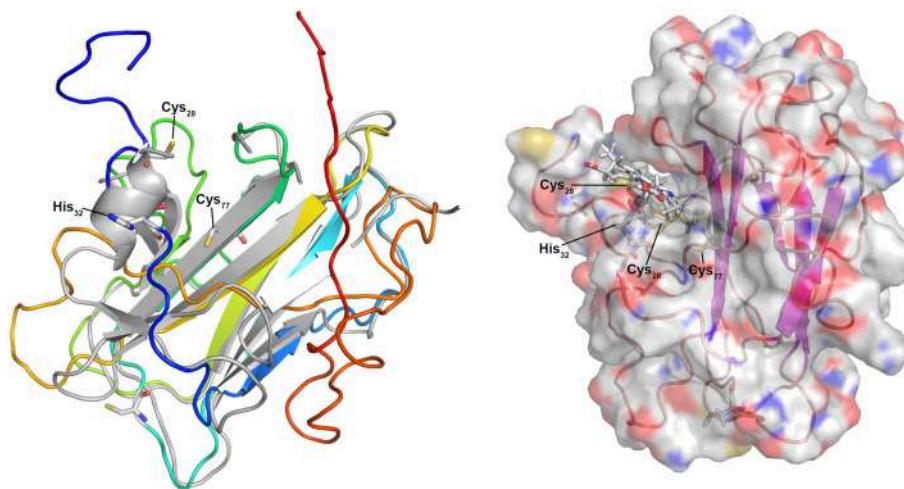


FIGURE 4 | 3D model of the predicted monoheme cytochrome c Igni_0759 from *I. hospitalis*. Left panel: 3D model created using the I-Tasser server (Roy et al., 2010; rainbow color, N-terminus blue) and superimposed over the modeling template

(heparin-binding reelin-N domain of f-spondin; PDB accession 3COO; Tan et al., 2008); **Right panel:** Igni_0759 3D model with heme c connected to Cys₂₈ and Cys₃₁ side chains with the iron connected to His₃₂.

four proteins, CcmB, CcmC, CcmE, and CcmF indicative of the presence of the entire CCM system. CcmH homologs were solely found in *Ferroglobus placidus*, while the remaining four proteins had homologs in 45 archaeal species, in which Cytc proteins were also predicted (Figure 2). Proteins of cytochrome maturation systems II–V were not identified in Archaea. It can be concluded from these results that the Cytc apoproteins are transferred at least over one membrane. Seventeen monoheme Cytc were predicted in species with either none (*Haloflexax spp.*, *Halogeometricum borinquense*, *Halosarcina pallida*, cluster 50) or only one maturation protein (*Methanocella spp.* *Pyrobaculum arsenaticum*) encoded in the genomes (Figure 2).

Cytochromes c in *Ignicoccus Hospitalis*

We had previously reported on the purification of three multiheme cytochromes c (MCC) from the hyperthermophilic archaeon *I. hospitalis* (Naß et al., 2014). We had also reported that one of those cytochromes was a membrane-bound MCC with four CxxCH motifs (locus tag Igni_0530) and that two octaheme MCCs were present both in the soluble and the membrane fractions (Igni_1359 and Igni_0955). We had further predicted an octaheme tetrathionate reductase-like protein (Igni_1130) and two so far hypothetical monoheme cytochromes c in the *I. hospitalis* proteome (Igni_0579 and Igni_1052; cluster 38). Here, we wanted to investigate in more detail whether the structure prediction used in Cytc identification in Archaea could substantiate this claim. We also extended structure prediction to the MCCs, again with the scope of extending the method more generally.

Igni_0579 and Igni_1052 are similar; Igni_1052 however has a second predicted TMH at its C-terminus not present in Igni_0759. Homologs occur in the related crenarchaeota *Pyrolobus fumarii* and *Hyperthermus butylicus*, both with a C-terminal TMH. The modeling servers (Phyre² and I-Tasser)

both used eukaryal spondin as the folding template (a non-heme protein, Tan et al., 2008) with high statistical confidence (100%). The models left a cleft in the molecule sufficient for heme accommodation with the cysteine side chains positioned at the top of the cleft (Figure 4), thus pointing to a space where the heme might be positioned. In further modeling steps, the heme moiety was added to the Igni_0759 model PDB coordinate file and connected to the side chains of Cys₂₈ and Cys₃₁. After energy minimization, the iron atom was connected to His₃₂ as proximal ligand and the protein was again subjected to energy minimization resulting in the model depicted in Figure 4. A further step connecting the iron to the side chain of Cys₇₇ as putative distal ligand failed. His₁₁₁ is a second candidate for the distal ligand and it is conserved in the homologs (cluster38_Igni_0759.fasta in the Archaea_Cytc.zip file). It was located beneath the β-sandwich forming the main structural body of the model so that we cannot presently decide, which of these two is correct. In summary, the model is congruent with the hypothesis that these *I. hospitalis* proteins are Cytc and they show that 3D structure prediction could be a valuable tool for the identification of unknown proteins, at least when applied to suspected monoheme Cytc.

Structure prediction was more difficult for the MCCs although Igni_1359 and Igni_0955 gave high-confidence (100%) full-length models with the *Nitrosomonas europeae* HAO 3D structure as template (PDB accession 1FGJ) with up to 28% sequence identity (not shown). Likewise, Igni_1130 gave a well-predicted model with the *Shewanella oneidensis* OTR (3SP3; not shown). However, significant 3D models were also created when the three proteins were modeled with non-homologous MCC templates (e.g., Igni_1130 with the HAO template) regardless of sequence similarity. The MCCs seem to be folded into multiple pre-existing 3D structures because high numbers of heme-binding sites predefine the folding of the apoproteins, thereby

restricting the predictive capabilities of structure modeling of MCCs. In consequence, a function prediction of MCC is at best difficult when trying to model non-homologous MCCs of unknown function, while monoheme Cytc give more reliable results.

Discussion

We present here a study for the identification of Cytc and their maturation proteins encoded in archaeal genomes using a computational approach coupled to an extensive manual evaluation of the results. We show that Cytc are not a common property of the majority of Archaea to our current knowledge and that they are not distributed equally, being restricted to 5–6 of the major taxa (**Figure 3**). In most Bacteria, Cytc are bound to cytoplasmic membranes or located in the periplasm or—in Gram-positives—in the space containing peptidoglycan and teichoic acids outside the cytoplasmic membrane, which is discussed to be equivalent to the periplasm of Gram-negatives (Matias and Beveridge, 2005). This is different in the compartmentalized Bacteria and Archaea. In the following discussion we will focus on two main questions:

1. What can we learn from the results of our computational study and the present state of knowledge about the distribution of Cytc, physiological patterns, and pathways in different archaeal lineages and about the acquisition of the genes during evolution?
2. What can we learn and predict about the localization and maturation of Cytc in Archaea and especially in double-membraned microorganisms like *I. hospitalis*?

Cytochromes c in Archaea

Forty-seven archaeal species or consortia of uncultured microorganisms were found encoding both Cytc and CCM maturation proteins in their genomes while 17 other species harbor hypothetical single Cytc candidates with little evidence for maturation proteins (**Table 1**, **Figure 2**). They belong to only five different orders of Archaea with the exception of two proteins from a single-cell genome of a *Thermoplasmatales* species. Some of the archaeal Cytc have numerous homologs in Bacteria (e.g., clusters 3 and 4) while others are specific for Archaea (e.g., cluster 1–2).

There are differences in the distribution within Cytc-containing archaeal orders and even within single genera: The *Archaeoglobales* are the only order, in which all species sequenced so far contain Cytc genes (**Figures 2, 3**). In contrast, out of 17 genome-sequenced *Thermoproteales* species only *Thermoproteus uzonensis* and 4–5 of 7 *Pyrobaculum* spp. contain Cytc genes (**Figure 2, Table S1**). For example, *Pyrobaculum* sp. strain 1860 and *Pb. oguniense* grow by iron and nitrate respiration (Nunoura et al., 2003; Mardanov et al., 2012) and contain several monoheme Cytc and MCCs obviously involved in various electron transport chains. Two heme-stained proteins were observed in gel electrophoresis of cell extracts of *Pyrobaculum aerophilum* (Feinberg et al., 2008). The authors proposed that they are identical to Cytc subunits of a three-subunit *bc*₁ complex

(PAE1347-9) and of a two-subunit NirS-type *cd*₁ nitrite reductase (PAE3598). We also found both proteins in this study although the apparent molecular mass of the nitrate-induced band did not match the calculated mass of PAE3598 (20 kDa w/o signal peptide). Protein identification was not given so that ORF numbers of the two heme-stained proteins remain tentative.

The only biochemically purified three-subunit crenarchaeal *bc* complex came from the microaerophilic species *A. pernix* (Kabashima and Sakamoto, 2011). In contrast, cyt *bc* complexes are absent in aerobic *Sulfolobales*, which have an analogous cytochrome *ba* electron transport complex instead (Bandeiras et al., 2009). The Cytc subunit was the only one to be identified (Ape_1719.1). The adjacent gene encodes a subunit of a terminal oxidase, whereas the genes for cytochrome *b* and a Rieske protein are close by but not in the same predicted operon (APE_1724.1 and APE_1725.1). Homologs of Ape_1719.1 are present in *Pyrolobus fumarii* and *Hyperthermus butylicus* but none of the cytochrome *b* and the Rieske proteins. It can be concluded that *Pyrobaculum* spp., *Thermoproteus uzonensis*, and *Aeropyrum* spp. encode canonical *bc* complexes, whereas the homologous Cytc plays a different role in *Pyrolobus* and *Hyperthermus*, it might be part of an unidentified electron transport complex. The distribution pattern is similar in the remaining archaeal orders with Cytc. Some species of the *Methanosaecinales* and *Halobacteriales* encode single or multiple Cytc and the corresponding *ccm* genes but not the majority of either of them.

Cytochromes c, Anaerobic Respiration, and Ammonia Oxidation

An exceptionally high number of Cytc was found in the euryarchaeota *Ferroglobus placidus* (**Figure 2**) and *Ca. "Methanoperedens nitroreducens"*. *F. placidus* (and also the crenarchaeote *Pyrolobus fumarii*) grow by Fe²⁺ oxidation with nitrate or Fe³⁺ reduction with various organic and inorganic electron donors, whereas *Ca. "Mp. nitroreducens"* grows by anaerobic oxidation of methane with nitrate (Hafenbrädl et al., 1996; Anderson et al., 2011; Haroon et al., 2013). Several ANME Archaea however couple anaerobic methane oxidation to iron or manganese reduction (Beal et al., 2009; Wankel et al., 2012) and the diversity of Cytc in these Archaea was noted in the respective metagenome papers (Meyer-Dierks et al., 2010; Wang et al., 2014). Some of the large multiheme and multidomain proteins from *F. placidus* and *Ca. "Mp. nitroreducens"* (clusters 17 and 65) have 5–8 CxxCH motifs in their N- or C-terminal Cytc domains. Modeling the none-Cytc domains separately, those parts can be folded into chains of successive beta sandwich domains comparable to surface layer proteins (not shown). The results suggest that these proteins might form extracellular conductive structures or pili as in *Shewanella* or *Geobacter*. Here, periplasmic, outer-membrane, or pilus-bound Cytc transfer electrons to and from the cells (reviewed for example in Gorby et al., 2006; Richter et al., 2012; Boesen and Nielsen, 2013; Smith et al., 2015). This might provide a structural and biochemical basis of the metal ion-reducing and the presumed electron-conductive capabilities of the iron-metabolizing Archaea. In a recent study, many heme-stained bands were found SDS gels

of extracts of Fe³⁺-grown *F. placidus* cells. The number of bands and of transcripts of Cyt c genes differed depending on the solution state of the iron: there were more Cyt c proteins and corresponding transcripts in cells grown on solid compared to soluble Fe³⁺ species; in addition there were numerous type IV pili suggesting close attachment of the cells to the substrate and/or electrically conductive pili (Smith et al., 2015). By analogy, the sulfate reducer *Archaeoglobus veneficus* with a total of 16 Cyt c genes should also be able to grow by metal respiration (Figure 2). In summary, metal ion respiration seems to be a predominant motif for the presence of high numbers of Cyt c genes in archaeal genomes.

Bacterial sulfate reducers are typical sources of a large variety of Cyt c (reviewed for example in Romão et al., 2012) and this seems also true for the *Archaeoglobi* but not for sulfate-reducing crenarchaeota (e.g., *Caldivirga maquilensis*), since we did not find any Cyt c genes in the latter microorganisms. Besides sulfate respiration, Cyt c play important roles in oxidative and reductive pathways of microbial sulfur and nitrogen cycles such as denitrification, nitrate ammonification, thiosulfate oxidation, and anaerobic ammonium oxidation (Anammox; Kartal et al., 2011; Simon et al., 2011; Kappler and Maher, 2013; van Teeseling et al., 2013). Surprisingly, no Cyt c were found in *Thaumarchaeota*, which represent a large phylum of Archaea characterized by their involvement in the global N cycle. *Thaumarchaeota* are proposed to be among the most abundant ammonia oxidizers in marine and in terrestrial ecosystems (Offre et al., 2013; Monteiro et al., 2014; Stieglmeier et al., 2014) and they might be implicated in denitrification as well (Jung et al., 2014). It is therefore surprising that the *Thaumarchaeota* seem to be (so far) devoid of Cyt c suggesting that other proteins with comparable activities fill in the gap and that they use different catalytic metal sites.

Methanogenesis

Other *Methanosarcinales* species beside the ANME group contain Cyt c as it was already discovered in the 1980s (Kuhn et al., 1983; Jussofie and Gottschalk, 1986). Two different Cyt c were found spectroscopically in membrane fractions of methanol-grown *Methanosarcina mazei* Gö1 cells but the proteins were not purified or identified (Kamlage and Blaut, 1992). We found three monoheme and one multiheme Cyt c gene in the *Ms. mazei* Gö1 genome (Table S1) but their assignment to the proteins reported by Kamlage and Blaut (1992) is currently not possible. Similarly, a multiheme Cyt c was found to participate in electron transport of *Ms. thermophila* (Wang et al., 2011). In both cases, the Cyt c were oxidized upon heterodisulfide addition (CoM-S-S-CoB) to membrane fractions however their precise role in the redox chains is not known. *Methanosarcinales* species are characterized by their utilization of various C₁ compounds and many can disproportionate acetate for methanogenesis and energy conservation. Now, *Methanosarcinales* are the only phylogenetic branch of methanogenic euryarchaeota (among at least six others) with both *b* and *c*-type cytochromes but their presence does not seem to be a prerequisite for growth on these substrates. Chemiosmotic coupling during methanogenesis from H₂/CO₂ is the most probable reason for the observed

higher growth yields in methanogens with cytochromes like *Methanosarcina barkeri* compared to those without (Thauer et al., 2008; Wang et al., 2011). The heterodisulfide reductase from *Ms. barkeri* contains a cytochrome *b* subunit (Heiden et al., 1994; Kunkel et al., 1997). We did not find cytochromes *c* in *Ms. barkeri* in our study here so that the cytochrome *b* subunit alone seems to be responsible for the growth yield effect and there is no other indication that Cyt c are integral players in this process. Generally, we observed here that only a small fraction of the known *Methanosarcinales* species contains Cyt c suggesting a different role for these proteins in energy metabolism.

The Haloarchaea Paradox

Electron transport components from halophilic Archaea (*Halobacteriales*) were studied since the 1960s (Lanyi, 1968; Cheah, 1970). Later, Scharf et al. (1997) characterized a membrane-bound 2-subunit *bc* complex (14 and 18 kDa, respectively) and a soluble 75 kDa Cyt c. A single Cyt c candidate was identified in our computational analysis: a 453 aa cytochrome *c*₅₅₁ peroxidase (MauG, cluster 29) is encoded in the genome together with *ccm* genes as in several other haloarchaea and *Methanosarcina* species (Tables S1, S2). This could explain the 75-kDa soluble heme-stained protein (Scharf et al., 1997). In contrast, we could not identify candidates for the heme-c protein of the *bc* complex. The situation was similar for *Halobacterium salinarum* and *Haloferax volcanii*. In both species, Cyt c were either purified (14 kDa protein in *Hbt. salinarum*), and/or spectroscopically characterized combined with heme-stained SDS gels (Sreeramulu et al., 1998; Tanaka et al., 2002; Sreeramulu, 2003). Two small proteins were found encoded in the *Hfx. volcanii* genome with little mutual sequence similarity and each with homologs in the same 12–13 haloarchaeal species (cluster 35 and 36; Figure 2) not including *Hbt. salinarum*. None of these species contain CCM. Both clusters gave low-confidence structure prediction hits (Table S2) so that independent evidence would be necessary for the identification of the Cyt c component of the haloarchaeal *bc* complexes. This leads to the conclusion that they might not be found using similarity and/or pattern searches and that they use non-standard amino acid patterns and heme c linkage.

There were several other haloarchaeal species with well-recognized and correctly annotated Cyt c and *ccm* genes; cluster 28 comprising 368–485 aa proteins with a monoheme domain and the already mentioned cluster 29 (MauG-type peroxidases). The observation that some haloarchaea contain genes for cluster 28 and 29 Cyt c only—the latter occurring in some of the *Methanosarcinales* as well—and the lack of MCCs suggests late gene acquisition from bacterial sources by horizontal gene transfer (HGT) as suggested earlier (Nelson-Sathi et al., 2012). A similar mechanism can be concluded for the metal-metabolizing archaeal species and the *Methanosarcinales*. In conclusion, the overall pattern suggests several events of horizontal transfer from Bacteria to Archaea as proposed as a general model of archaeal gene acquisition (Nelson-Sathi et al., 2015). In addition, the occurrence of Cyt c genes seems to match physiological constraints rather than phylogenetic relationships.

Cytochromes c and Cell Morphology

The majority of Archaea with cytochromes *c*—predicted in this study or biochemically proven—display the “standard” archaeal cell architecture: a cytoplasmic membrane covered with a proteinaceous surface (S-) layer anchored in the membrane (König et al., 2007). S-layers are protein canopies anchored in the cytoplasmic membrane encompassing a “quasi-” or “pseudo-periplasmic space” (Baumeister and Lembcke, 1992; König et al., 2007; Klingl, 2014), which can accommodate membrane-bound and soluble proteins (Baumeister et al., 1989; Veith et al., 2009; Protze et al., 2011; Klingl, 2014). It is therefore to be expected that Cyt_c are located in this space between cytoplasmic membrane and protein canopy and that they are retained either by pores in the protein lattice or by C-terminal membrane anchors as seen in many of the Cyt_c candidates described here (Table S1). Similarly, maturation of Cyt_c should also take place in this environment.

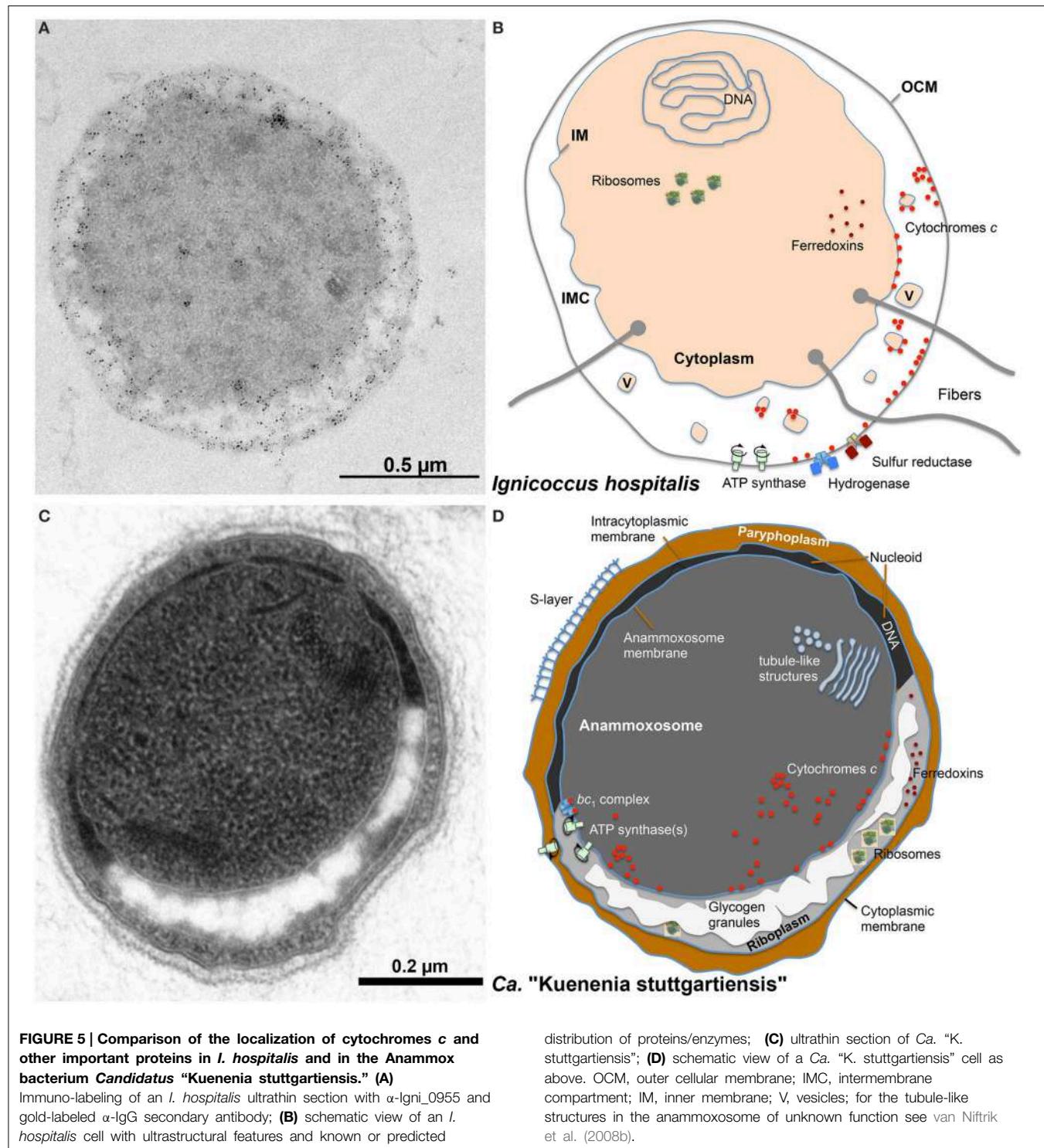
With their two membranes and the lack of an S-layer, *Ignicoccus* species are an exception to the typical archaeal cell architecture (Figures 1, 5). For Cyt_c, this encompasses the localization of the proteins, the location of the CCM machinery and last but not least the pathways of electron transport from the OCM to the inner compartment. Similar questions arise for the growing number of known double-membraned Archaea including the tiny *Parvarchaeota* of the ARMAN group (Comolli et al., 2009), the *Methanoplasmatales* (methanogens of the *Thermoplasmata* phylum; Dridi et al., 2012; Paul et al., 2012) and the uncultured SM1 euryarchaeota from a newly defined order *Candidatus “Altarchaeales”* (Figure 1; Probst et al., 2014a,c). The distribution of proteins between the compartments and electron transfer is also unknown in those species. *Ignicoccus* spp. however, are the only double-membraned Archaea with cytochrome *c*. Immuno-labeling had shown that the octaheme MCCs Igni_0955 and Igni_1359 are localized at both membranes (and eventually also at vesicles in the intermembrane compartment (IMC); Figure 5; Naß et al., 2014).

The organisms of the bacterial phylum *Planctomycetes* display ostensibly similar cell morphologies and the question is whether that is comparable to the double-membraned Archaea and whether we can make deductions for protein distribution and electron pathways from these bacteria. *Planctomycetes* species are known to have an inner and outer membrane encompassing a “paryphoplasm” in addition to a protein S-layer (Lindsay et al., 2001; van Teeseling et al., 2014). The paryphoplasm was defined as a structural description of “a unique, peripheral ribosome-free region of cytoplasm” in order to distinguish it from the “riboplasm,” the central compartment containing ribosomes and the nucleoid surrounded by an IM (Lindsay et al., 2001). There is also an ongoing discussion whether or not their membrane organization “is not different from, but an extension of, the “classical” Gram-negative bacterial membrane system” (Santarella-Mellwig et al., 2013; Sagulenko et al., 2014; Jeske et al., in press; van Teeseling et al., in press). Even more complex are Anammox bacteria like *Ca. “Kuenenia stuttgartiensis”*, also belonging to the *Planctomycetes* and also with an S-layer (Figure 5; van Teeseling et al., 2014). The cells have an additional cellular compartment, the anammoxosome within the riboplasm,

which contains the proteins required for anaerobic ammonium oxidation including numerous cytochromes *c* like hydroxylamine and hydrazine oxidoreductases. This compartment is the place of energy conversion; the anammoxosome membrane comprises a pH gradient across its ATP synthase-containing membrane with the positive (p) side inside (van Niftrik et al., 2008a, 2010; van der Star et al., 2010; Neumann et al., 2014). Therefore, it is reasonable to assume that maturation of the Cyt_c (using the system II) takes place in the anammoxosome and that the apoproteins are transported inside. The localization of Cyt_c is unknown in non-Anammox planctomycetes.

I. hospitalis differs in several aspects from the planctomycetes: it does not have an S-layer or a morphologically defined nucleoid and of course nothing equivalent to the anammoxosome. Also, the IMC is very lightly contrasted in electron microscopy pictures suggesting a low concentration of biomolecules. The same seems to be true for the *Methanoplasmatales* (Figure 1; Dridi et al., 2012). In contrast, the paryphoplasm of the planctomycetes usually is much darker in electron microscopy (“electron-dense”) than the IMC (Figure 5; Lindsay et al., 2001). And third, the *I. hospitalis* ATP synthase and a heterologous hydrogenase/sulfur reductase complex are localized in the OCM (Küper et al., 2010). From this, we have to assume that the P-side is outside of the OCM. Maturation of the Cyt_c at the OCM however would require a transfer of the apoprotein and the heme moiety across two membrane systems (Figure 5). The latter cannot be excluded, however the mature Cyt_c would have to go back in to reach the IM, where they were found as well (Figure 5). A more easy explanation would be to assume that the apoprotein is transferred co-translationally across the IM via the sec pathway and that maturation would occur prior to further transport. This conclusion however would imply that the maturation takes place in the IMC but at the negative side of the cytoplasmic membrane unless there is an additional proton gradient across the IM. None of that is at present resolved (Huber et al., 2012).

A different question is about the function of the Cyt_c in *I. hospitalis*. We have proposed that the membrane-bound tetraheme Cyt_c Igni_0530 might be part of the sulfur reductase, however this is still hypothetical (Naß et al., 2014). Likewise hypothetical is the hypothesis that the Cyt_c might act as electron relay from the OCM-bound hydrogenase to oxidoreductases in the cytoplasm. We had measured reduction of Igni_0955 (and to a lesser extent Igni_1359) by the native hydrogenase supporting this assumption but it will have to be confirmed independently. At present we would disregard ferredoxins as electron transfer proteins because there are no ferredoxins with twin arginine signal peptides encoded in the *I. hospitalis* genome, which would be required for membrane transport of iron-sulfur proteins. The same observation was made for ferredoxins of *Ca. “K. stuttgartiensis”*, which lead to the—tentative—placement of the ferredoxins in the inner compartment or in the riboplasm, respectively, in the schematic drawings of Figure 5. We also did not find quinones by solvent extraction (Naß et al., 2014). Therefore, the abundantly available Cyt_c are good candidates for electron transfer from the OCM to the inner compartment in *I. hospitalis*.



We can conclude about the comparison of *I. hospitalis* to the Anammox planctomycetales that the anammoxosomes of those bacteria are distinctly different structures and that the pathways of electron flow and the localization of Cyt c is fundamentally different. Unfortunately, we do not know the localization of the respiratory chain(s) in the non-anammox planctomycetes, but they seem to be a system better comparable to the situation in

I. hospitalis especially regarding Cyt c distribution and electron flow.

Acknowledgments

We thank Alexandra Perras and Christine Moissl-Eichinger (Medical University of Graz, Graz, Austria) for the samples

for preparing the EM picture of *Candidatus “Altiaarchaeum hamiconexum”* (**Figure 1**). TH, JF, and RR were supported by the Deutsche Forschungsgemeinschaft (DFG HU703/2-2). Special thanks are due to Felicitas Pfeifer (Darmstadt) for continuous support and discussion.

Supplementary Material

The Supplementary Material for this article can be found online at: <http://journal.frontiersin.org/article/10.3389/fmicb.2015.00439/abstract>

References

- Allen, J. W. (2011). Cytochrome *c* biogenesis in mitochondria–Systems III and V. *FEBS J.* 278, 4198–4216. doi: 10.1111/j.1742-4658.2011.08231.x
- Allen, J. W., Harvat, E. M., Stevens, J. M., and Ferguson, S. J. (2006). A variant System I for cytochrome *c* biogenesis in Archaea and some Bacteria has a novel CcmE and no CcmH. *FEBS Lett.* 580, 4827–4834. doi: 10.1016/j.febslet.2006.07.073
- Anderson, I., Risso, C., Holmes, D., Lucas, S., Copeland, A., Lapidus, A., et al. (2011). Complete genome sequence of *Ferroglobus placidus* AEDII12DO. *Stand. Genomic Sci.* 5, 50–60. doi: 10.4056/sigs.2225018
- Bandeiras, T. M., Refojo, P. N., Todorovic, S., Murgida, D. H., Hildebrandt, P., Bauer, C., et al. (2009). The cytochrome *ba* complex from the thermoacidophilic crenarchaeote *Acidianus ambivalens* is an analog of *bc*(1) complexes. *Biochim. Biophys. Acta* 1787, 37–45. doi: 10.1016/j.bbabi.2008.09.009
- Baumeister, W., and Lembecke, G. (1992). Structural features of archaeabacterial cell envelopes. *J. Bioenerg. Biomembr.* 24, 567–575. doi: 10.1007/BF00762349
- Baumeister, W., Wildhaber, W., and Phipps, B. M. (1989). Principles of organization in eubacterial and archaeabacterial surface proteins. *Can. J. Microbiol.* 35, 215–227. doi: 10.1139/m89-034
- Beal, E. J., House, C. H., and Orphan, V. J. (2009). Manganese- and iron-dependent marine methane oxidation. *Science* 325, 184–187. doi: 10.1126/science.1169984
- Bendtsen, J. D., Nielsen, H., Widdick, D., Palmer, T., and Brunak, S. (2005). Prediction of twin-arginine signal peptides. *BMC Bioinformatics* 6:167. doi: 10.1186/1471-2105-6-167
- Bertini, I., Cavallaro, G., and Rosato, A. (2006). Cytochrome *c*: occurrence and functions. *Chem. Rev.* 106, 90–115. doi: 10.1021/cr050241v
- Boesen, T., and Nielsen, L. P. (2013). Molecular dissection of bacterial nanowires. *MBio* 4, e00270–e00213. doi: 10.1128/mBio.00270-13
- Burghardt, T., Nather, D. J., Junglas, B., Huber, H., and Rachel, R. (2007). The dominating outer membrane protein of the hyperthermophilic Archaeum *Ignicoccus hospitalis*: a novel pore-forming complex. *Mol. Microbiol.* 63, 166–176. doi: 10.1111/j.1365-2958.2006.05509.x
- Cheah, K. S. (1970). Properties of the membrane-bound respiratory chain system of *Halobacterium salinarium*. *Biochim. Biophys. Acta* 216, 43–53. doi: 10.1016/0006-2728(70)90157-X
- Comolli, L. R., Baker, B. J., Downing, K. H., Siegerist, C. E., and Banfield, J. F. (2009). Three-dimensional analysis of the structure and ecology of a novel, ultra-small archaeon. *ISME J.* 3, 159–167. doi: 10.1038/ismej.2008.99
- Delano, W. L. (2002). *The PyMOL Molecular Graphics System*. San Carlos, CA: DeLano Scientific.
- de Vitry, C. (2011). Cytochrome *c* maturation system on the negative side of bioenergetic membranes: CCB or system IV. *FEBS J.* 278, 4189–4197. doi: 10.1111/j.1742-4658.2011.08373.x
- Dirmeier, R., Keller, M., Frey, G., Huber, H., and Stetter, K. O. (1998). Purification and properties of an extremely thermostable membrane-bound sulfur-reducing complex from the hyperthermophilic *Pyrodictium abyssi*. *Eur. J. Biochem.* 252, 486–491. doi: 10.1046/j.1432-1327.1998.2520486.x
- Dridi, B., Farreau, M. L., Ollivier, B., Raoult, D., and Drancourt, M. (2012). *Methanomassiliicoccus luminyensis* gen. nov., sp. nov., a methanogenic archaeon isolated from human faeces. *Int. J. Syst. Evol. Microbiol.* 62, 1902–1907. doi: 10.1099/ijss.0.033712-0
- Feinberg, L. F., and Holden, J. F. (2006). Characterization of dissimilatory Fe(III) versus NO₃- reduction in the hyperthermophilic archaeon *Pyrobaculum aerophilum*. *J. Bacteriol.* 188, 525–531. doi: 10.1128/JB.188.2.525-531.2006
- Feinberg, L. F., Srikanth, R., Vachet, R. W., and Holden, J. F. (2008). Constraints on anaerobic respiration in the hyperthermophilic Archaea *Pyrobaculum islandicum* and *Pyrobaculum aerophilum*. *Appl. Environ. Microbiol.* 74, 396–402. doi: 10.1128/AEM.02033-07
- Forterre, P. (2013). The common ancestor of archaea and eukarya was not an archaeon. *Archaea* 2013:372396. doi: 10.1155/2013/372396
- Gorby, Y. A., Yanina, S., McLean, J. S., Rosso, K. M., Moyles, D., Dohnalkova, A., et al. (2006). Electrically conductive bacterial nanowires produced by *Shewanella oneidensis* strain MR-1 and other microorganisms. *Proc. Natl. Acad. Sci. U.S.A.* 103, 11358–11363. doi: 10.1073/pnas.0604517103
- Guy, L., and Ettema, T. J. (2011). The archaeal ‘TACK’ superphylum and the origin of eukaryotes. *Trends Microbiol.* 19, 580–587. doi: 10.1016/j.tim.2011.09.002
- Hafenbradl, D., Keller, M., Dirmeier, R., Rachel, R., Rossnagel, P., Burggraf, S., et al. (1996). *Ferroglobus placidus* gen. nov., sp. nov., A novel hyperthermophilic archaeum that oxidizes Fe²⁺ at neutral pH under anoxic conditions. *Arch. Microbiol.* 166, 308–314. doi: 10.1007/s002030050388
- Haroon, M. F., Hu, S., Shi, Y., Imelfort, M., Keller, J., Hugenholtz, P., et al. (2013). Anaerobic oxidation of methane coupled to nitrate reduction in a novel archaeal lineage. *Nature* 500, 567–570. doi: 10.1038/nature12375
- Heiden, S., Hedderich, R., Setzke, E., and Thauer, R. K. (1994). Purification of a two-subunit cytochrome-*b*-containing heterodisulfide reductase from methanol-grown *Methanosarcina barkeri*. *Eur. J. Biochem.* 221, 855–861. doi: 10.1111/j.1432-1033.1994.tb18800.x
- Hirokawa, T., Boon-Chieng, S., and Mitaku, S. (1998). SOSUI: classification and secondary structure prediction system for membrane proteins. *Bioinformatics* 14, 378–379. doi: 10.1093/bioinformatics/14.4.378
- Huber, H., Küper, U., Daxer, S., and Rachel, R. (2012). The unusual cell biology of the hyperthermophilic Crenarchaeon *Ignicoccus hospitalis*. *Antonie Van Leeuwenhoek* 102, 203–219. doi: 10.1007/s10482-012-9748-5
- Jeske, O., Schuffler, M., Schumann, P., Schneider, A., Boedeker, C., Jogler, M., et al. (in press). Planctomycetes do possess a peptidoglycan cell wall. *Nat. Comm.*
- Jung, M. Y., Well, R., Min, D., Giesemann, A., Park, S. J., Kim, J. G., et al. (2014). Isotopic signatures of N₂O produced by ammonia-oxidizing archaea from soils. *ISME J.* 8, 1115–1125. doi: 10.1038/ismej.2013.205
- Jussofie, A., and Gottschalk, G. (1986). Further-studies on the distribution of cytochromes in Methanogenic bacteria. *FEMS Microbiol. Lett.* 37, 15–18. doi: 10.1111/j.1574-6968.1986.tb01758.x
- Kabashima, Y., and Sakamoto, J. (2011). Purification and biochemical properties of a cytochrome *bc* complex from the aerobic hyperthermophilic archaeon *Aeropyrum pernix*. *BMC Microbiol.* 11:52. doi: 10.1186/1471-2180-11-52
- Kamlage, B., and Blaut, M. (1992). Characterization of cytochromes from *Methanosarcina* strain GÖl and their involvement in electron transport during growth on methanol. *J. Bacteriol.* 174, 3921–3927.
- Kappler, U., and Maher, M. J. (2013). The bacterial SoxAX cytochromes. *Cell. Mol. Life Sci.* 70, 977–992. doi: 10.1007/s00018-012-1098-y
- Kartal, B., Maalke, W. J., de Almeida, N. M., Cirpus, I., Gloerich, J., Geerts, W., et al. (2011). Molecular mechanism of anaerobic ammonium oxidation. *Nature* 479, 127–130. doi: 10.1038/nature10453

- Katoh, K., and Standley, D. M. (2014). MAFFT: iterative refinement and additional methods. *Methods Mol. Biol.* 1079, 131–146. doi: 10.1007/978-1-62703-646-7_8
- Kelley, L. A., and Sternberg, M. J. (2009). Protein structure prediction on the web: a case study using the Phyre server. *Nat. Protoc.* 4, 363–371. doi: 10.1038/nprot.2009.2
- Kern, M., Klotz, M. G., and Simon, J. (2011). The *Wolinella succinogenes* mcc gene cluster encodes an unconventional respiratory sulphite reduction system. *Mol. Microbiol.* 82, 1515–1530. doi: 10.1111/j.1365-2958.2011.07906.x
- Klingl, A. (2014). S-layer and cytoplasmic membrane - exceptions from the typical archaeal cell wall with a focus on double membranes. *Front. Microbiol.* 5:624. doi: 10.3389/fmicb.2014.00624
- König, H., Rachel, R., and Claus, H. (2007). “Proteinaceous surface layers of Archaea: ultrastructure and biochemistry,” in *Archaea. Molecular and Cellular Biology*, ed R. Cavicchioli (Washington, DC: ASM Press), 315–340.
- Krogh, A., Larsson, B., Von Heijne, G., and Sonnhammer, E. L. (2001). Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J. Mol. Biol.* 305, 567–580. doi: 10.1006/jmbi.2000.4315
- Kuhn, W., Fiebig, K., Hippe, H., Mah, R. A., Husser, B. A., and Gottschalk, G. (1983). Distribution of cytochromes in Methanogenic bacteria. *FEMS Microbiol. Lett.* 20, 407–410. doi: 10.1016/0378-1097(83)90105-2
- Kunkel, A., Vaupel, M., Heim, S., Thauer, R. K., and Hedderich, R. (1997). Heterodisulfide reductase from methanol-grown cells of *Methanosarcina barkeri* is not a flavoenzyme. *Eur. J. Biochem.* 244, 226–234. doi: 10.1111/j.1432-1033.1997.00226.x
- Küper, U., Meyer, C., Müller, V., Rachel, R., and Huber, H. (2010). Energized outer membrane and spatial separation of metabolic processes in the hyperthermophilic Archaeon *Ignicoccus hospitalis*. *Proc. Natl. Acad. Sci. U.S.A.* 107, 3152–3156. doi: 10.1073/pnas.0911711107
- Lanyi, J. K. (1968). Studies of the electron transport chain of extremely halophilic bacteria. I. Spectrophotometric identification of the cytochromes of *Halobacterium cutirubrum*. *Arch. Biochem. Biophys.* 128, 716–724. doi: 10.1016/0003-9861(68)90080-5
- Lindsay, M. R., Webb, R. I., Strous, M., Jetten, M. S., Butler, M. K., Forde, R. J., et al. (2001). Cell compartmentalisation in planctomycetes: novel types of structural organisation for the bacterial cell. *Arch. Microbiol.* 175, 413–429. doi: 10.1007/s002030100280
- Lockwood, C. W., Clarke, T. A., Butt, J. N., Hemmings, A. M., and Richardson, D. J. (2011). Characterization of the active site and calcium binding in cytochrome c nitrite reductases. *Biochem. Soc. Trans.* 39, 1871–1875. doi: 10.1042/BST20110731
- Mardanov, A. V., Gumerov, V. M., Slobodkina, G. B., Beletsky, A. V., Bonch-Osmolovskaya, E. A., Ravin, N. V., et al. (2012). Complete genome sequence of strain 1860, a crenarchaeon of the genus *Pyrobaculum* able to grow with various electron acceptors. *J. Bacteriol.* 194, 727–728. doi: 10.1128/JB.06465-11
- Matias, V. R., and Beveridge, T. J. (2005). Cryo-electron microscopy reveals native polymeric cell wall structure in *Bacillus subtilis* 168 and the existence of a periplasmic space. *Mol. Microbiol.* 56, 240–251. doi: 10.1111/j.1365-2958.2005.04535.x
- Mayer, F., Küper, U., Meyer, C., Daxer, S., Müller, V., Rachel, R., et al. (2012). AMP-forming acetyl coenzyme A synthetase in the outermost membrane of the hyperthermophilic crenarchaeon *Ignicoccus hospitalis*. *J. Bacteriol.* 194, 1572–1581. doi: 10.1128/JB.06130-11
- Mayer, F., and Müller, V. (2014). Adaptations of anaerobic archaea to life under extreme energy limitation. *FEMS Microbiol. Rev.* 38, 449–472. doi: 10.1111/1574-6976.12043
- Meyer, C., Heimerl, T., Wirth, R., Klingl, A., and Rachel, R. (2014). The Iho670 fibers of *Ignicoccus hospitalis* are anchored in the cell by a spherical structure located beneath the inner membrane. *J. Bacteriol.* 196, 3807–3815. doi: 10.1128/JB.01861-14
- Meyerdierks, A., Kube, M., Kostadinov, I., Teeling, H., Glockner, F. O., Reinhardt, R., et al. (2010). Metagenome and mRNA expression analyses of anaerobic methanotrophic archaea of the ANME-1 group. *Environ. Microbiol.* 12, 422–439. doi: 10.1111/j.1462-2920.2009.02083.x
- Moissl, C., Rachel, R., Briegel, A., Engelhardt, H., and Huber, R. (2005). The unique structure of archaeal ‘hami’ highly complex cell appendages with nano-grappling hooks. *Mol. Microbiol.* 56, 361–370. doi: 10.1111/j.1365-2958.2005.04294.x
- Monteiro, M., Seneca, J., and Magalhaes, C. (2014). The history of aerobic ammonia oxidizers: from the first discoveries to today. *J. Microbiol.* 52, 537–547. doi: 10.1007/s12275-014-4114-0
- Mooser, D., Maneg, O., Macmillan, F., Malatesta, F., Soulimane, T., and Ludwig, B. (2006). The menaquinol-oxidizing cytochrome bc complex from *Thermus thermophilus*: protein domains and subunits. *Biochim. Biophys. Acta* 1757, 1084–1095. doi: 10.1016/j.bbabi.2006.05.033
- Naß, B., Pöll, U., Langer, J. D., Kreuter, L., Küper, U., Flechsler, J., et al. (2014). Three multihaem cytochromes c from the hyperthermophilic archaeon *Ignicoccus hospitalis*: purification, properties and localization. *Microbiology* 160, 1278–1289. doi: 10.1099/mic.0.077792-0
- Nelson-Sathi, S., Dagan, T., Landan, G., Janssen, A., Steel, M., Mcinerney, J. O., et al. (2012). Acquisition of 1000 eubacterial genes physiologically transformed a methanogen at the origin of Haloarchaea. *Proc. Natl. Acad. Sci. U.S.A.* 109, 20537–20542. doi: 10.1073/pnas.1209119109
- Nelson-Sathi, S., Sousa, F. L., Roettger, M., Lozada-Chavez, N., Thiergart, T., Janssen, A., et al. (2015). Origins of major archaeal clades correspond to gene acquisitions from bacteria. *Nature* 517, 77–80. doi: 10.1038/nature13805
- Neumann, S., Wessels, H. J., Rijpstra, W. I., Sinnighe Damste, J. S., Kartal, B., Jetten, M. S., et al. (2014). Isolation and characterization of a prokaryotic cell organelle from the anammox bacterium *Kuenenia stuttgartiensis*. *Mol. Microbiol.* 94, 794–802. doi: 10.1111/mmi.12816
- Noor, M. R., and Soulimane, T. (2013). Structure of *caa*(3) cytochrome c oxidase—a nature-made enzyme-substrate complex. *Biol. Chem.* 394, 579–591. doi: 10.1515/hzs-2012-0343
- Nounoura, T., Sako, Y., Wakagi, T., and Uchida, A. (2003). Regulation of the aerobic respiratory chain in the facultatively aerobic and hyperthermophilic archaeon *Pyrobaculum oguniense*. *Microbiology* 149, 673–688. doi: 10.1099/mic.0.26000-0
- Offre, P., Spang, A., and Schleper, C. (2013). Archaea in biogeochemical cycles. *Annu. Rev. Microbiol.* 67, 437–457. doi: 10.1146/annurev-micro-092412-155614
- Paul, K., Nonoh, J. O., Mikulski, L., and Brune, A. (2012). “*Methanoplasmatales*” *Thermoplasmatales*-related archaea in termite guts and other environments, are the seventh order of methanogens. *Appl. Environ. Microbiol.* 78, 8245–8253. doi: 10.1128/AEM.02193-12
- Perras, A. K., Wanner, G., Klingl, A., Mora, M., Auerbach, A. K., Heinz, V., et al. (2014). Grappling archaea: ultrastructural analyses of an uncultivated, cold-loving archaeon, and its biofilm. *Front. Microbiol.* 5:397. doi: 10.3389/fmicb.2014.00397
- Petersen, T. N., Brunak, S., von Heijne, G., and Nielsen, H. (2011). SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nat. Methods* 8, 785–786. doi: 10.1038/nmeth.1701
- Petitjean, C., Deschamps, P., Lopez-Garcia, P., and Moreira, D. (2014). Rooting the domain archaea by phylogenomic analysis supports the foundation of the new kingdom *Proteoarchaeota*. *Genome Biol. Evol.* 7, 191–204. doi: 10.1093/gbe/evu274
- Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C., et al. (2004). UCSF Chimera—a visualization system for exploratory research and analysis. *J. Comput. Chem.* 25, 1605–1612. doi: 10.1002/jcc.20084
- Pihl, T. D., Black, L. K., Schulman, B. A., and Maier, R. J. (1992). Hydrogen-oxidizing electron transport components in the hyperthermophilic archaeabacterium *Pyrodictium brockii*. *J. Bacteriol.* 174, 137–143.
- Probst, A. J., Birarda, G., Holman, H. Y., Desantis, T. Z., Wanner, G., Andersen, G. L., et al. (2014a). Coupling genetic and chemical microbiome profiling reveals heterogeneity of archaeome and bacteriome in subsurface biofilms that are dominated by the same archaeal species. *PLoS ONE* 9:e99801. doi: 10.1371/journal.pone.0099801
- Probst, A. J., Weinmaier, T., Raymann, K., Perras, A., Emerson, J. B., Rattei, T., et al. (2014c). Biology of a widespread uncultivated archaeon that contributes to carbon fixation in the subsurface. *Nat. Commun.* 5, 5497. doi: 10.1038/ncomms6497
- Protze, J., Müller, F., Lauber, K., Nass, B., Mentele, R., Lottspeich, F., et al. (2011). An extracellular Tetrathionate Hydrolase from the Thermoacidophilic Archaeon *Acidianus ambivalens* with an activity optimum at pH 1. *Front. Microbiol.* 2:68. doi: 10.3389/fmicb.2011.00068

- Rachel, R., Meyer, C., Klingl, A., Gurster, S., Heimerl, T., Wasserburger, N., et al. (2010). Analysis of the ultrastructure of archaea by electron microscopy. *Methods Cell Biol.* 96, 47–69. doi: 10.1016/S0091-679X(10)96003-2
- Richter, K., Schicklberger, M., and Gescher, J. (2012). Dissimilatory reduction of extracellular electron acceptors in anaerobic respiration. *Appl. Environ. Microbiol.* 78, 913–921. doi: 10.1128/AEM.06803-11
- Romão, C. V., Archer, M., Lobo, S. A., Louro, R. O., Pereira, I. A. C., Saraiva, L. M., et al. (2012). “Diversity of heme proteins in sulfate reducing bacteria,” in *Handbook of Porphyrin Science*, eds K. M. Kadish, K. M. Smith, and R. Guilard (Singapore: World Scientific Publishing Co.), 139–229.
- Roy, A., Kucukural, A., and Zhang, Y. (2010). I-TASSER: a unified platform for automated protein structure and function prediction. *Nat. Protoc.* 5, 725–738. doi: 10.1038/nprot.2010.5
- Sagulenko, E., Morgan, G. P., Webb, R. I., Yee, B., Lee, K. C., and Fuerst, J. A. (2014). Structural studies of planctomycete Gemmata obscuriglobus support cell compartmentalisation in a bacterium. *PLoS ONE* 9:e91344. doi: 10.1371/journal.pone.0091344
- Santarella-Mellwig, R., Pruggnaller, S., Roos, N., Mattaj, I. W., and Devos, D. P. (2013). Three-dimensional reconstruction of bacteria with a complex endomembrane system. *PLoS Biol.* 11:e1001565. doi: 10.1371/journal.pbio.1001565
- Schäfer, G. (2004). “Respiratory chains in Archaea: from minimal systems to supercomplexes,” in *Respiration in Archaea and Bacteria. Diversity of Prokaryotic Respiratory Systems*, ed D. Zannoni (Dordrecht: Springer), 1–33.
- Schäfer, G., Engelhard, M., and Müller, V. (1999). Bioenergetics of the Archaea. *Microbiol. Mol. Biol. Rev.* 63, 570–620.
- Scharf, B., Wittenberg, R., and Engelhard, M. (1997). Electron transfer proteins from the haloalkaliphilic archaeon *Natronobacterium pharaonis*: possible components of the respiratory chain include cytochrome bc and a terminal oxidase cytochrome ba3. *Biochemistry* 36, 4471–4479. doi: 10.1021/bi962312d
- Seiler, M., Mehrle, A., Poustka, A., and Wiemann, S. (2006). The 3of5 web application for complex and comprehensive pattern matching in protein sequences. *BMC Bioinformatics* 7:144. doi: 10.1186/1471-2105-7-144
- Sharma, S., Cavallaro, G., and Rosato, A. (2010). A systematic investigation of multiheme c-type cytochromes in prokaryotes. *J. Biol. Inorg. Chem.* 15, 559–571. doi: 10.1007/s00775-010-0623-4
- Simon, J., and Hederstedt, L. (2011). Composition and function of cytochrome c biogenesis System II. *FEBS J.* 278, 4179–4188. doi: 10.1111/j.1742-4658.2011.08374.x
- Simon, J., Kern, M., Hermann, B., Einsle, O., and Butt, J. N. (2011). Physiological function and catalytic versatility of bacterial multihaem cytochromes c involved in nitrogen and sulfur cycling. *Biochem. Soc. Trans.* 39, 1864–1870. doi: 10.1042/BST20110713
- Smith, J., Aklujkar, M., Risso, C., Leang, C., Giloteaux, L., and Holmes, D. (2015). Mechanisms involved in Fe(III) respiration by the hyperthermophilic archaeon *Ferroglobus placidus*. *Appl. Environ. Microbiol.* 81, 2735–2744. doi: 10.1128/AEM.04038-14
- Sreeramulu, K. (2003). Purification and partial characterization of a cytochrome c_{552} from *Halobacterium salinarum*. *Ind. J. Biochem. Biophys.* 40, 274–277.
- Sreeramulu, K., Schmidt, C. L., Schäfer, G., and Anemüller, S. (1998). Studies of the electron transport chain of the euryarchaeon *Halobacterium salinarum*: indications for a type II NADH dehydrogenase and a complex III analog. *J. Bioenerg. Biomembr.* 30, 443–453. doi: 10.1023/A:1020538129400
- Stevens, J. M., Mavridou, D. A., Hamer, R., Kritsiligkou, P., Goddard, A. D., and Ferguson, S. J. (2011). Cytochrome c biogenesis system I. *FEBS J.* 278, 4170–4178. doi: 10.1111/j.1742-4658.2011.08376.x
- Stieglmeier, M., Klingl, A., Alves, R. J., Rittmann, S. K., Melcher, M., Leisch, N., et al. (2014). Nitrososphaera viennensis gen. nov., sp. nov., an aerobic and mesophilic, ammonia-oxidizing archaeon from soil and a member of the archaeal phylum Thaumarchaeota. *Int. J. Syst. Evol. Microbiol.* 64, 2738–2752. doi: 10.1099/ijss.0.063177-0
- Tan, K., Duquette, M., Liu, J. H., Lawler, J., and Wang, J. H. (2008). The crystal structure of the heparin-binding reelin-N domain of f-spondin. *J. Mol. Biol.* 381, 1213–1223. doi: 10.1016/j.jmb.2008.06.045
- Tanaka, M., Ogawa, N., Ihara, K., Sugiyama, Y., and Mukohata, Y. (2002). Cytochrome aa(3) in *Haloferax volcanii*. *J. Bacteriol.* 184, 840–845. doi: 10.1128/JB.184.3.840-845.2002
- Thauer, R. K., Kaster, A. K., Seedorf, H., Buckel, W., and Hedderich, R. (2008). Methanogenic archaea: ecologically relevant differences in energy conservation. *Nat. Rev. Microbiol.* 6, 579–591. doi: 10.1038/nrmicro1931
- van der Star, W. R., Dijkema, C., de Waard, P., Picóreau, C., Strous, M., and Van Loosdrecht, M. C. (2010). An intracellular pH gradient in the anammox bacterium *Kuenenia stuttgartiensis* as evaluated by 31P NMR. *Appl. Microbiol. Biotechnol.* 86, 311–317. doi: 10.1007/s00253-009-2309-9
- van Niftrik, L., Geerts, W. J., van Donselaar, E. G., Humber, B. M., Webb, R. I., Fuerst, J. A., et al. (2008a). Linking ultrastructure and function in four genera of anaerobic ammonium-oxidizing bacteria: cell plan, glycogen storage, and localization of cytochrome c proteins. *J. Bacteriol.* 190, 708–717. doi: 10.1128/JB.01449-07
- van Niftrik, L., Geerts, W. J., van Donselaar, E. G., Humber, B. M., Yakushevska, A., Verkleij, A. J., et al. (2008b). Combined structural and chemical analysis of the anammoxosome: a membrane-bounded intracytoplasmic compartment in anammox bacteria. *J. Struct. Biol.* 161, 401–410. doi: 10.1016/j.jsb.2007.05.005
- van Niftrik, L., van Helden, M., Kirchen, S., van Donselaar, E. G., Harhangi, H. R., Webb, R. I., et al. (2010). Intracellular localization of membrane-bound ATPases in the compartmentalized anammox bacterium ‘*Candidatus Kuenenia stuttgartiensis*’. *Mol. Microbiol.* 77, 701–715. doi: 10.1111/j.1365-2958.2010.07242.x
- van Teeseling, M. C., de Almeida, N. M., Klingl, A., Speth, D. R., Op den Camp, H. J., Rachel, R., et al. (2014). A new addition to the cell plan of anammox bacteria: “*Candidatus Kuenenia stuttgartiensis*” has a protein surface layer as the outermost layer of the cell. *J. Bacteriol.* 196, 80–89. doi: 10.1128/JB.00988-13
- van Teeseling, M. C. F., Mesman, R. J., Kuru, E., Espaillat, A., Cava, F., Brun, Y. V., et al. (in press). Anammox Planctomycetes have a peptidoglycan cell wall. *Nat. Comm.*
- van Teeseling, M. C., Neumann, S., and van Niftrik, L. (2013). The anammoxosome organelle is crucial for the energy metabolism of anaerobic ammonium oxidizing bacteria. *J. Mol. Microbiol. Biotechnol.* 23, 104–117. doi: 10.1159/000346547
- Veith, A., Klingl, A., Zolghadr, B., Lauber, K., Mentele, R., Lottspeich, F., et al. (2009). *Acidianus*, *Sulfolobus* and *Metallosphaera* surface layers: structure, composition and gene expression. *Mol. Microbiol.* 73, 58–72. doi: 10.1111/j.1365-2958.2009.06746.x
- Verissimo, A. F., and Daldal, F. (2014). Cytochrome c biogenesis system I: an intricate process catalyzed by a maturase supercomplex? *Biochim. Biophys. Acta* 1837, 989–998. doi: 10.1016/j.bbabi.2014.03.003
- Wang, F. P., Zhang, Y., Chen, Y., He, Y., Qi, J., Hinrichs, K. U., et al. (2014). Methanotrophic archaea possessing diverging methane-oxidizing and electron-transporting pathways. *ISME J.* 8, 1069–1078. doi: 10.1038/ismej.2013.212
- Wang, M., Tomb, J. F., and Ferry, J. G. (2011). Electron transport in acetate-grown *Methanosaerica aceticivorans*. *BMC Microbiol.* 11:165. doi: 10.1186/1471-2180-11-165
- Wankel, S. D., Adams, M. M., Johnston, D. T., Hansel, C. M., Joye, S. B., and Girguis, P. R. (2012). Anaerobic methane oxidation in metalliferous hydrothermal sediments: influence on carbon flux and decoupling from sulfate reduction. *Environ. Microbiol.* 14, 2726–2740. doi: 10.1111/j.1462-2920.2012.02825.x
- Wu, M. L., Van Teeseling, M. C., Willemse, M. J., Van Donselaar, E. G., Klingl, A., Rachel, R., et al. (2012). Ultrastructure of the denitrifying methanotroph “*Candidatus Methyloirabilis oxyfera*,” a novel polygon-shaped bacterium. *J. Bacteriol.* 194, 284–291. doi: 10.1128/JB.05816-11
- Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.
- Copyright © 2015 Kletzin, Heimerl, Flechsler, van Niftrik, Rachel and Klingl. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



S-layer and cytoplasmic membrane – exceptions from the typical archaeal cell wall with a focus on double membranes

Andreas Klingl*

Plant Development, Department of Biology, Biocenter LMU Munich – Botany, Ludwig Maximilian University Munich, Munich, Germany

Edited by:

Sonja-Verena Albers, University of Freiburg, Germany

Reviewed by:

Jerry Eichler, Ben-Gurion University of the Negev, Israel

Benjamin Harry Meyer, University of Freiburg, Germany

***Correspondence:**

Andreas Klingl, Plant Development, Department of Biology, Biocenter LMU Munich – Botany, Ludwig Maximilian University Munich, Grosshaderner Street 2–4, Planegg-Martinsried 82152, Munich, Germany
e-mail: andreas.klingl@biologie.uni-muenchen.de

The common idea of typical cell wall architecture in archaea consists of a pseudo-crystalline proteinaceous surface layer (S-layer), situated upon the cytoplasmic membrane. This is true for the majority of described archaea, hitherto. Within the crenarchaea, the S-layer often represents the only cell wall component, but there are various exceptions from this wall architecture. Beside (glycosylated) S-layers in (hyper)thermophilic cren- and euryarchaea as well as halophilic archaea, one can find a great variety of other cell wall structures like proteoglycan-like S-layers (Halobacteria), glutaminylglycan (Natronococci), methanochondroitin (*Methanosarcina*) or double layered cell walls with pseudomurein (*Methanothermus* and *Methanopyrus*). The presence of an outermost cellular membrane in the crenarchaeal species *Ignicoccus hospitalis* already gave indications for an outer membrane similar to Gram-negative bacteria. Although there is just limited data concerning their biochemistry and ultrastructure, recent studies on the euryarchaeal methanogen *Methanomassiliicoccus luminyensis*, cells of the ARMAN group, and the SM1 euryarchaeon delivered further examples for this exceptional cell envelope type consisting of two membranes.

Keywords: archaea, S-layer, outer membrane, cytoplasmic membrane, cell wall

INTRODUCTION

Microorganisms and especially archaea can be found in almost any kind of extreme environment, although they are not limited to them: high temperature, high acidity, high pressure, anoxic, no organic substrates. In those habitats, various species of hyper-thermophilic or more generally extremophilic archaea were found and described. Therefore, the general cell plan of the majority of these extremophilic archaea and especially their cell wall architecture might represent the most basic and archaic version: a pseudo-crystalline proteinaceous surface layer (S-layer), a so called S-layer which is situated upon a single cytoplasmic membrane which is enclosing the cytoplasm. This simple cell plan was found to be present in the majority of described archaeal species. Because of its simplicity and widespread distribution within the major groups of archaea and bacteria, it was already stated by Albers and Meyer (2011) that the S-layer might be the cell wall variant that has evolved the earliest. Especially within the crenarchaea, the S-layer usually depicts the only cell wall component. S-layer glycoproteins were first discovered and extensively studied in halophilic archaea, namely *Halobacterium salinarum* as well as in *Haloflexax volcanii* (Houwink, 1956; Mescher and Strominger, 1976a,b; Lechner and Sumper, 1987; Sumper and Wieland, 1995; Sumper et al., 1990) and *Halococcus* (Brown and Cho, 1970) or methanogens like *Methanosarcina* (Kandler and Hippe, 1977), *Methanothermus fervidus* (Kandler and König, 1993; Kärcher et al., 1993) and *Methanococcus* species like *Methanococcus vannielii* and *Methanococcus thermolithotrophicus* (Koval and Jarrell, 1987; Nußer and König, 1987). Amongst others, several studies were carried out focusing on the S-layer in various *Sulfolobus* species. The members of the order Sulfolobales, e.g., *Sulfolobus solfataricus* or *Metallosphaera sedula*, represent model organisms for the basic

structure of this kind of cell wall (Veith et al., 2009; Albers and Meyer, 2011).

But as various examples in the past could show, the archaeal cell wall architecture is not always that simple. Beside the (glycosylated) S-layers in halophilic, thermophilic and hyperthermophilic eury – as well as crenarchaea, one can find a great variety of totally different cell wall structures that sometimes resemble biological substances also found in eukaryotes and bacteria, e.g., glutaminylglycan in Natronococci, methanochondroitin in *Methanosarcina* or double layered cell walls containing pseudomurein in *Methanothermus* and *Methanopyrus* to name just a few (König et al., 2007; Albers and Meyer, 2011; Klingl et al., 2013).

In addition, the finding of an energized outermost cellular membrane in the well described *Ignicoccus hospitalis* and related species already indicated the possibility of an outer membrane (OM), as it is present in Gram-negative bacteria. Furthermore, recent results on the SM1 euryarchaeon, ultra-small ARMAN cells and *Methanomassiliicoccus luminyensis* strengthened the idea of a real archaeal OM and, besides others, will also be discussed here (Comolli et al., 2009; Dridi et al., 2012; Perras et al., 2014). And in this concern, the possible functions of an OM in regard to the bacterial version as well as challenges concerning energetic problems become apparent.

ARCHAEAL CELL WALLS

Similar to bacteria, the cytoplasm in archaea is enclosed by a cytoplasmic membrane built up mainly of glycerol phosphate phospholipids, although with slight differences in membrane lipid composition (Kates, 1992; Albers and Meyer, 2011; Klingl et al., 2013). But instead of fatty acids linked to the (*sn*)-1,2 positions of glycerol via ester bonds, the lipid core of archaea consists of

C_5 isoprenoid units coupled to glycerol via ether bonds in an archaeal cell wall variants (Figure 1). According to some recent findings, there will be a special focus on archaea that could be shown to be surrounded by double membranes.

S-LAYER

Most commonly, the archaeal cell envelope consists of a protein or glycoprotein S-layer, a so called S-layer, forming a 2-D pseudo-crystalline array on the cell surface with a distinct symmetry (Kandler and König, 1985; Beveridge and Graham, 1991; Baumeister and Lembcke, 1992; Messner and Sleytr, 1992; Kandler and König, 1993; Sumper and Wieland, 1995; Veith et al., 2009; Albers and Meyer, 2011; Klingl et al., 2013). They are usually composed of one type of (glyco-)protein forming a central crystal unit consisting of two, three, four, or six subunits which equates p2-, p3-, p4- or p6-symmetry, respectively (Figure 2; Sleytr et al., 1988; Sleytr et al., 1999; Eichler, 2003).

This protein array is usually anchored in the cytoplasmic membrane via stalk like structures forming a quasi-periplasmic space. The lattice constants for those S-layer crystals were shown to vary between 11 and 30 nm with protein masses between 40 and

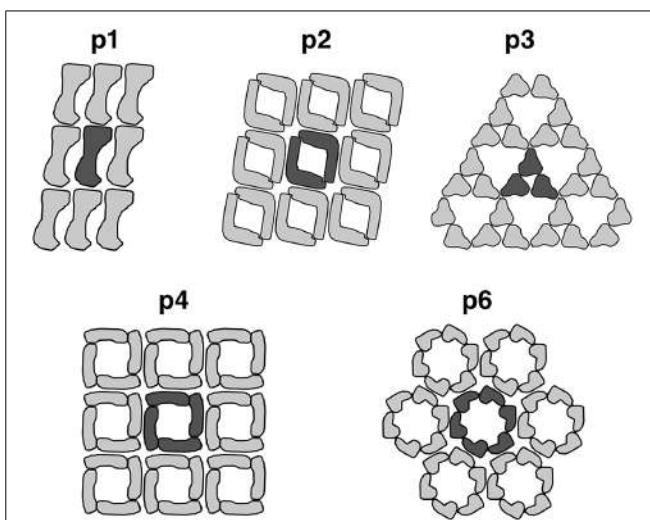


FIGURE 2 | Arrangement of S-layer subunits in the respective symmetry types. The unit cell of the 2D-protein crystals are indicated in dark grey. Reproduced from Sleytr et al. (1999) with the permission of WILEY-VCH Verlag GmbH, Weinheim, Germany

325 kDa (Messner and Sleytr, 1992; König et al., 2007). With some limitations, the S-layer symmetry as well as the center-to-center spacing can be used as a taxonomic trait (König et al., 2007; Klingl et al., 2011). For example, all S-layer proteins of *Sulfolobus* species

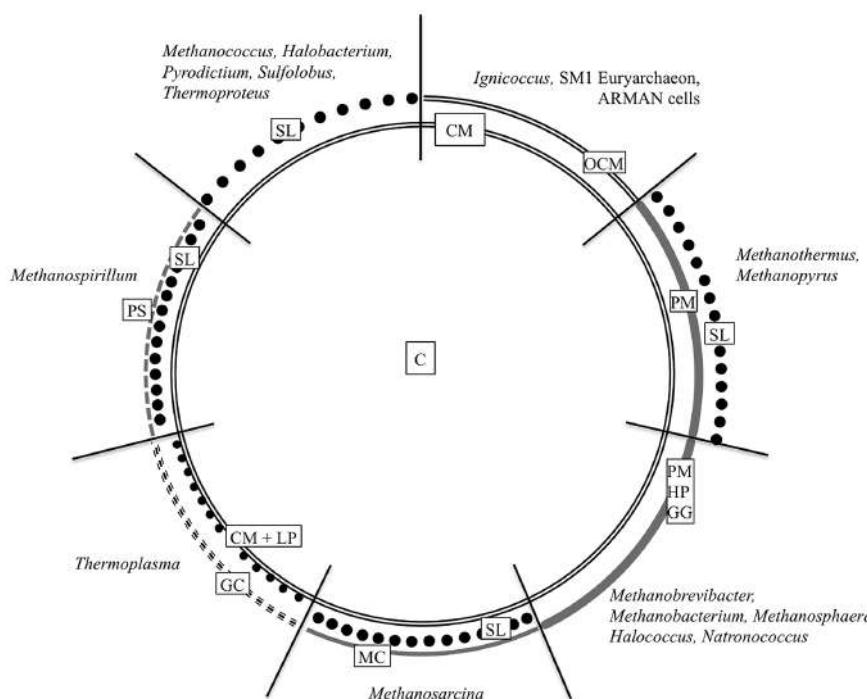


FIGURE 1 | Cell wall organization of Archaea. The schematic illustration summarizes the most common archaeal cell wall types including the most relevant genera. C, cytoplasm; CM, cytoplasmic membrane; GC, glycocalyx; GG, glutaminylglycan; HP, heteropolysaccharide; LP,

lipoglycans; MC, methanochondroitin; OCM, outermost cellular membrane or outer membrane; PM, pseudomurein; PS, protein sheath; SL, S-layer. Based on König et al. (2007) ASM Press, Washington, DC, USA.

described so far revealed a very rare p3-symmetry and a spacing around 21 nm (König et al., 2007; Veith et al., 2009). This symmetry was thought to be unique for the Sulfolobales until recent findings concerning the S-layer of *Nitrososphaera viennensis* could show that this member of the phylum Thaumarchaeota also has a surface protein with p3-symmetry (Stieglmeier et al., 2014).

The S-layer protein of *Halobacterium salinarum* was not only the first glycoprotein discovered in prokaryotes but also exemplifies the fact that S-layers are often highly glycosylated (Mescher and Strominger, 1976a,b; Kandler and König, 1998; König et al., 2007; Veith et al., 2009; Albers and Meyer, 2011). The glycosylation of halophilic S-layer proteins increases protein stability and also prevents degradation (Yurist-Doutsch et al., 2008). Besides the situation in halophilic archaea, the glycosylation may also contribute to a thermal stabilization of S-layer proteins as mentioned in Jarrell et al. (2014).

Concerning the potential function of S-layer proteins, several possibilities have been discussed (Engelhardt, 2007a,b): protection against high temperature, salinity (osmoprotection), low pH and maintenance of cell shape (exoskeleton). Especially within the crenarchaea, they comprise high temperature stability as they have to withstand temperatures around 80°C and pH below 2 in case of Sulfolobales (e.g., Veith et al., 2009). Herein, a high portion of charged amino acids as well as ionic interactions may play an important role (Haney et al., 1999). Another example for the high stability of S-layer proteins was shown for *Thermoproteus tenax* and *Thermofilum pendens*, where the rigid S-layer sacculus even withstands treatment with 2% SDS at 100°C for 30 min (König and Stetter, 1986; Wildhaber and Baumeister, 1987; König et al., 2007). In most euryarchaeota, the situation is totally different with highly labile S-layer proteins (e.g., *Archaeoglobus fulgidus*, König et al., 2007), which also makes it difficult to isolate the proteins. An exception from these findings is the S-layer of *Picrophilus*, which may be a side effect of its high acid stability.

For additional information on general properties of S-layer proteins, their genetic background and characteristic features, the reader's attention should be pointed to some general reviews on this topic (e.g., Claus et al., 2001, 2005; König et al., 2007; Albers and Meyer, 2011). Furthermore, there are also several more focused studies on S-layer proteins of mesophilic and extremely thermophilic archaea (Claus et al., 2002) as well as mesophilic, thermophilic and extremely thermophilic methanococci (Akça et al., 2002).

PSEUDOMUREIN, METHANOCHONDROITIN, AND PROTEIN SHEATHS

Furthermore, pseudomurein, a polymer which maintains the cell shape and perhaps also protects the cells, can be found as an additional second cell wall compound in all species of *Methanothermus* and *Methanopyrus* (König et al., 2007). It shows similarity to bacterial peptidoglycan but usually consists of L-N-acetyltaulosaminuronic acid with a β -1,3 linkage to D-N-acetylglucosamine instead of N-acetylmuramic acid linked β -1,4 to D-N-acetylglucosamine as it is the case in bacterial murein (peptidoglycan). In addition, the crosslinking amino acids in pseudomurein are represented by L-amino acids (glutamic acid,

alanine, lysine) instead of D-amino acids in murein (Kandler and König, 1993; König et al., 1994; Albers and Meyer, 2011).

In contrast to single cells, aggregates of *Methanosarcina* spp. produce a substance called methanochondroitin covering the S-layer with the latter one also being present in single cells (Kreisl and Kandler, 1986; Albers and Meyer, 2011). Methanochondroitin, which is similar to chondroitin in the connecting tissue of vertebrates (Kjellen and Lindahl, 1991), consists of a repeating trimer of two N-acetylgalactosamines and one glucuronic acid but differing from vertebral chondroitin in the molar ratio of the monomers and the fact that it is not sulfated (Albers and Meyer, 2011).

The methanogenic archaeal species *Methanospirillum hungatei* and *Methanosaeta concilii* form long chains that are surrounded by a proteinaceous sheath (Zeikus and Bowen, 1975). Beside its high stability against proteases and detergents, it also revealed a paracrystalline structure and functioning as a micro sieve (Kandler and König, 1978; Sprott and McKellar, 1980; König et al., 2007). The specialty of this sheath is that it is surrounding the whole chain and not just the single cells. Each cell is surrounded separately by an inner cell wall consisting of an S-layer (*Methanospirillum hungatei*) or an amorphous granular layer (*Methanosaeta concilii*; Zeikus and Bowen, 1975; Sprott et al., 1979; Zehnder et al., 1980; Beveridge et al., 1985, 1986; Shaw et al., 1985; Beveridge and Graham, 1991; Firtel et al., 1993; Albers and Meyer, 2011).

GLUTAMINYLGlyCAN AND HALOMUCIN

In similarity to poly- γ -D-glutamyl polymers in *Bacillus*, *Sporosarcina* and *Planococcus*, such polymers were also found within the genus *Natronococcus* (Niemetz et al., 1997). In *Natronococcus occultus*, polyglutamin is forming the cell wall but in contrast to similar polymers found in bacteria, the wall polymer in this archaeum is glycosylated. It is consisting of approximately 60 monomers, which are linked via the γ -carboxylic group (König et al., 2007).

In the square shaped extremely halophilic euryarchaeon *Halococcus walsbyi*, cells are surrounded by an S-layer upon the cytoplasmic membrane. Depending on the strain C23^T or HBSQ001, the cells of *H. walsbyi* are surrounded by one or, even more complex, two S-layers, respectively (Burns et al., 2007). In addition, another protein called halomucin is present which is highly similar to mammalian mucin and probably helps the cells to thrive under conditions of up to 2 M MgCl₂ (Bolhuis et al., 2006). Because of the presence of respective genes, *M. walsbyi* is most likely also surrounded by a poly- γ -glutamate capsule (Bolhuis et al., 2006; Albers and Meyer, 2011).

TWO LAYERED CELL WALLS

For both *Methanothermus fervidus* and *Methanopyrus kandleri*, a cell envelope consisting of two distinct layers has been described (Stetter et al., 1981; Kurr et al., 1991; König et al., 2007). In the former case, it is formed by a pseudomurein layer (thickness 15–20 nm) covered by an external S-layer glycoprotein with p6-symmetry. In the latter case, the situation is similar except that no regular arrangement of the outermost layer could be shown for *Methanopyrus* (König et al., 2007). At this point it has to be mentioned that two layered cell walls are not just

limited to *Methanothermus fervidus* and *Methanopyrus kandleri* because other archaea can also possess two cell walls, e.g., *Methanoscincus* species are covered with an S-layer and an optional layer of methanochondroitin. Another example is the previously mentioned *H. walsbyi* strain HBSQ001 that is covered by two S-layers.

DOUBLE MEMBRANES

There are just a few examples of archaeal species described so far, which do not possess one of the previously mentioned cell wall polymers and structures. Members of the Thermoarchaeales like *Ferroplasma acidophilum* completely lack a cell wall, despite growing under harsh conditions like elevated temperatures and low pH. It is therefore thought that a glycocalyx, lipoglycans, or membrane-associated glycoproteins substitute the function of a cell wall for these organisms (Albers and Meyer, 2011). The hyperthermophilic sulfur-oxidizing crenarchaeal species *Ignicoccus hospitalis* was the first archaeon, for which a double membrane system was described (Huber et al., 2002, 2012; Rachel et al., 2002; Näther and Rachel, 2004; Jungblut et al., 2008; Küper et al., 2010). This is also true for all other species within the genus *Ignicoccus* investigated up to date. It is a highly complex and dynamic system leading to a compartmentalized cell with a huge periplasm enclosed by both membranes. The width of this periplasm can vary from 20 up to 1000 nm (König et al., 2007; Huber et al., 2012). There are some clear differences between both membranes. The inner membrane (IM) consists of archaeol as well as caldarchaeol with the latter one forming tetraether lipids and therefore cannot be separated in freeze fracturing experiments (Rachel et al., 2002, 2010; Burghardt et al., 2007; Huber et al., 2012; Klingl et al., 2013) while the outermost cellular membrane contains archaeol. In addition, most of the polar head groups are glycosylated (Jahn et al., 2004). Interestingly, the ATP synthase as well as the S⁰-H oxidoreductase were shown to be located in this outermost membrane and not in the cytoplasmic membrane, as it could have been expected; *Ignicoccus hospitalis* therefore exhibits an energized outer cellular membrane (Küper et al., 2010).

Beside the two membranes of *Ignicoccus hospitalis* and other closely related species of the genus *Ignicoccus*, recent studies on other archaea could also confirm a double membrane system on these organisms. Three-dimensional cryo electron tomography on cells of some ultra-small archaea belonging to the phylogenetically deeply branching and uncultivated ARMAN lineage revealed an inner and an OM enclosing a periplasm (Comolli et al., 2009). In this special case, they also found indications for cytochromes in the IM. During a study attempting to isolate human-associated archaea, a new genus named *Methanomassiliicoccus luminyensis* was described (Dridi et al., 2012). Although the quality of data concerning the ultrastructure of this organism was poor, it was still possible to recognize an electron dense layer outside the cytoplasmic membrane, most likely represented by an OM. The thick transparent layer mentioned in this study might depict the periplasm of *Methanomassiliicoccus luminyensis*. In a recent study concerning the ultrastructure of the cold-loving archaeal isolate SM1, an outer cellular membrane in addition to

the cytoplasmic membrane could be documented as well (Perras et al., 2014).

With a second, outermost membrane, you get at least two separated compartments like in Gram-negative bacteria: the cytoplasm and the (pseudo)periplasm (Rigel and Silhavy, 2012). In Gram-negative bacteria, the periplasm can make up about 10% of the cell volume and constitutes an oxidizing environment, containing soluble proteins, the thin peptidoglycan layer and usually no ATP (Ruiz et al., 2006). In the special case of *Ignicoccus*, the volume of the intermembrane compartment as an analog to the bacterial periplasm can even be higher than that of the cytoplasm (Küper et al., 2010). Like in bacteria, the presence of membrane proteins and pores makes the OM a permeable and selective barrier (Rigel and Silhavy, 2012). Although there are differences in lipid and protein composition of the inner and outermost cellular membrane in *Ignicoccus hospitalis* (Burghardt et al., 2007; Küper et al., 2010) it still has to be elucidated if there is also an asymmetric OM containing LPS (lipopolysaccharide) present in archaea. In Gram-negative bacteria, one can find a phospholipid bilayer (IM) and usually an asymmetric bilayer in case of the OM, including proteins like transporters or channels (Ruiz et al., 2006; Rigel and Silhavy, 2012). In the OM, the inner leaflet is composed of phospholipids; the outer leaflet is mainly composed of LPS, which is essential for the barrier function of the OM (Ruiz et al., 2006): lipid A, a core oligosaccharide and an O-antigen polysaccharide with variations in length. In similarity to Gram-negative bacteria, archaea with two membranes are featuring several problems: They need lipoproteins and integral OM proteins (OMPs) in the OM. The latter ones are essential for intake of nutrients and export of waste products as they serve as channels (Ruiz et al., 2006). Furthermore, it also shows the importance of a specific system for the biogenesis of OMs and the secretion system in archaea as it was described for *Escherichia coli*, for example (Tokuda, 2009).

SUMMARY AND OUTLOOK

Although a cytoplasmic membrane superimposed by an S-layer depicts the most common cell wall architecture in archaea, there are various other cell wall versions present in cren- as well as in euryarchaeota. As they were isolated from totally different biotopes, it cannot be generalized that one certain environmental condition leads to a certain kind of cell wall (König et al., 2007), this is true for halophilic archaea in particular and all other archaea in general. With the increasing number of archaea, which were described to be surrounded by two membranes like ultra-small ARMAN cells, *Methanomassiliicoccus luminyensis* or the SM1 euryarchaeon, particular attention should be paid to this topic. For example the SM1 euryarchaeon was already known for more than 10 years, without having data about its cell wall structure (Rudolph et al., 2001).

Interestingly, a common feature of all archaea that posses a double membrane cell wall architecture is that they are closely interacting with other organisms (archaea, bacteria, eukaryotes), as already mentioned by Perras et al. (2014), and that they are difficult to cultivate or even not cultivatable at all. At this point, it can still be discussed if the S-layer (Albers and Meyer, 2011) or an OM is the more archaic cell wall compound. With recently

developed and refined isolation and preparation methods, ongoing investigations should be able to shed light on further structural and biochemical features of archaeal outermost cellular membranes. Especially the localization of protein complexes like the ATPase in the cytoplasmic membrane like in Gram-negative bacteria or in the outermost cellular membrane like in *Ignicoccus hospitalis* (Küper et al., 2010) seems to be crucial in this concern.

ACKNOWLEDGMENTS

The LOEWE Research Centre for Synthetic Microbiology (Synmikro) supported this work. I wish to thank Uwe-G. Maier for allocation of the EM facility in Marburg and Marion Debus for technical assistance.

REFERENCES

- Akça, E., Claus, H., Schultz, N., Karbach, G., Schlott, B., Debaerdemaeker, T., et al. (2002). Genes and derived amino acid sequences of S-layer proteins from mesophilic, thermophilic and extremely thermophilic methanococci. *Extremophiles* 6, 351–358. doi: 10.1007/s00792-001-0264-1
- Albers, S.-V., and Meyer, B. H. (2011). The archaeal cell envelope. *Nat. Rev. Microbiol.* 9, 414–426. doi: 10.1038/nrmicro2576
- Baumeister, W., and Lembcke, G. (1992). Structural features of archaeabacterial cell envelopes. *J. Bioenerg. Biomembr.* 24, 567–575. doi: 10.1007/BF00762349
- Beveridge, T. J., and Graham, L. L. (1991). Surface layers of bacteria. *Microbiol. Rev.* 55, 684–705.
- Beveridge, T. J., Patel, G. B., Harris, B. J., and Sprott, G. D. (1986). The ultrastructure of *Methanothrix concilii*, a mesophilic aceticlastic methanogen. *Can. J. Microbiol.* 32, 703–710. doi: 10.1139/m86-128
- Beveridge, T. J., Stewart, M., Doyle, R. J., and Sprott, G. D. (1985). Unusual stability of the *Methanospirillum hungatei* sheath. *J. Bacteriol.* 162, 728–737.
- Bolhuis, H., Palm, P., Wende, A., Falb, M., Rampp, M., Rodriguez-Valera, F., et al. (2006). The genome of the square archaeon *Haloquadratum walsbyi*: life at the limits of water activity. *BMC Genomics* 7:169. doi: 10.1186/1471-2164-7-169
- Brown, A. D., and Cho, K. J. (1970). The cell walls of extremely halophilic cocci. Gram-positive bacteria lacking muramic acid. *J. Gen. Microbiol.* 62, 267–270. doi: 10.1099/00221287-62-2-267
- Burghardt, T., Näther, D. J., Junglas, B., Huber, H., and Rachel, R. (2007). The dominating outer membrane protein of the hyperthermophilic archaeum *Ignicoccus hospitalis*: a novel pore-forming complex. *Mol. Microbiol.* 63, 166–176. doi: 10.1111/j.1365-2958.2006.05509.x
- Burns, D. G., Janssen, P. H., Itoh, T., Kamekura, M., Li, Z., Jensen, G., et al. (2007). *Haloquadratum walsbyi* gen. nov., sp. nov., the square haloarchaeon of Walsby, isolated from saltern crystallizers in Australia and Spain. *Int. J. Syst. Evol. Microbiol.* 57, 387–392. doi: 10.1099/ijst.0.64690-0
- Claus, H., Akça, E., Debaerdemaeker, T., Evrard, C., Declercq, J. P., Harris, J. R., et al. (2005). Molecular organization of selected prokaryotic S-layer proteins. *Can. J. Microbiol.* 51, 731–743. doi: 10.1139/w05-093
- Claus, H., Akça, E., Debaerdemaeker, T., Evrard, C., Declercq, J. P., and König, H. (2002). Primary structure of selected archaeal mesophilic and extremely thermophilic outer surface proteins. *Syst. Appl. Microbiol.* 25, 3–12. doi: 10.1078/0723-2020-00100
- Claus, H., Akça, E., Schultz, N., Karbach, G., Schlott, B., Debaerdemaeker, T., et al. (2001). “Surface (glyco-)proteins: primary structure and crystallization under microgravity conditions,” in *Proceedings of the First European Workshop on Exo-/Astro-Biology, ESA SP-496*, Frascati, 806–809.
- Comolli, L. R., Baker, B. J., Downing, K. H., Soegerist, C. E., and Banfield, J. F. (2009). Three-dimensional analysis of the structure and ecology of a novel, ultra-small archaeon. *ISME J.* 3, 159–167. doi: 10.1038/ismej.2008.99
- Dridi, B., Fardeau, M.-L., Ollivier, B., Raoult, D., and Drancourt, M. (2012). *Methanomassiliicoccus luminyensis* gen. nov., sp. nov., a methanogenic archaeon isolated from human faeces. *Int. J. Syst. Evol. Microbiol.* 62, 1902–1907. doi: 10.1099/ijst.0.033712-0
- Eichler, J. (2003). Facing extremes: archaeal surface-layer (glyco)proteins. *Microbiology* 149, 3347–3351. doi: 10.1099/mic.0.26591-0
- Engelhardt, H. (2007a). Are S-layers exoskeletons? The basic function of protein surface layers revisited. *J. Struct. Biol.* 160, 115–124. doi: 10.1016/j.jsb.2007.08.003
- Engelhardt, H. (2007b). Mechanism of osmoprotection by archaeal S-layers: a theoretical study. *J. Struct. Biol.* 160, 190–199. doi: 10.1016/j.jsb.2007.08.004
- Firtel, M., Southam, G., Harauz, G., and Beveridge, T. J. (1993). Characterization of the cell wall of the sheathed methanogen *Methanospirillum hungatei* Gp1 as an S-layer. *J. Bacteriol.* 175, 7550–7560.
- Haney, P. J., Badger, J. H., Buldak, G. L., Reich, C. I., Woese, C. R., and Olsen, G. J. (1999). Thermal adaption analyzed by comparison of protein sequences from mesophilic and extremely thermophilic *Methanococcus* species. *Proc. Natl. Acad. Sci. U.S.A.* 96, 3578–3583. doi: 10.1073/pnas.96.7.3578
- Houwink, A. L. (1956). Flagella, gas vacuoles and cell-wall structure in *Halobacterium halobium*; an electron microscope study. *J. Gen. Microbiol.* 15, 146–150. doi: 10.1099/00221287-15-1-146
- Huber, H., Hohn, M. J., Rachel, R., Fuchs, T., Wimmer, V. C., and Stetter, K. O. (2002). A new phylum of archaea represented by a nanosized hyperthermophilic symbiont. *Nature* 417, 63–67. doi: 10.1038/417063a
- Huber, H., Küper, U., Daxer, S., and Rachel, R. (2012). The unusual cell biology of the hyperthermophilic Crenarchaeon *Ignicoccus hospitalis*. *Antonie Van Leeuwenhoek* 102, 203–219. doi: 10.1007/s10482-012-9748-5
- Jahn, U., Summons, R., Sturt, H., Grossjean, E., and Huber, H. (2004). Composition and source of the lipids of *Nanoarchaeum equitans* and their origin in the cytoplasmic membrane of its host *Ignicoccus* sp. KIN4I. *Arch. Microbiol.* 182, 404–413. doi: 10.1007/s00203-004-0725-x
- Jarrell, K. G., Ding, Y., Meyer, B. H., Albers, S.-V., Kaminski, L., and Eichler, J. (2014). N-linked glycosylation in archaea: a structural, functional, and genetic analysis. *Microbiol. Mol. Biol. Rev.* 78, 304–341. doi: 10.1128/MMBR.00052-13
- Junglas, B., Briegel, A., Burghardt, T., Walther, P., Wirth, R., Huber, H., et al. (2008). *Ignicoccus hospitalis* and *Nanoarchaeum equitans*: ultrastructure, cell-cell interaction, and 3D reconstruction from serial sections of freeze-substituted cells and by electron cryotomography. *Arch. Microbiol.* 190, 395–408. doi: 10.1007/s00203-008-0402-6
- Kandler, O., and Hippe, H. (1977). Lack of peptidoglycan in the cell walls of *Methanosarcina barkeri*. *Arch. Microbiol.* 113, 57–60. doi: 10.1007/BF00428580
- Kandler, O., and König, H. (1978). Chemical composition of the peptidoglycan-free cell walls of methanogenic bacteria. *Arch. Microbiol.* 118, 141–152. doi: 10.1007/BF00415722
- Kandler, O., and König, H. (1985). “Cell envelopes of archaeabacteria,” in *The Bacteria. A Treatise on Structure and Function. Archaeabacteria*, Vol. VIII, C. R. Woese and R. S. Wolfe (New York: Academic Press), 413–457.
- Kandler, O., and König, H. (1993). “Cell envelopes of archaea: structure and chemistry,” in *The Biochemistry of Archaea (Archaeabacteria)*, eds M. Kates, D. Kushner, and A. T. Matheson (Amsterdam: Elsevier Science Publication), 223–333.
- Kandler, O., and König, H. (1998). Cell wall polymers in archaea (Archaeabacteria). *Cell Mol. Life Sci.* 54, 305–308. doi: 10.1007/s000180050156
- Kärcher, U., Schröder, H., Haslinger, E., Allmeier, G., Schreiner, R., Wieland, F., et al. (1993). Primary structure of the heterosaccharide of the surface glycoprotein of *Methanothermus fervidus*. *J. Biol. Chem.* 268, 26821–26826.
- Kates, M. (1978). The phytanyl ether-linked polar lipids and isoprenoid neutral lipids of extremely halophilic bacteria. *Prog. Chem. Fats Other Lipids* 15, 301–342. doi: 10.1016/0079-6832(77)90011-8
- Kates, M. (1992). Archaeabacterial lipids: structure, biosynthesis and function. *Biochem. Soc. Symp.* 58, 51–72.
- Kjellen, L., and Lindahl, U. (1991). Proteoglycans: structures and interactions. *Annu. Rev. Biochem.* 60, 443–475. doi: 10.1146/annurev.bi.60.070191.002303
- Klingl, A., Flechsler, J., Heimerl, T., and Rachel, R. (2013). Archaeal Cells. In: eLS. John Wiley & Sons, Ltd: Chichester.
- Klingl, A., Moissl-Eichinger, C., Wanner, G., Zweck, J., Huber, H., Thomm, M., et al. (2011). Analysis of the surface proteins of *Acidithiobacillus ferrooxidans* strain SP5/1 and the new, pyrite-oxidizing *Acidithiobacillus* isolate HV2/2, and their possible involvement in pyrite oxidation. *Arch. Microbiol.* 193, 867–882. doi: 10.1007/s00203-011-0720-y
- König, H., Hartmann, E., and Kärcher, U. (1994). Pathways and principles of the biosynthesis of methanobacterial cell wall polymers. *Syst. Appl. Microbiol.* 16, 510–517. doi: 10.1016/S0723-2020(11)80320-6

- König, H., Rachel, R., and Claus, H. (2007). "Proteinaceous surface layers of archaea: ultrastructure and biochemistry," in *Archaea: Molecular and Cell Biology*, ed. R. Cavicchioli (Washington, DC: American Society of Microbiology Press), 315–340.
- König, H., and Stetter, K. O. (1986). Studies on archaeabacterial S-layers. *Syst. Appl. Microbiol.* 7, 300–309. doi: 10.1016/S0723-2020(86)80023-6
- Koval, S., and Jarrell, K. (1987). Ultrastructure and biochemistry of the cell wall of *Methanococcus voltae*. *J. Bacteriol.* 169, 1298–1306.
- Kreisl, P., and Kandler, O. (1986). Chemical structure of the cell wall polymer of *Methanosarcina*. *Syst. Appl. Microbiol.* 7, 293–299. doi: 10.1016/S0723-2020(86)80022-4
- Küper, U., Meyer, C., Müller, V., Rachel, R., and Huber, H. (2010). Energized outer membrane and spatial separation of metabolic processes in the hyperthermophilic archaeon *Ignicoccus hospitalis*. *Proc. Natl. Acad. Sci. U.S.A.* 107, 3152–3156. doi: 10.1073/pnas.0911711107
- Kurr, M., Huber, R., König, H., Jannasch, H. W., Fricke, H., Trincone, A., et al. (1991). *Methanopyrus kandleri*, gen. and sp. nov. represents a novel group of hyperthermophilic methanogens, growing at 110° C. *Arch. Microbiol.* 156, 239–247. doi: 10.1007/BF00262992
- Lechner, J., and Sumper, M. (1987). The primary structure of a prokaryotic glycoprotein. Cloning and sequencing of the cell surface glycoprotein gene of halobacteria. *J. Biol. Chem.* 262, 9724–9729.
- Mescher, M. F., and Strominger, J. L. (1976a). Purification and characterisation of a prokaryotic glycoprotein from the cell envelope of *Halobacterium salinarium*. *J. Biol. Chem.* 251, 2005–2014.
- Mescher, M. F., and Strominger, J. L. (1976b). Structural (shape-maintaining) role of the cell surface glycoprotein of *Halobacterium salinarium*. *Proc. Natl. Acad. Sci. U.S.A.* 73, 2687–2691. doi: 10.1073/pnas.73.8.2687
- Messner, P., and Sleytr, U. B. (1992). Crystalline bacterial cell-surface layers. *Adv. Microbial. Physiol.* 33, 213–274. doi: 10.1016/S0065-2911(08)60218-0
- Näther, D. J., and Rachel, R. (2004). The outer membrane of the hyperthermophilic archaeon *Ignicoccus*: dynamics, ultrastructure and composition. *Biochem. Soc. Trans.* 32, 199–203. doi: 10.1042/BST0320199
- Niemetz, R., Kärcher, U., Kandler, O., Tindall, B., and König, H. (1997). The cell wall polymer of the extremely halophilic archaeon *Natronococcus occultus*. *Eur. J. Biochem.* 249, 905–911. doi: 10.1111/j.1432-1033.1997.0905.x
- Nußer, E., and König, H. (1987). S-layer studies on three species of *Methanococcus* living at different temperatures. *Can. J. Microbiol.* 33, 256–261. doi: 10.1139/m87-043
- Perras, A. K., Wanner, G., Klingl, A., Mora, M., Auerbach, A. K., Heinz, V., et al. (2014). Grappling archaea: ultrastructural analyses of an uncultivated, cold-loving archaeon, and its biofilm. *Front. Microbiol.* 5:397. doi: 10.3389/fmicb.2014.00397
- Rachel, R., Meyer, C., Klingl, A., Gürster, S., Heimerl, T., Wasserburger, N., et al. (2010). Analysis of the ultrastructure of archaea by electron microscopy. *Methods Cell Biol.* 96, 47–69. doi: 10.1016/S0091-679X(10)96003-2
- Rachel, R., Wyschkony, I., Riehl, S., and Huber, H. (2002). The ultrastructure of *Ignicoccus*: evidence for a novel outer membrane and for intracellular vesicle budding in an archaeon. *Archaea* 1, 9–18. doi: 10.1155/2002/307480
- Rigel, N. W., and Silhavy, T. J. (2012). Making a beta-barrel: assembly of outer membrane proteins in Gram-negative bacteria. *Curr. Opin. Microbiol.* 15, 189–193. doi: 10.1016/j.mib.2011.12.007
- Rudolph, C., Wanner, G., and Huber, R. (2001). Natural communities of novel archaea and bacteria growing in cold sulfuric springs with a string-of-pearls-like morphology. *Appl. Environ. Microbiol.* 67, 2336–2344. doi: 10.1128/AEM.67.5.2336–2344.2001
- Ruiz, N., Kahne, D., and Silhavy, T. J. (2006). Advances in understanding bacterial outer-membrane biogenesis. *Nat. Rev. Microbiol.* 4, 57–66. doi: 10.1038/nrmicro1322
- Shaw, P. J., Hills, G. J., Henwood, J. A., Harris, J. E., and Archer, D. B. (1985). Three-dimensional architecture of the cell sheath and septa of *Methanospirillum hungatei*. *J. Bacteriol.* 161, 750–757.
- Sleytr, U. B., Messner, P., Pum, D., and Sára, M. (1988). "Crystalline bacterial cell surface layers," in *Proceedings of EMBO Workshop on Crystalline Bacterial Surface Layers*. Berlin: Springer Verlag.
- Sleytr, U. B., Messner, P., Pum, D., and Sára, M. (1999). Crystalline bacterial cell surface layers (S-layers): from supramolecular cell structure to biomimetics and nanotechnology. *Angew. Chem. Int. Ed. Engl.* 38, 1034–1054. doi: 10.1002/(SICI)1521-3773(19990419)38:8<1034::AID-ANIE1034>3.0.CO;2-#
- Sprott, G. D., Colvin, J. R., and McKellar, R. C. (1979). Spheroblasts of *Methanospirillum hungatai* formed upon treatment with dithiothreitol. *Can. J. Microbiol.* 25, 730–738. doi: 10.1139/m79-106
- Sprott, G. D., and McKellar, R. C. (1980). Composition and properties of the cell wall of *Methanospirillum hungatei*. *Can. J. Microbiol.* 26, 115–120. doi: 10.1139/m80-017
- Stetter, K. O., Thomm, M., Winter, J., Wildgruber, G., Huber, H., Zillig, W., et al. (1981). *Methanothermus fervidus*, sp. nov., a novel extremely thermophilic methanogen isolated from an Icelandic hot spring. *Zbl. Bakt. Hyg. I. Abt. Orig. C2*, 166–178.
- Stiegelmeyer, M., Klingl, A., Rittmann, S., Alves, R. E., Melcher, M., Leisch, N., et al. (2014). *Nitrososphaera viennensis* sp. nov., an aerobic and mesophilic ammonia oxidizing archaeon from soil and member of the novel archaeal phylum Thaumarchaeota. *Int. J. Syst. Evol. Microbiol.* 64, 2738–2752. doi: 10.1099/ijss.0.063172-0
- Sumper, M., Berg, E., Mengelkoch, R., and Strobel, I. (1990). Primary structure and glycosylation of the S-layer protein of *Haloferax volcanii*. *J. Bacteriol.* 172, 7111–7118.
- Sumper, M., and Wieland, F. T. (1995). "Bacterial glycoproteins," in *Glycoproteins*, eds J. Montreuil, J. F. G. Vliegenthart, and H. Schachter (Amsterdam: Elsevier), 455–473.
- Tokuda, H. (2009). Biogenesis of outer membranes in Gram-negative bacteria. *Biosci. Biotechnol. Biochem.* 73, 465–473. doi: 10.1271/bbb.80778
- Veith, A., Klingl, A., Zolghadr, B., Lauber, K., Mentele, R., Lottspeich, F., et al. (2009). *Acidianus*, *Sulfolobus* and *Metallocphaera* surface layers: structure, composition and gene expression. *Mol. Microbiol.* 73, 58–72. doi: 10.1111/j.1365-2958.2009.06746.x
- Wildhaber, I., and Baumeister, W. (1987). The cell envelope of *Thermoproteus tenax*: three-dimensional structure of the surface layer and its role in shape maintenance. *EMBO J.* 6, 1475–1480.
- Yurist-Doutsch, S., Chaban, B., VanDyke, D. J., Jarrell, K. F., and Eichler, J. (2008). Sweet to the extreme: protein glycosylation in archaea. *Mol. Microbiol.* 68, 1079–1084. doi: 10.1111/j.1365-2958.2008.06224.x
- Zehnder, A. J. B., Huser, B. A., Brock, T. B., and Wuhrmann, K. (1980). Characterization of an acetate carboxylating, non hydrogen oxidizing methane bacterium. *Arch. Microbiol.* 124, 1–11. doi: 10.1007/BF00407022
- Zeikus, J. G., and Bowen, V. G. (1975). Fine structure of *Methanospirillum hungatei*. *J. Bacteriol.* 121, 373–380.

Conflict of Interest Statement: The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 26 September 2014; accepted: 31 October 2014; published online: 25 November 2014.

Citation: Klingl A (2014) S-layer and cytoplasmic membrane – exceptions from the typical archaeal cell wall with a focus on double membranes. *Front. Microbiol.* 5:624. doi: 10.3389/fmicb.2014.00624

This article was submitted to Microbial Physiology and Metabolism, a section of the journal Frontiers in Microbiology.

Copyright © 2014 Klingl. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Mining proteomic data to expose protein modifications in *Methanosarcina mazei* strain Gö1

Deborah R. Leon^{1†}, A. Jimmy Ytterberg^{1†}, Pinmanee Boontheung^{1†}, Unmi Kim^{2†}, Joseph A. Loo^{1,3,4}, Robert P. Gunsalus^{2,4*} and Rachel R. Ogorzalek Loo^{3,4*}

¹ Department of Chemistry and Biochemistry, University of California, Los Angeles, Los Angeles, CA, USA

² Microbiology, Immunology, and Molecular Genetics, University of California, Los Angeles, Los Angeles, CA, USA

³ Department of Biological Chemistry, David Geffen School of Medicine, University of California, Los Angeles, Los Angeles, CA, USA

⁴ UCLA-DOE Institute for Genomics and Proteomics, University of California, Los Angeles, Los Angeles, CA, USA

Edited by:

Sonja-Verena Albers, University of Freiburg, Germany

Reviewed by:

Marco Moracci, National Research Council of Italy, Italy

Benjamin Harry Meyer, University of Freiburg, Germany

***Correspondence:**

Robert P. Gunsalus, Microbiology, Immunology, and Molecular Genetics, University of California, Los Angeles, 1602 Molecular Science Bldg., Los Angeles, CA 90095, USA

e-mail: robg@microbio.ucla.edu;
Rachel R. Ogorzalek Loo, Molecular Biology Institute, University of California, 406 Boyer Hall, Los Angeles, Los Angeles, CA 90095, USA

e-mail: rloo@mednet.ucla.edu

†Present address:

Deborah R. Leon, Mass Spectrometry Resource, Boston University School of Medicine, Boston, USA;

A. Jimmy Ytterberg, Department of Medical Biochemistry and Biophysics, Karolinska Institute, Stockholm, Sweden;

Pinmanee Boontheung, Halliburton, Houston, USA;

Unmi Kim, BP Biofuels, San Diego, USA

Proteomic tools identify constituents of complex mixtures, often delivering long lists of identified proteins. The high-throughput methods excel at matching tandem mass spectrometry data to spectra predicted from sequence databases. Unassigned mass spectra are ignored, but could, in principle, provide valuable information on unanticipated modifications and improve protein annotations while consuming limited quantities of material. Strategies to "mine" information from these discards are presented, along with discussion of features that, when present, provide strong support for modifications. In this study we mined LC-MS/MS datasets of proteolytically-digested concanavalin A pull down fractions from *Methanosarcina mazei* Gö1 cell lysates. Analyses identified 154 proteins. Many of the observed proteins displayed post-translationally modified forms, including O-formylated and methyl-esterified segments that appear biologically relevant (i.e., not artifacts of sample handling). Interesting cleavages and modifications (e.g., S-cyanation and trimethylation) were observed near catalytic sites of methanogenesis enzymes. Of 31 *Methanosarcina* protein N-termini recovered by concanavalin A binding or from a previous study, only *M. mazei* S-layer protein MM1976 and its *M. acetivorans* C2A orthologue, MA0829, underwent signal peptide excision. Experimental results contrast with predictions from algorithms SignalP 3.0 and Exprot, which were found to over-predict the presence of signal peptides. Proteins MM0002, MM0716, MM1364, and MM1976 were found to be glycosylated, and employing chromatography tailored specifically for glycopeptides will likely reveal more. This study supplements limited, existing experimental datasets of mature archaeal N-termini, including presence or absence of signal peptides, translation initiation sites, and other processing. *Methanosarcina* surface and membrane proteins are richly modified.

Keywords: S-layers, archaeal surface proteins, *Methanosarcina mazei*, prokaryotic glycosylation, membrane proteins, concanavalin A

INTRODUCTION

Knowledge about Archaea and their proteins is limited, making their characterization important. Fortunately, tools are available to identify proteins at high throughput, while bioinformatic analyses can overlay existing knowledge onto this kingdom. Nevertheless, protein modifications unique to these organisms and/or rare in well-studied microbes may elude us, primarily because high throughput proteomic methods focus on matching peptide fragment data to what can be anticipated, primarily from the genome sequence.

Progress in understanding archaeal cell surface structures has been hindered by the limited availability of experimental results,

and any new protein modifications that are revealed may hint at function. Here peptide tandem mass spectrometry (MS/MS) datasets from previous investigations of *Methanosarcina* S-layer and surface-exposed proteins (Francoleon et al., 2009; Rohlin et al., 2012), as well as from mixtures recovered by concanavalin A binding were selected for further analysis, motivated by interest in how protein modifications can impact organisms' interactions with their environment and with other organisms.

The above-mentioned emphasis on matching high throughput proteomic data to predictions means that *unassigned mass spectra, (often 90% of all data) are ignored* (Savitski et al., 2005; Baumgartner et al., 2008; Falkner et al., 2008; Menschaert et al.,

2009; Hahne et al., 2013). In principle, unassigned proteomic data could be a treasure trove. In practice, its value depends on sample complexity (whether it contains ~5000 or 50,000 tryptic peptides), the protein and/or peptide separation strategies employed for its acquisition [single-dimension liquid chromatography (LC) of tryptic peptides vs. two-dimensional polyacrylamide gel electrophoresis (2D-PAGE) of proteins followed by peptide LC], and on the quality of the tandem mass spectrometry (MS/MS) data.

There are many reasons why peptide tandem mass spectra may be unassigned in these complex experiments:

1. The spectra may be poor in quality and/or low in information content.
2. The MS/MS spectra may be derived from a mixture, rather than from a single peptide.
3. The spectra may be derived from a peptide that is modified in a manner not considered by the algorithm.

High throughput workflows compromise between *speed* and *depth of analysis*, with “success” generally assessed by the number of proteins identified at a specified peptide or protein false discovery rate (FDR). This gene-centric approach does not differentiate between modified and/or processed forms of proteins. Typically, only modifications related to sample handling are considered (Schmidt et al., 2005) to balance *sensitivity* (number of peptides or proteins identified) with *specificity* (low false discovery rate). If post-translational modifications (PTMs) are pursued in high-throughput environments, modified peptides are typically enriched *via* PTM-specific immunoprecipitation or other affinity capture methods; e.g., phosphopeptide binding to TiO₂.

Skillful data mining can recover valuable information about unanticipated modifications from high throughput proteomic data. Some strategies include:

1. Performing an error-tolerant search (Creasy and Cottrell, 2002); i.e., a second search limited to proteins detected in the initial search. It attempts to match mass spectra by considering one additional modification-type from the UniMod collection or single amino acid substitutions (Creasy and Cottrell, 2004). It also relaxes enzyme specificity, e.g., only the N- or C-terminus of a peptide must conform to trypsin’s known cleavage specificity, rather than both.
2. Searching against the genome sequence translated in 6-reading frames (three forward and three reverse strand translations). This approach overcomes DNA sequencing errors, missed open reading frames, incorrect start or stop sites, and alternate initiation. It can also be employed in an error-tolerant mode to overcome nucleobase substitutions.
3. Attempting to match spectra by assuming that one residue’s mass has been shifted by some amount (within a specified range) or by clustering related spectra (Savitski et al., 2006; Bandeira et al., 2007; Falkner et al., 2008; Wilhelm and Jones, 2014).
4. Manual or computer-assisted *de novo* interpretation of mass spectra.

Evaluating accuracy is challenging for any data mining strategy. When searches are limited to relatively common modifications, it may be possible to calculate separate false discovery rates for the modified peptides. But when almost anything is possible, confirmatory information must be sought elsewhere.

Here, we illustrate how proteomic datasets can be mined to recover information about *Methanosarcina mazei* protein modifications and describe some characteristic mass signatures that assist in validating that modifications are present. Clearly, methods including antibody blotting, functional group specific staining, and chemical derivatization also provide essential verification. The focus of this manuscript is on mining existing data to recover information about unanticipated modifications. The knowledge may suggest protein forms (proteoforms) to track in future studies, follow-up experiments to confirm the modifications, it may hint at protein function, or it may simply improve protein annotations.

MATERIALS AND METHODS

CELL CULTIVATION

M. mazei Gö1 was grown at 37°C as single cells (non-aggregated) in pH 6.8 basal mineral medium prepared by the Hungate technique and supplemented with 0.05 M methanol as the sole source of carbon and energy (Sowers et al., 1993). Medium osmolarity was defined by 0.2 M NaCl. Cultivation employed 10-mL anaerobic tubes (Difco, Sparks, MD) sealed with a N₂-CO₂ (4:1) atmosphere. Cultures were harvested at an average OD₆₀₀ of ~1.5.

CRUDE EXTRACT/LYSATE PREPARATION

Eight tubes of 10-mL anaerobic cultures were unsealed and their contents were transferred to 15-mL Falcon™ centrifuge tubes. Cells were sedimented at room temperature for 10 min in a swing bucket rotor at 1125 × g. After centrifugation, 500 µL of chilled lysis buffer was added to each tube [2% (w/v) CHAPS (3-(3-cholamidopropyl) dimethylammonio-1-propanesulfonate) in pH 7.5, 50 mM Tris, 0.15 M NaCl, 1 mM CaCl₂, and 1 mM MnCl₂ supplemented with 2.25 µL Sigma P8465 protease inhibitor]. The cell pellets were disrupted further by multiple freeze/thaw cycles interspersed with vortexing. The lysates were transferred to 1.5 mL microcentrifuge tubes for centrifugation at 16,000 × g for 15 min at 4°C. The soluble lysate was retained for analysis.

CONCANAVALIN A GLYCOPROTEIN ENRICHMENT

Our previous purification was modified slightly (Francoleon et al., 2009). Briefly, each of four 5-mL centrifuge columns (Pierce, Rockford, IL) was loaded with a 2 mL slurry of Con A-coupled agarose beads (Vector Laboratories, Burlingame, CA). The beads were washed with 3 mL of 50 mM Tris, 0.15 M NaCl (pH 7.5) seven times, followed by six equilibrating washes with 2 mL binding buffer [BB, 50 mM Tris, 0.15 M NaCl, 1 mM CaCl₂, 1 mM MnCl₂ (pH 7.5)]. After equilibration, 1 mL of lysate and 1 mL of BB buffer were added to each column for incubation in a room temperature rotor. After 30 min, the column flow through was collected by centrifugation and discarded. The protein-bound beads were washed 10 times to minimize non-specific binding: (1) five washes, each with 2 mL of BB buffer supplemented

with 0.1 % Tween-20, followed by (2) five 2-mL washes with 50 mM of (NH₄)HCO₃ (*pH* 7.8). Glycoproteins were eluted from the lectin media twice: (1) Con A beads were incubated in 2 mL of elution buffer [50 mM (NH₄)HCO₃/0.2 M methyl- α -D-mannopyranoside/0.2 M methyl- α -D-glucopyranoside (*pH* 7.8)] for 10 min at room temperature, and centrifuged to recover the eluate. (2) The elution was repeated with 1 mL of buffer. The combined eluate was concentrated to 1.6 μ g/ μ L (Pierce BCA assay) by ultrafiltration through an Amicon® 50 kDa cut-off cellulose membrane (Millipore, Billerica, MA).

Multiple enrichments were also performed on a smaller scale in a manner similar to that described above, but without ultrafiltration.

IN-SOLUTION TRYPSIN, GLU-C AND ASP-N PROTEOLYSIS

Each proteolytic digestion used 50 μ L of concentrated Con A eluate (~78 μ g of total protein) which, prior to digestion, was precipitated at -20°C overnight in 9 volumes of chilled acetone. Protein precipitate was recovered by centrifugation at 4°C, 16,000 \times *g* for 20 min. Pellets were washed in 500 μ L of chilled 80% acetone/10% methanol/0.2% acetic acid.

For trypsin digestion, the protein was resuspended in 20 μ L of room temperature dimethyl sulfoxide (DMSO) with 600 rpm shaking for 30 min (Ytterberg et al., 2006). The solution was diluted to 30% DMSO/50 mM (NH₄)HCO₃ and sequencing grade trypsin (Promega, Madison, WI) was added at a 1:20 enzyme:protein ratio (*w/w*). Digestion proceeded overnight at 37°C with 300 rpm shaking.

For Glu-C digestion, precipitated protein was resuspended in 131 μ L of *pH* 7.8, 25 mM (NH₄)HCO₃ and 1:20 (*w/w*) Glu-C:protein (sequencing grade *Staphylococcus aureus* Protease V-8, Roche, Indianapolis, IN). The reaction proceeded for 4 h with 300 rpm shaking in a 25°C incubator.

Asp-N cleavage employed sequencing grade *Pseudomonas fragi* Asp-N (Roche) at 1:20 Asp-N:protein (*w/w*). The protein pellet, resuspended in 131 μ L of *pH* 7.5, 25 mM Tris buffer, was incubated for 4 h at 37°C with 300 rpm shaking.

Nano-HPLC and data dependent MS/MS

Aliquots (~5 μ g) of trypsin- and Glu-C-digested peptides were dried by vacuum centrifugation and resuspended in 125 μ L of 5% formic acid (FA) to ~40 ng/ μ L. Asp-N digested proteins (~10 μ g) were desalting using a C-18 spin column (Pierce), dried, and resuspended to ~40 ng/ μ L in 5% formic acid.

Peptide mixtures were analyzed by liquid chromatography-tandem mass spectrometry (LC-MS/MS) with ESI (electrospray ionization) on an Applied BioSystems QSTAR® Pulsar XL quadrupole time-of-flight mass spectrometer equipped with a nanoelectrospray interface (Protana), a Proxeon (Odense) nano-bore stainless steel emitter (30 μ m ID), and an LC Packings nano-LC system as described previously (Francoleon et al., 2009). Homemade pre- (150 \times 5 mm) and analytical (75 μ m \times 150 mm) columns were packed with Jupiter Proteo C12 4- μ m resin (Phenomenex). Typically 6 μ L of sample was loaded onto the precolumn, washed with 0.1% FA for 4 min and transferred to the LC column. The 200 nL/min mobile phase gradient employed 3–6% B in 6 s, 6–24% B in 18 min, 24–36% B in 6 min, 36–80%

B in 2 min, and 80% B for 7.9 min. The column was equilibrated with 3% B for 15 min prior to the next run. Eluents used were 0.1% FA (aq) (solvent A) and 95% CH₃CN containing 0.1% FA (solvent B).

Peptide product ion spectra were recorded automatically by IDA (information-dependent analysis) software on the mass spectrometer. Protein sequence searches employed a conservative mass tolerance of 0.3 Da for both precursor and product ions, and 1 (trypsin) or 2 (Asp-N and Glu-C) missed cleavages. Proteins hits were accepted based on ≥ 2 ascribed peptides, at least of one which possessed a MOWSE score ≥ 26 (*p* \leq 0.02) with MuDPIT scoring. Identifications based on single peptides are presented separately. Correspondences between MS/MS spectra and all ascribed sequences were also verified manually.

Nano-HPLC and data independent acquisition MS/MS

The peptide mixtures described above were also analyzed by LC-ESI-MS/MS on a Xevo™ quadrupole time-of-flight MS (Waters Corporation) equipped with a Universal NanoFlow Sprayer interface and pre-cut Pico Tip Emitter (360 μ m OD \times 20 μ m ID, 10 μ m tip; 2.5" long), connected on-line to a nanoACQUITY® UltraPerformance® HPLC system (Waters Corporation). The nanoACQUITY® system was equipped with Waters' 5 μ m Symmetry C₁₈, 180 μ m \times 20 mm reversed-phase trap and 1.7 μ m BEH130 C₁₈, 75 μ m \times 100 mm reversed-phase analytical columns. Both columns were maintained at 40°C. Typically 3 μ L of samples were injected onto the precolumn in aqueous 1% CH₃CN/0.1% FA at a flow rate of 5 μ L/min for 3 min. Mobile phase A was water with 0.1% FA (*v/v*) and mobile phase B was CH₃CN with 0.1% FA. After desalting and concentrating in the trap column, peptides were transferred to the analytical column and resolved by a gradient of 3–60% mobile phase B delivered over 30 min at a flow rate of 300 nL/min, followed by a 15 min wash with 95% B and 15 min re-equilibration at the initial conditions (3% B, 97% A).

The Xevo™ quadrupole time-of-flight MS was operated in positive ion, V-mode with an average resolution of 9500 FWHM. Full scan mass spectra were acquired from 50 to 2000 *m/z*. LC-MS and LC-MS^E data were collected in alternating low and high collision energy modes throughout the run (Silva et al., 2005), with each spectrum acquired for 1 s per mode.

Proteins were identified using the ProteinLynx Global SERVER™ version 2.4 search engine (PLGS, Waters Corporation). All ions were lock mass corrected, de-isotoped, and decombulated (charge state reduced). PLGS software ascribed collision induced dissociation (CID) product ions to their precursor peptides by time-aligning low- and high-energy-detected ions with a retention time tolerance of approximately ± 0.05 min. Sequence searches were restricted to fully tryptic products with up to one missed cleavage, variable methionine oxidation and N-terminal Gln/Glu conversion to pyro-Glu, and peptide and product ion tolerances of 10 and 25 ppm, respectively. Proteins hits were accepted based on ≥ 2 ascribed peptides, each with 3 or more product ions, and at least seven fragment ions per protein. Correspondences between MS/MS spectra and ascribed sequences were also evaluated manually.

RESULTS AND DISCUSSION

We identified 154 proteins from concanavalin A pull down fractions and cell surface labeling. The following sections describe these proteins and the additional information that can be recovered by data mining.

IDENTIFICATION OF CONCANAVALIN A INTERACTING PROTEINS

The archaeal cytoplasmic membrane has been described as fulfilling the role of the eukaryotic endoplasmic reticulum (Yurist-Doutsch et al., 2008) because archaeal glycosylation machinery is membrane-bound. Cell N-linked glycoproteins are expected to localize to the cytoplasmic membrane or the associated outer cell envelope region (Eichler, 2003; Albers et al., 2006; Messner, 2009). Thus, glycoprotein capture methods (e.g., lectin affinity chromatography) complement cell surface labeling (Francoleon et al., 2009) for enriching archaeal surface and membrane proteins.

To support ongoing studies characterizing protein glycosylation, *M. mazei* cell lysate proteins were lectin affinity-captured through direct and indirect binding using Con A, for which the subset of direct binders typically includes glycoproteins containing α -D-mannose and α -D-glucose (Kornfeld and Ferris, 1975; Baenziger and Fiete, 1979; Debray et al., 1981; Jaipuri et al., 2008). Here, we describe what can be learned from LC-MS/MS analyses of the Con A eluate with subsequent data mining. Further studies that employed additional dimensions of separation [e.g., hydrophilic interaction chromatography (HILIC) followed by reversed phase liquid chromatography], in concert with advanced ion activation techniques including infrared multiphoton dissociation (Zubarev, 2004; Cooper et al., 2005) will be discussed elsewhere (Leon et al., in preparation). Those methods greatly increase the numbers of glycoproteins observed, improve ability to localize glycosylation sites, and characterize individual glycan chains. Nevertheless, useful information can be gleaned from these initial LC-MS/MS experiments.

Not all proteins recovered from the Con A eluate are necessarily glycosylated. Indirect binding partners do not bind Con A directly, but associate with one or more direct (glycosylated) interactors; e.g., other subunits of a non-covalent complex. From *M. mazei* G₀, 99 Con A eluate proteins were identified by LC-MS/MS with ≥ 2 peptides (**Table 1**). An additional 55 protein identifications based on a single peptide are presented in Supplemental Table S-1.

Standard proteomic search strategies do not reveal unknown glycopeptides, because it is not possible to predict the amount by which peptide masses are incremented. However, manual MS/MS analysis of, e.g., HILIC fractions, reveals some precursor masses (peptides) that are obviously glycosylated, because they dissociate to release low mass-to-charge ratio (*m/z*) ions, known as oxonium ions, that are characteristic of different sugars (Mechref, 2012). Ions at 163.06 and 127.06 *m/z*; e.g., signal that hexoses are present, while 204.09, 186.09, and 168.09 *m/z* reflect present *N*-acetylhexosamines.

Bona fide glycosylated proteins, revealed by their oxonium ions in peptide MS/MS spectra, were identified from Con A eluate as MM0002, MM0716, MM1364, and S-layer protein MM1976. It is unlikely that these are the only glycoproteins within *M. mazei*.

Indeed, glycoprotein-specific staining (Francoleon et al., 2009) highlights many bands. Supplemental Table S-3 lists, from the proteins observed, those first predicted by SignalP 3.0 to be secreted and then by the NetNGlyc server (Blom et al., 2004) to be potentially glycosylated. The table lists 41 candidate glycoproteins. Interestingly, protein MM0002 was not predicted by SignalP to contain a signal peptide, although it is a candidate for leaderless secretion by SecretomeP, and NetNGlyc does suggest 2 Asn sites as potentially glycosylated. Hence, the list of potentially *N*-glycosylated proteins can be longer, should proteins secreted leaderlessly also be considered. The predictions from SignalP version 4.0 comprise a subset of the version 3.0 predictions. All but two of the proteins absent from the newer algorithm's list of signal peptide-containing sequences (MM1329 and MM2033) are also candidates for leaderless secretion (according to the SecretomeP algorithm), and thus potentially *N*-glycosylated. Additional glycoproteins likely await discovery.

It should also be clarified that the tandem MS conditions and protein quantities required for high-throughput peptide identification differ from those employed in oligosaccharide and glycoprotein analysis. Glycoprotein analyses will employ 10–100 fold more material and the MS/MS conditions will be customized for each analyte.

The utility of data mining is that, by employing a relatively simple enrichment method without exhaustive chromatographic purifications, these four proteins (MM0002, MM0716, MM1364, and MM1976) were revealed as glycosylated and bearing at least hexose and *N*-acetylhexosamine saccharides. It also indicates how glycopeptide knowledge can be extracted from tandem mass spectrometry data, even without upstream enrichment, bonus knowledge when experiments are not specifically targeting glycosylation. Adding chromatographic dimensions will increase the number of glycopeptides detected, because glycopeptide intensities are often suppressed by co-eluting non-glycopeptides, and because the chromatographic conditions that best resolve different glycopeptides differ from those best resolving peptides, in general. Analyses probing more *M. mazei* glycopeptides and at greater depth are underway.

PREDICTION OF ARCHAEL PROTEIN *N*-TERMINI

Signal peptides are key participants in protein translocation, but our ability to predict them confidently, especially for archaeal species, has been limited by availability of experimental data (Armengaud, 2009). Even among bacteria, e.g., *Mycobacterium smegmatis*, predictions were found to be erroneous for up to 19% of protein *N*-termini (Gallien et al., 2009). Complications delineating open reading frames (ORFs) arise when multiple AUG codons lie near the DNA 5' end (ambiguity that often prompts annotators to select the longest open reading frame) (Prats et al., 1992; Kozak, 1997; Meinnel and Giglione, 2008; Fournier et al., 2012), when organisms employ alternative initiation codons; e.g., GUG or UUG (Klunker et al., 2003; Falb et al., 2006; Yamazaki et al., 2006; Meinnel and Giglione, 2008; Running and Reilly, 2009; Elzanowski and Ostell, 2013), or when translation initiates at multiple sites (Prats et al., 1992). Modifications to *N*-termini are rarely predictable, and their prevalence varies with organism. Strikingly, 14–19% of protein *N*-termini in

Table 1 | proteins detected by two or more peptides from concanavalin A eluate.

Uniprot accession	MM#	Name	Unique peptides observed	Mascot protein score	Exprot ^a	SignalP ^b	SecP ^c	LipoP ^d
Q8Q0X9	MM0001	Dipeptide ABC transporter	5	229	1		Y	
Q8Q0X8	MM0002	Dipeptide ABC transporter	20	555	1		Y	
Q8Q0S0	MM0066	Hypothetical protein	5	216	2	Y	Y	
Q8Q0K8	MM0128	Lon protease, membrane-bound	8	276				
Q8Q0F2	MM0184	Rps3Ae	2	51				
Q8Q035	MM0303	Hypothetical protein	2	39	1		Y	
Q8PZM5	MM0467	S-layer paralog	17	606		Y	Y	
Q8PZ66	MM0628	Methylenetetrahydromethanopterin reductase	6	276				
Q8PZ61	MM0633	Multiheme cytochrome c, hypothetical	2	90			Y	
Q8PZ47	MM0647	Oligosaccharyl transferase	2	91			Y	
Q8PYZ8	MM0700	HppA	3	39		Y*	Y	
Q8PYY5	MM0714	Fructose bis phosphate aldolase	2	29				
Q8PYY3	MM0716	Hypothetical protein	3	105	1	Y	Y	
Q60187	MM0779	AtpB	6	189				
Q60186	MM0780	AtpA	11	368				
Q60184	MM0782	AtpC	4	154				
Q60183	MM0783	AtpE	4	34,13				
F1SVJ4	MM0784	AtpK	5	365	1	Y*	Y	
O59659	MM0785	AtpI	4	115			Y	
Q8PYS4	MM0786	AtpH	2	75				
Q8PYJ5	MM0866	Periplasmic serine protease	4	242	1		Y	
Q8PYIO	MM0882	Hypothetical protein	3	106			Y	
Q8PYE0	MM0922	Eif5a, IF5A	2	131				
Q8PY73	MM0991	Thioredoxin	5	150	2	Y		Y
Q8PY52	MM1012	Rpl1P	3	62				
Q8PY39	MM1025	ThiC	3	58				
Q8PXZ6	MM1070	MtaA1 methylcobalamin:CoM methyltransferase	7	374				
Q8PXZ5	MM1071	4Fe:4S ferredoxin, hypothetical	2	121				
Q8PXZ3	MM1073	MtaC2 methyl corrinoid protein	6	230				
Q8PXZ2	MM1074	MtaB2	9	250				
Q8PXZ1	MM1075	Putative regulatory protein	2	92	1		Y	
Q8PXX0	MM1096	Thermosome, gamma subunit	5	137				
Q8PXW0	MM1106	Putative phosphoserine phosphatase	5	81				
Q8PXV8	MM1108	Mtd, F420-dependent methylenetetrahydromethanopterin dehydrogenase	3	126				
Q8PXi2	MM1236	HtpX protease	2	25		Y	Y	
Q8PXH8	MM1240	Methyl CoM reductase, alpha subunit	7	237				
Q8PXH7	MM1241	Methyl CoM reductase, gamma subunit	10	563				
Q8PXH4	MM1244	Methyl CoM reductase, beta subunit	13	467				
Q8PXB8	MM1250	Cation transporter	2	66		Y*	Y	
Q8PXB7	MM1251	Cation transporter	2	109		Y*	Y	
Q8PXB9	MM1333	Zinc ABC transporter	10	486	2	Y		Y
Q8PXB2	MM1340	Pyruvate synthase, alpha subunit	2	61				
Q8PXB0	MM1362	Putative aliphatic sulfonate binding	11	556	2	Y	Y	
Q8PXB8	MM1364	S-layer paralog	28	1093	1	Y	Y	
Q8PXB3	MM1379	Thermosome, alpha subunit	10	329				
Q8PWZ8	MM1424	SecF	2	85	1	Y	Y	
Q8PWZ7	MM1425	SedD	3	144		Y	Y	
Q8PWX0	MM1456	USP-like	4	153				
P80650	MM1540	MtrH	9	276				
P80656	MM1541	MtrG	4	192			Y	
P80654	MM1542	MtrF	2	105			Y	
O59640	MM1543	MtrA	6	309				

(Continued)

Table 1 | Continued

Uniprot accession	MM#	Name	Unique peptides observed	Mascot protein score	Exprot ^a	SignalP ^b	SecP ^c	LipoP ^d
P80655	MM1544	MtrB	4	222			Y	
O59638	MM1545	MtrC	5	127			Y	
P80653	MM1546	MtrD	2	133		Y	Y	
P80651	MM1547	MtrE	4	117		Y	Y	
Q8PWN5	MM1549	Na/proline symporter	4	211		Y*	Y	
Q8PWN3	MM1551	Hypothetical protein	4	75	1	Y	Y	
Q8PWM3	MM1561	ABC transporter, tungsten binding protein	2	72		Y*	Y	
Q8PWJ6	MM1589	Surface layer protein B	3	128	1	Y	Y	
Q8PWE2	MM1647	MtaB1	14	357				
Q8PWE1	MM1648	MtaC1	10	533				
Q8PWA1	MM1695	Hypothetical protein	10	425	2	Y	Y	Y
Q8PW83	MM1713	Hypothetical protein	2	173			Y	
Q8PW41	MM1760	Rps2p 30S ribosomal protein S2	3	104			Y	
Q8PW15	MM1789	Hypothetical protein	2	64			Y	
Q8PVY9	MM1816	Conserved protein	6	238	1	Y	Y	
Q8PVW4	MM1843	HdrE, heterodisulfide reductase	6	210		Y	Y	
Q8PVV1	MM1859	DdpA, ABC transporter	3	29	1	Y	Y	
Q8PMV4	MM1939	Glutamine binding protein	2	115	2	Y		Y
Q8PVI7	MM1976	S-layer protein, (identified)	151	5310	1	Y	Y	
Q8PVI6	MM1977	Hypothetical protein	3	118	2	Y	Y	Y
Q8PVG6	MM1999	Hypothetical protein	4	103	2	Y	Y	Y
Q8PVG5	MM2000	Hypothetical protein	4	82	2	Y	Y	Y
Q8PVD2	MM2033	Stomatin-like protein	2	52		Y*		
Q8PVA2	MM2069	Iron III dicitrate binding protein	5	179	2	Y	Y	Y
Q8PV27	MM2147	SecY	2	67			Y	
Q8PV17	MM2157	Rpsllp, 30S ribosomal protein S11p	2	81			Y	
Q50227	MM2171	Cytochrome b, F420 non-reducing hydrogenase II	2	87			Y	
Q8PUU5	MM2234	Hypothetical protein	2	94	1	Y	Y	
Q8PUR8/	MM2264	Ef1A	10	301				
AAM31960								
Q8PUM8	MM2305	Na/proline symporter	2	51		Y	Y	
Q8PUL4	MM2320	EchA	3	75		Y*	Y	
Q8PU81	MM2460	Dipeptide oligopeptide binding protein	16	556	2	Y	Y	Y
F1SVE1	MM2479	FpoO, F420H2 dehydrogenase	2	83				
F1SVH9	MM2481	FpoM, F420H2 dehydrogenase	2	63		Y	Y	
F1SVK0	MM2482	FpoL, F420H2 dehydrogenase	2	91		Y	Y	
F1SVE0	MM2483	FpoK, F420H2 dehydrogenase	2	64		Y	Y	
Q8PU59	MM2487	FpoH, F420H2 dehydrogenase	3	97		Y*	Y	
P27094	MM2505	DnaK, chaperone protein	13	451				
Q8PU26	MM2531	Hypothetical protein	3	113	1			Y
Q8PTZ0	MM2567	ABC transporter, periplasmic binding protein	14	420	2	Y		
Q8PTY1	MM2576	Ferrous iron transport protein B	8	276			Y	
Q8PTS5	MM2637	Ion channel transporter	3	58			Y	
Q8PTQ7	MM2656	Peptidyl-prolyl cis-trans isomerase	2	28				
Q8PTI7	MM2728	Hypothetical protein	2	61	1	Y*	Y	
Q8PT25	MM2893	Hypothetical protein	6	193	1	Y	Y	
Q8PSQ2	MM3024	Hypothetical protein	3	65	1	Y	Y	
Q8PS84	MM3198	Lipoprotein, hypothetical	5	134				

^aExprot (Saleh et al., 2010); 1, Predicted type I signal peptidase substrate; 2, Predicted type II signal peptidase substrate.^bSignalP (Bendtsen et al., 2004b); Y, Predicted signal peptide by SignalP 3.0.^cSecP (Bendtsen et al., 2004a, 2005); Y, Predicted substrate for leaderless secretion by SecretomeP 2.0.^dLipoP (Juncker et al., 2003); Y, Predicted lipoprotein.^{*}Predicted as secreted only by the SignalP eukaryotic predictor.

halophiles *Halobacterium salinarum* and *Natronomonas pharaonis* are acetylated (Fallb et al., 2006). Clearly, experimentally characterizing the *N*-termini of microbial proteins is important.

In a computational approach, SignalP 3.0 (Bendtsen et al., 2004b), LipoP (Juncker et al., 2003), and SecretomeP 2.0 (Bendtsen et al., 2004a, 2005) algorithms were employed to predict secreted proteins, while Exprot predictions for *M. mazei* Gö1 were obtained from Supplemental information in Saleh et al. (2010). SignalP used gram-positive, gram-negative, and eukaryotic-type models, while SecretomeP was applied employing bacterial models (Bendtsen et al., 2004a, 2005). Predictions for proteins recovered by Con A binding are displayed in Supplemental Table S-2. Predicted and experimental results are compared in the following section and indicate a tendency of prediction algorithms to over-predict signal peptide excision.

EXPERIMENTALLY DETECTED *N*- AND *C*-TERMINI

Clearly, experimental approaches tailored to recovering as many *N*-terminal peptides as possible (Ogorzalek Loo et al., 2002; Gevaert et al., 2003; Dormeyer et al., 2007; Shen et al., 2007; Russo et al., 2008; Yamaguchi et al., 2008; Gallien et al., 2009; Xu and Jaffrey, 2010; Fournier et al., 2012; Kim et al., 2013; Venne et al., 2013) provide large datasets for evaluating and enhancing prediction algorithms and improving protein database annotations. Nevertheless, datasets acquired by other experimental approaches with different goals can be harvested to yield equivalent information for a smaller number of proteins, some of which are missed by large-scale “terminalomics” studies, especially because many large-scale approaches recover only free amino termini. Data harvests are also more likely to reveal instances where multiple *N*-terminal forms are present (e.g., modified and unmodified).

Information about protein *N*-termini is not automatically returned by database searching algorithms. Although most algorithms now consider both excised and retained initiator methionines when attempting to match MS/MS spectra, *N*-terminal acetylation is only considered if specified in the search parameters. Because each variable modification (i.e., one which may be present or absent) that must be considered in the search process adds to the time required for completion and often reduces specificity, only abundant variable modifications are usually considered by high-throughput proteomics studies.

From our Con A eluate studies, some LC-MS/MS spectra spanned *M. mazei* protein *N*- or *C*-termini, allowing us to compile that information in **Table 2**. *M. acetivorans* C2A and *M. mazei* Gö1 *N*- and *C*-termini information obtained previously (Francoleon et al., 2009), are also included. Information was recovered by semi-trypic and error-tolerant searches. Semi-trypic searches seek matches for MS/MS spectra to peptides in which only one terminus matches trypsin’s known cleavage specificity. Examples may include (i) peptides with non-Lys or Arg *C*-termini, but with *N*-termini reflecting cleavage after Lys or Arg, or (ii) peptide *N*-termini inconsistent with cleavage after Lys or Arg, but with *C*-terminal Lys or Arg. Error-tolerant searches, described in the Introduction, consider a large range of potential modifications.

Of *M. mazei* protein *N*-termini recovered by concanavalin A binding, only *S*-layer protein MM1976 underwent signal peptide

excision. These experimental results contrasted with those from prediction algorithms SignalP 3.0 (Bendtsen et al., 2004b) and Exprot (Saleh et al., 2010), which were found to over-predict the presence of signal peptides. Although SignalP 3.0 correctly predicted the MM1976 signal peptide, it also predicted leaders for 7 other proteins, while Exprot predicted signal peptides for 6 other proteins (Supplemental Table S-2). The newer algorithm SignalP 4.0 predicted leaders for only MM1362 and MM1547, in addition to MM1976. The poor correlation between prediction and experiment underscores previous conclusions about major problems predicting proteins lacking signal peptides (Antelmann et al., 2001).

Interestingly, SecretomeP 2.0, a machine-learning approach developed to predict non-classically secreted proteins in mammals and bacteria; i.e., proteins exported without a classical *N*-terminal signal peptide, predicted leaderless secretion of 14 proteins in Supplemental Table S-2 (Bendtsen et al., 2004a). Overlap for 8 of these SecretomeP 2.0 predictions with those by SignalP 3.0 is reasonable, because SecretomeP was trained on datasets of secreted protein sequences that had their signal peptides deleted. Detection of the other six proteins (MM0866, MM1009, MM1075, MM1221, MM1542, and MM1362) by our study would seem to verify these SecretomeP predictions.

M. mazei *N*-terminal peptides were recovered from 28 proteins over the course of this and previous studies. These protein identifications were supported by MS/MS spectra from multiple peptides in all but 3 cases. (Observing multiple peptides from a protein is one criterion for assessing confidence in the protein’s identification.) **Table 2** also includes data for 3 *M. acetivorans* *N*-termini. Heterogeneous *N*-termini were observed from 4 of the 31 proteins; e.g., Acetyl-VDAASTGLFLDAAGMK and Acetyl-MVDAASTGLFLDAAGMK were observed from A₁A₀ H⁺ ATPase subunit K (MM0784), indicating partial methionine excision prior to acetylation. MM0784 matched predictions in all other respects; the 4 tryptic peptides observed verified 100% of the 8-kDa proteolipid’s sequence. Strong *b*₁ ions (i.e., peptide fragments corresponding to protonated Ac-Val - H₂O or protonated Ac-Met - H₂O (*m/z* 142.09 and 174.06, respectively) in the MS/MS spectra confirmed the modification as *N*-acetylation (Yalcin et al., 1995). Because *b*₁ ions are generally observed in MS/MS spectra of *N*-terminally acetylated peptides, but otherwise relatively rare, their presence provides powerful validation of the modified peptide that is independent of the statistical score; i.e., search algorithms do not accord special significance to *b*₁ ions.

Partial methionine processing has been noted in other archaeal species (Fallb et al., 2006). That the *N*-terminus of the A₁ ATPase proteolipid subunit was found to be blocked was not surprising; blocked *N*-termini have frequently been observed for F₀ proteolipids; e.g., *N*-formylmethionine in bacteria (*E. coli* and *Bacillus*, UniProt Accessions P68699 and P00845, respectively), yeast mitochondria (Sebald et al., 1979), and wheat (Howe et al., 1982) and spinach (P69447) chloroplasts.

Putative regulatory protein MM1075 (MtaR) was detected as M SENAGTSTVIVDK (where M.S denotes methionine excision) and a formylated version, M SENAGTSTVIVDK, modified at S7 or T8. MS/MS spectra revealed ions *y*₁-*y*₅, and *y*₇-*y*₁₁, localizing the modification to S7 or T8. (See **Figure 1**). Technically, T → E,

Table 2 | *M. mazei* and *M. acetivorans* protein termini recovered.

Accession	MM#	Protein	N-Termini	C-Termini
Q8Q0X9	MM0001	Dipeptide ABC transporter		A.DIFGSILEWKRA.-
Q8Q024	MM0314	Valyl tRNA synthase		GAVVIVR.-
Q8PZY5	MM0353	Hypothetical protein		E.MILAYLKGN.-
Q8PZ47	MM0647	Oligosaccharyl transferase		E.EAVLNGQEIKI.-
F1SVJ4	MM0784	AtpK, A ₁ A ₀ H ⁺ ATPase subunit	Ac-VDAASTGLFLDAAGMK.A and Ac-MVDAASTGLFLDAAGMK.A	K.GLILTVPETIVIFGLVALLINQ.-
Q8PY54	MM0786	AtpH		E.FERAVLNE.-
Q8PY55	MM0866	Periplasmic serine protease	-.MNDENVNSDGVDSK.N	
Q8PY77	MM0987	Conserved protein		R.AGQFNNASAENTCAEHEHHH.-
Q8PY73	MM0991	Thioredoxin		R.AILYHENVNK.-
Q8PY70	MM0994	Precorrin-8X mutase		E.LVIAIRESEE.-
Q8PY55	MM1009	SecE	M.VESALEPNITTK.S	
Q8PY39	MM1025	ThiC		K.IVNQNLYNLAK.-
Q8PXZ6	MM1070	MtaA1		K.ALVAARDEFYA.-
Q8PXZ3	MM1073	MtaC2 methyl corrinoid protein	-.MLDLKLE.D	K.IADAIVAGTTDIAALREEFHKH.-
Q8PXZ2	MM1074	MtaB2	M.VKYTSMAYANADE.L	S.DNLTKFQEVAVFKPENYGL.-
Q8PXZ1	MM1075	Putative regulatory protein	M.SENAGTSTVIVDK.T	
Q8PXZ1	MM1075	Putative regulatory protein	M.SENAGTSTVIVDK.T + Formyl or T8→E8	
Q8PXX0	MM1096	Thermosome, gamma	M.AAQPIFILRE.G	
Q8PXV8	MM1108	Mtd, F420-dependent	M.VNIGFIK.M	
Q8PXJ7	MM1221	Mc1 chromosomal protein	M.SNTRNFVLRDEEGNE.H	
Q8PXH8	MM1240	Methyl CoM reductase, alpha subunit	Ac-AADIFSK.F	E.RSLIIPAK.-
Q8PXH4	MM1244	Methyl CoM reductase, beta subunit	M.SDTVDIYDDRGK.L	G.DVFGSIDEFRQPIKAVAGAL.-
Q8PXG8	MM1250	Cation transporter		R.LGLPLGLVAMTIYFIVLFILKI.-
Q8PXG7	MM1251	Cation transporter	M.ALEPGILAGFLVIFLAVLLGPFK.I	R.LGVPLGLVAMAIYFVVIFVLGI.-
Q80X89	MM1333	Zinc ABC transporter		E.NMDSVSDIFARNLV.-
Q8PX60	MM1362	Putative aliphatic sulfonate binding	M.MAGNLDVAYVGAAPAITALSQGLDAK.I	
Q8PX43	MM1379	Thermosome, alpha	M.AGQPIFILKE.G	
Q8PWZ8	MM1424	SecF	M.ATGLTEFLDNFVK.S	
Q8PWW2	MM1464	Hypothetical protein		E.LGYTKGIQTI.-
P80650	MM1540	MtrH	-.MFKFDKKQEVFE.L	
P80650	MM1540	MtrH	Supposed internal peptide 232–239 observed after trypsin M.ASAWDWLR.K	
P80654	MM1542	MtrF	M.AEEHEKGVPMLAPQMGAI.D (Asp-N product)	
O59640	MM1543	MtrA	M.ADKREPAPGWPLKGEYEVG.D	K.VEGAMIGLTITISLLGLLLLGR.-
P80655	MM1544	MtrB	M.SIVRIAPE.I	
O59638	MM1545	MtrC		R.DAYAVVGTGLLPSAEELQ.-
P80653	MM1546	MtrD		K.AVISSVVASILCAIVAVIAQLGGI.-
P80651	MM1547	MtrE	-.MEPLIGMGVLAL.I	K.AYGPYKEDKAEEASA.-
Q8PWE2	MM1647	MtaB1	M.AAATRFTKMYASADE.M	K.VFLPENYGF.-
Q8PWE1	MM1648	MtaC1	.MLDFTEASLKK.V	
Q8PVI7	MM1976	S-layer protein	(signal pept).ADVIEIR.G	
Q8PVI7	MM1976	S-layer protein	.ADVIEIR.G + 14 Da	
Q8PVI7	MM1976	S-layer protein	.XADVIEIR.G X = 40 Da	
Q8PVI7	MM1976	S-layer protein	.XADVIEIR.G X = 42 Da	
Q50226* or	MM2170* or	F420 non-reducing	M.VNVTVDPPLTR	
Q8PUM1*	MM2313*	hydrogenase		
Q8PUR8	MM2264	Ef1A	M.AADKPHMNLAVIGHIDHGK.S	

(Continued)

Table 2 | Continued

Accession	MM#	Protein	N-Termini	C-Termini
Q8PUL4	MM2320	EchA		K.LTTLGYAISIILVIALAGGVL.-
Q8PU74	MM2467	Rpl7ae	M.AQLAKFDVPEELTNKALE.A	
F1SVH9	MM2481	FpoM	.MLPVASLLILVPLIFAVVTFFTK.T	
F1SVK0	MM2482	FpoL		K.IQTGVVQTYATALIAGVSLLILVKLIMEVL.-
F1SVE0	MM2483	FpoK	M.TAIPLTFYLGAAALLFSIGLYGVMTHK.S	L.DELNILRW.-
Q8PTZ0	MM2567	ABC transporter, periplasmic binding protein		E.DPWLNLSGVYLERK.-
Q9C4S2	MA0456	Co-methyl transferase	-.MLDFTEASLKK.V	
Q9C4S2	MA0456	Co-methyl transferase	Ac-MLDFTEASLKK.V	
Q8TSG7	MA0829	S-layer protein	(signal pept.).VDVIEIR.G	
Q8TSG7	MA0829	S-layer protein		
Q8TRS7	MA1092	Ribosomal protein S5	.XVDVIEIR.G X=42 Da	R.LPVYYGKEEV.-
P58868	MA2971	Corrinoid protein	M.TNTEIFEK.L	K.VALNIMK.- ox
Q8TKI9	MA3416	Ribosomal protein S3Ae		R.KTEVETIPAAPAPAAA.-
Q8TIE5	MA4207	Ammonium transporter		R.AYPEYLMREE.-
Q8THH0	MA4547	Methyl CoM reductase, gamma subunit		R.TSYGFQPK.-

*Several LC-MS/MS experiments were unable to differentiate between the two proteins; none of the 6 peptides identified was unique.

Red entries correspond to identifications based on a single peptide.

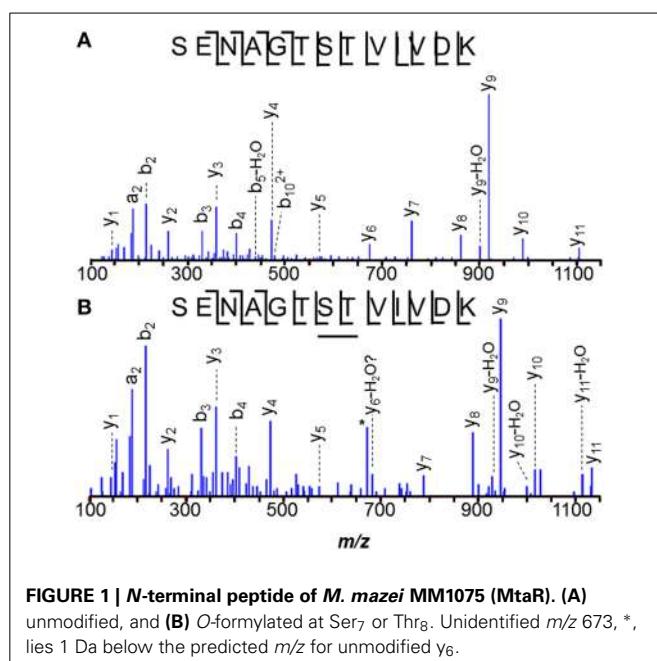


FIGURE 1 | N-terminal peptide of *M. mazei* MM1075 (MtaR). (A) unmodified, and **(B)** O-formylated at Ser₇ or Thr₈. Unidentified m/z 673, *, lies 1 Da below the predicted m/z for unmodified y_6 .

or S→D substitutions or C₂H₄ addition would also match the incremented mass, but the multiple base substitutions required to convert a Thr (ACC codon) to Glu (GAA/GAG) or a Ser (UCU) to Asp (GAU/GAC) are not easily reconciled, leading us to favor interpretation as formylation. High mass accuracy measurements can distinguish addition of CO (formylation) from C₂H₄ (27.995 vs. 28.031 Da).

O-formylation may be important to controlling MM1075's function, adapting cells to acetate- and methanol-dependent growth. Its mRNA levels are 200–500 times higher in methanol-

vs. trimethylamine-grown cells, and still higher for acetate culture (Hoey et al., 2005; Krätzer et al., 2009), while the operon's other genes, MM1073 and MM1074, are among the most highly regulated genes known in the Archaea (Bose et al., 2006). It will be interesting to monitor modifications of MM1075 as culture conditions are varied. For example, evidence that the ratio of formylated:unformylated MM1075 varies with; e.g., substrate or with length of time since the substrate was switched, would support a role in adaptation for the modification.

SURFACE LAYER PROTEIN MODIFICATIONS

The *M. mazei* sheath or S-layer protein, MM1976, is one of the most abundant proteins made by the cell (Francoleon et al., 2009; Rohlin et al., 2012). Con A binding enriched cell lysates for this protein, permitting characterization of low stoichiometry modifications. N-termini for the S-layer protein were especially varied, although all were consistent with signal peptide cleavage after residue 24. By abundance, the major N-terminal peptide observed was ADVIEIR, although peptides 14, 40, and 42 Da heavier were also found. The modified peptides followed ADVIEIR in elution by <1, 8, and 3 min, respectively.

N-terminal addition of 42 Da, localized by y_1 - y_6 and b_2 ions, (see Francoleon et al., 2009, Figures 3C,D) was initially attributed to α -amino acetylation. However, none of the tandem mass spectra acquired for this precursor yielded a b_1 -ion, generally considered diagnostic of N-terminal acetylation (Yalcin et al., 1995). Careful mass measurements on b_2 product ions better matched modifications of composition C₂H₂O, rather than C₃H₆. The *M. acetivorans* ortholog, MA0829, also displayed evidence of signal peptide cleavage, yielding the N-terminal tryptic peptides VDVIEIR and a +42 Da variant, similarly lacking its b_1 product ion by MS/MS (Francoleon et al., 2009 see Francoleon et al., 2009, Figures 3A,B). Further investigation is required to

confidently ascribe *N*-terminal modifications for these low abundance variants.

MS/MS spectra suggest that the *M. mazei* +40 Da modification also localizes to the *N*-terminus, as y_1 - y_6 , a_2 , and b_2 - b_5 ions were observed. Again, the b_1 -ion was not seen. The late elution of the +40 Da species relative to unmodified ADVIEIR, leads us to questions whether it might arise from in-source collision-induced dissociation (CID) of an exceptionally labile, hydrophobic group, leaving behind only a residual $-N = CH-CH = O$, $-N = C(CH_3)_2$, or $-N = CH-CH_2-CH_3$ *N*-terminus. Available mass accuracy narrowed consideration to the latter two possibilities. As an alternative to production by CID, the C_3H_4 increment could correspond to a Schiff base formed by addition of propionaldehyde, although the mechanism for such a modification is unclear. Further effort is required to characterize this C_3H_4 modification.

In some MM1976 ADVIEIR peptides, Glu₅ was incremented by 14 Da, consistent with methyl esterification. Elsewhere, artifactual methyl adducts were attributed to incubation in acidic methanol during gel fixation or staining (Parker et al., 1998; Xing et al., 2008). Here the analyses displaying the modification were performed on proteins digested in solution; i.e., not subjected to staining. Their only exposure to methanol was shorter, at lower concentration and at lower temperature than conditions known to esterify. Thus, the 14 Da adducts are unlikely to be artifacts. Other instances of relevant methyl esters have been described previously (Hoelz et al., 2006).

OTHER MODIFIED PROTEINS

The predicted *N*-terminal peptide for MM1540, subunit *H* of tetrahydrosarcinopterin *S*-methyl transferase (MtrH) was confirmed to be MFKFDKKQE. Interestingly, peptide M. ASAHDWLR (residues 232–239) was also observed, potentially reflecting protein processing, anomalous cleavage, or alternate initiation, although we could find no rationalization for the latter possibility. That MM1540 is the catalytic subunit of the *S*-methyl transferase and binds methyl tetrahydrosarcinopterin (Hippler and Thauer, 1999) encourages speculation that an inadvertent methyl transfer to Met₂₃₁ instead of the coenzyme M thiol might lead to a sulfonium-activated cyanogen bromide-like cleavage at the observed position.

The *N*-terminal peptide of *M. acetivorans* C2A MA0456 (MtaC1, methanol-5-hydroxybenzimidazolyl cobamide co-methyl transferase) was previously recovered as MLDFTESLKR and in its methionine sulfoxide form. Methionines are often oxidized under experimental conditions. A related peptide 58-Da heavier than the unmodified species was also observed (Figure 2). Tandem MS of the latter peptide localized the +58-Da to the first residue, revealing intense 190- and 126-Da product ions, consistent with the b_1 ion from an *N*-terminally acetylated peptide and a corresponding 64-Da neutral loss product, unique to methionine sulfoxide. As discussed earlier, b_1 observations strongly suggest *N*-terminal acetylation. Larger b -ions also showed 64-Da neutral loss products, establishing the variant *M. acetivorans* peptide as Ac- M^{Ox} LDFTESLKK. Peptide Ac- M^{Ox} LDFTESLKK, was also observed. Previous mass analyses of intact *M. acetivorans* proteins reported only the

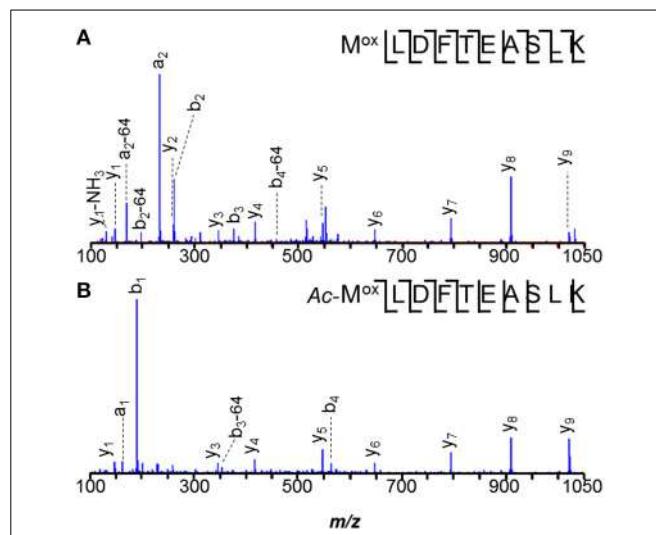


FIGURE 2 | *N*-terminal peptides of *M. acetivorans* MA0456 (MtaC1). (A) Unmodified MLDFTESLKR, **(B)** Low abundance variant Ac- M^{Ox} LDFTESLKK. Intense 190-Da ions correspond to the b_1 product from the *N*-terminally acetylated peptide. Larger b -ions show 64-Da neutral loss products characteristic of methionine sulfoxide. [Based on its measured mass, the product ion observed at m/z 129.11 is attributed to y_1-H_2O (m/z 129.10), rather than unmodified b_1 (m/z 129.01)].

free amino terminal form (Patrie et al., 2006). Interestingly, we observed only non-acetylated MLDFTESLKR from the *M. mazei* ortholog, MM1648, despite the larger protein quantity available for analysis. Note that *M. mazei* initiates translation 10 residues downstream of the originally annotated position (Deppenmeier et al., 2002).

Careful data mining revealed additional modifications to MA0456 (MtaC1). Three versions of peptide 149–159 were found: ANGYDVVDLGR, the Asn₂ deamidated peptide, and a peptide 42-Da heavier than predicted, with its modification localized to the first residue by y_{10} and b_2 ions (Figure 3). The absence of a b_1 -ion and accurate mass measurement ruled out N^{α} -acetylation for the heavier peptide, but could support an $A \rightarrow I/L$ substitution or N^{α} -trimethylation. However, amino acid substitution is hard to rationalize from the nucleotides coding Ala (GCA) vs. those coding Leu/Ile (CTX, TTA, TTG/ATT, ATC, ATA), as is sub-stoichiometric substitution. N^{α} -trimethylation would require prior cleavage of the protein to expose *N*-terminal ANGYDVVDLGR for modification. Interestingly, the peptide lies near His₁₃₆, an axial ligand to the Co²⁺ of the MtaC1 corrinoid co-factor that accepts CH₃ from methanol and subsequently transfers it to coenzyme M (Sauer et al., 1997; Randaccio et al., 2007). Neutral losses of 59-Da, sometimes observed from trimethylated residues, were not apparent in these spectra. Additional experiments are required to verify the source of this substitution; e.g., genetic drift (DNA), sloppy transcription (RNA), miscoding (translation), or post-translational modification, but at present trimethylation is favored.

Despite the larger protein amounts available for the two *M. mazei* orthologs, MM1648 (MtaC1) and MM1073

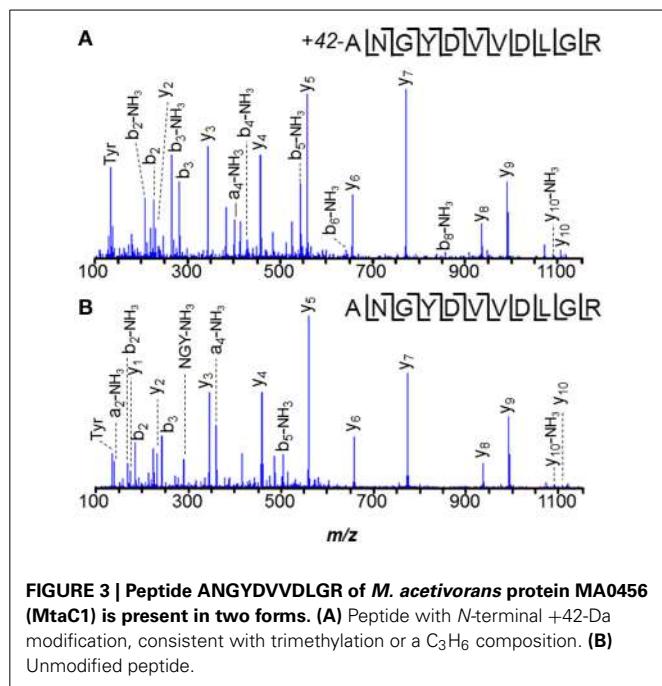


FIGURE 3 | Peptide ANGYDVVDLGR of *M. acetivorans* protein MA0456 (MtaC1) is present in two forms. (A) Peptide with *N*-terminal +42-Da modification, consistent with trimethylation or a C₃H₆ composition. **(B)** Unmodified peptide.

(MtaC2), 42-Da incremented peptides ANGYNVVDLGR and ANGYDVVDLGR, respectively, were not observed. Only unmodified and -17 Da variants were observed, with the latter reflecting succinimides, unremarkable for Asn-Gly bonds. Instead, a semi-tryptic peptide was observed in *M. mazei* MM1073 (MtaC2). Peptide 138–150, C*HVAEGDVHDIGK was incremented by 25-Da at its *N*-terminus, consistent with cyanylation see the MS/MS spectrum displayed in Figure 4. Such modification seems remarkable, but may reflect a radical-induced side reaction given that this region binds the corrinoid cofactor. A classic chemical cleavage scheme relies on *S*-cyanocysteine's base-catalyzed ability to cleave the *N*-terminal peptide bond to yield iminothiazolidinyl peptides (Jacobson et al., 1973; Degani and Patchornik, 1974; Nefsky and Bretscher, 1989; Wu and Watson, 1997). Uncleaved *S*-cyanocysteine 134–150 (GTVVC*HVAEGDVHDIGK) was not observed by LC-MS. At present, we cannot differentiate *in vivo* modification from modification induced by exposure to air during sample handling.

Limited experiments performed on *M. acetivorans* MA0456 (MtaC1), revealed only unmodified tryptic peptide GTVVCHVAEGDVHDIGK (124–140), but those studies cannot be considered conclusive, because smaller quantities of protein and a narrower range of sample handling conditions were pursued. In particular, the *M. acetivorans* samples were reduced with dithiothreitol (DTT), whereas cyanylated *M. mazei* protein was observed from untreated samples. Excess DTT removes the cyanide group from internal cysteines, thereby stopping the cleavage reaction (Degani and Patchornik, 1974; Nefsky and Bretscher, 1989).

PROTEIN COMPLEXES FOR ENERGY TRANSFER-ACQUISITION

Numerous protein complexes were recovered by Con A fractionation, including, tetrahydromethanopterin *S*-methyl

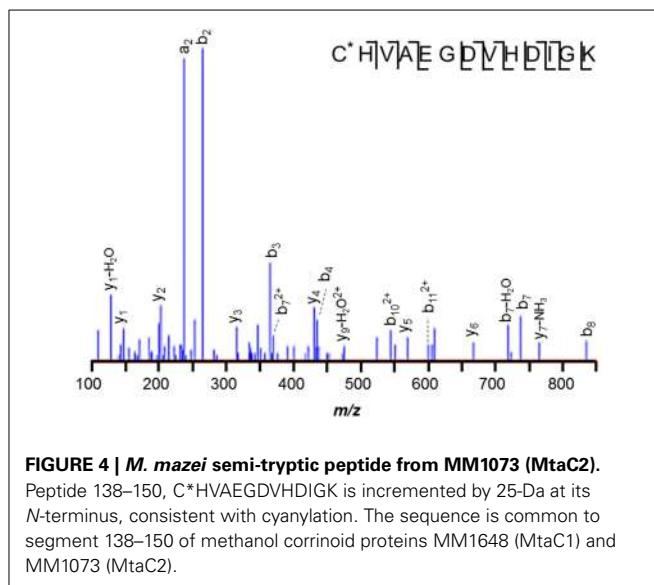


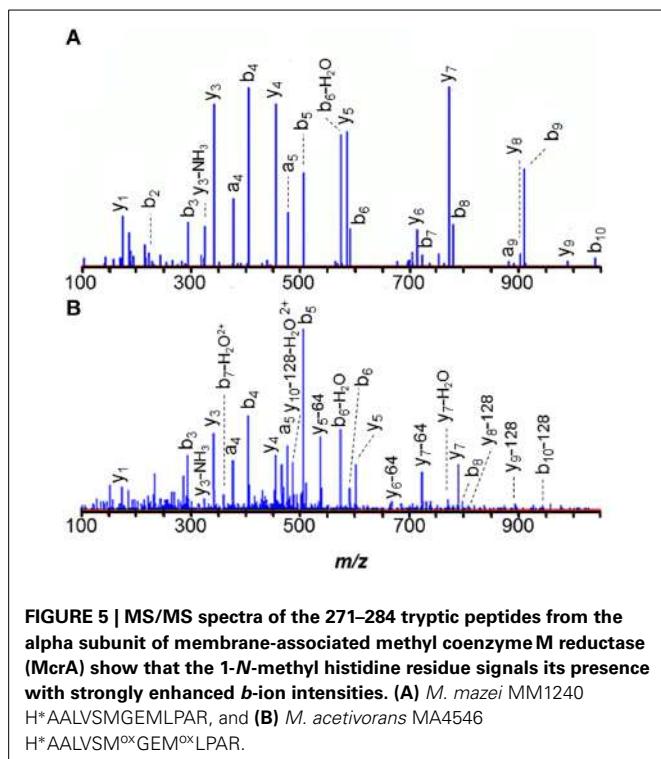
FIGURE 4 | *M. mazei* semi-tryptic peptide from MM1073 (MtaC2).

Peptide 138–150, C*HVAEGDVHDIGK is incremented by 25-Da at its *N*-terminus, consistent with cyanylation. The sequence is common to segment 138–150 of methanol corrinoid proteins MM1648 (MtaC1) and MM1073 (MtaC2).

transferase (*mtr*), methylcobalamin:CoM methyltransferase (*mta*), methyl CoM reductase (*mcr*), F₄₂₀H₂ dehydrogenase (*fpo*), heterodisulfide reductase (*hdr*), and A₁A₀ H⁺ ATPase (*aha*). Conceivably, Con A fractionation could show utility for enriching select *M. mazei* protein complexes in support of other studies or for additional characterization. Hence, we considered which subunits were identified in the eluate. A second reason for interest in the proteins and complexes detected is that mass spectrometrists often discount the presence of certain proteins in mixtures as reflective of contamination. These assumptions are not always justified. Here we sought evidence that proteins not annotated as surface- or membrane-localized, or not known to be glycosylated might rationally be carried along in the Con A fractionation by interactions with other proteins. Should most of the detected proteins be rationalized, there would be additional impetus to explore the cellular localization of any remaining proteins.

All eight tetramethyl methanopterin *S*-methyltransferase subunits (MtrABCDEFGHI) were detected, including integral membrane proteins MtrC, MtrD, and MtrE (Fischer et al., 1992; Lienard et al., 1996; Lienard and Gottschalk, 1998; Thauer, 1998; Kahnt et al., 2007). The abundance of this complex enabled *N*- and *C*-termini to be defined for many of the subunits, as well as providing some unexpected observations. Tryptic peptides IVTDEDKGIFDR (40–51) and IVTD~~X~~DK (40–46), with X corresponding to Glu-28 Da, were observed from catalytic subunit MtrH (MM1540). The latter peptide could reflect decarboxylation of Glu₄₄, perhaps from exposure to reactive species.

Strikingly, Con A binding recovered representatives from all types of the organism's membrane-bound hydrogenases involved in electron transfer: the F₄₂₀ non-reducing hydrogenase (Vht), Ech hydrogenase, and the F₄₂₀H₂ dehydrogenase (Fpo) (Kunkel et al., 1997, 1998). These included seven proteins encoded in the F₄₂₀H₂ dehydrogenase gene cluster MM2479-2491, the hydrogenase components EchA and EchB (MM2320, MM2321), and F₄₂₀-non-reducing hydrogenase MM2171 (VhtC). Previously, we



recovered a large subunit of the F₄₂₀H₂ dehydrogenase [MM2170 (VhtA) or MM2313] by biotin-tagging (Francoleon et al., 2009). Membrane-bound heterodisulfide reductase proteins HdrE and HdrD (MM1843 and MM1844) were also recovered, as was MM0628, provider of reduced F₄₂₀ (Bäumer et al., 2000; Thauer et al., 2008).

The methyl coenzyme M product (CH₃-S-CoM) is reductively demethylated by coenzyme B (CoB-SH) in a process catalyzed by the membrane-associated methyl coenzyme M reductase complex, Mcr (Hoppert and Mayer, 1990). All three α , β , and γ subunits (MM1240, McrA; MM1244, McrB; MM1241, McrC) were eluted from concanavalin A. From α -subunit McrA, 1-N-methyl-histidine was observed in peptide HAALVSMGEMLPAR (271–284), consistent with previous observations in *M. barkeri* (Grabarse et al., 2000). Because peptides spanning residues 285, 465, and 472 were not recovered, we could not determine if the unusual amino acids 5-methylarginine, S-methylcysteine, and thioglycine were present, as in other methanogens (Grabarse et al., 2000; Kahnt et al., 2007). However, the N-terminus was found to be acetylated. The peptide 271–284 MS/MS spectrum, illustrated in Figure 5A, demonstrates that the 1-N-methyl histidine residue enhances b-ion intensities from tryptic peptides. Peptides 271–284 and 1–7 were only observed in modified form. Similarly, 1-N-methyl histidine was found in the active site region of the *M. acetivorans* ortholog, MA4546. (See Figure 5B).

That protein products from 8 of 10 A₁A₀ ATPase-related genes were recovered via Con A elution, (*ahaABCDE*, *ahaHIK*) suggests that this capture method may be useful in isolating these unstable complexes (Müller et al., 1999). Previous experiments (Lemker et al., 2001, 2003) purifying ATPase subcomplexes from F₁F₀ ATPase-negative *E. coli* cells over-expressing the

M. mazei *ahaA-ahaG* operon recovered subunits A, B, C, D, and F. Questions persist regarding the participation in ATPase of *M. mazei* AhaG, considered authentic based on observation of an appropriately migrating SDS-PAGE band following heterologous expression in *Escherichia coli* (Lemker et al., 2001), but homologs of which are absent in several archaea. In our studies, recovery of AhaG may have been reduced, because the small (6.3-kDa) protein is expected to yield only two tryptic peptides larger than 900 Da, one of which is very hydrophobic. Although the mass spectrometer is capable of detecting tryptic peptides below 900 Da in mass, the peptides are often lost upstream in reversed phase HPLC, because their hydrophilicity causes them to elute with salt. Large hydrophobic peptides are also problematic for the chromatography because they fail to elute during the analysis. Thus, we cannot rule out AhaG as a component of the A₁A₀ ATPase.

PROTEIN COMPLEXES FOR METHANOL METABOLISM

Observed in these studies were products from two of the three operons coding methanol-specific methyl cobalamin:CoM methyltransferases: (1) MtaA1 (MM1070) with heterodimeric MtaB2/MtaC2 (MM1074/MM1073), and (2) MtaB1/MtaC1 (MM1647/MM1648). The MtaB/MtaC complexes transfer a methyl group from methanol to the corrinoid cofactor of MtaC. Subsequently, that methyl is transferred to coenzyme M (HS-CoM), catalyzed by MtaA1. Alternative enzymes catalyzing transfer to HS-CoM (MtaA2 and Mtba), were not observed. That Mtba was absent is unsurprising, as it is specific for growth on H₂/CO₂ or trimethylamine (Harms and Thauer, 1996; Hovey et al., 2005), and Ding et al. (2002) did not identify ortholog MtaA2 from methanol-cultivated *M. thermophila*. However, methanol-induced expression of MtaB3 and MtaC3 was established in *M. thermophila* (Ding et al., 2002), leading us to address their absence in our data. First, we would not expect these methyl transferases to bind concanavalin A or associate with Con A binders, because MtaB3 and MtaC3 are generally considered soluble. Also, several Mtab3 (MM0175) peptides are non-unique, (shared with other isozymes) complicating its identification from complex mixtures. DNA microarray analyses indicated that *mtaB1/mtaC1* were induced 10–33X in methanol, while *mtaB2/mtaC2* were induced only in acetate (Hovey et al., 2005). Quantifying roughly, by comparing numbers of peptides recovered, we see that the trend in protein abundances follows the same direction as the transcripts: 14 peptides vs. 9 for MtaB1 vs. MtaB2 and 10 peptides vs. 6 for MtaC1 vs. MtaC2.

ADDITIONAL PROTEINS RECOVERED BY CON A

MM0633, a hypothetical protein containing a multi-heme cytochrome c domain suggested as part of a membrane-bound complex, belongs to a gene cluster showing elevated expression under aceticlastic growth (Hovey et al., 2005). In these methylo trophic studies, however, MM0633 was the only cluster member observed. As it lacks transmembrane regions, we may wonder if its presence reflects interaction with some other membrane protein or glycosylation.

The oligosaccharyl transferase (MM0647) detected in the ConA pull down experiment (Table 1) is a product of one of three *aglB* homologs encoded in the *M. mazei* genome (MM646,

MM0647, MM2210) (Magidovich and Eichler, 2009). Its detection makes MM0647 a logical candidate for the AglB oligosaccharyl transferase that links glycans to asparagines on surface layer protein MM1976 and on other N-linked glycoproteins. Two minor S-layer proteins similar to MM1976, MM0467, and MM1364, were identified where the latter was also shown to be glycosylated (described below). Ongoing analyses of *M. mazei* N-linked glycans will reveal more about the oligosaccharides transferred.

Proteins with roles in cobalt and iron uptake were also observed: MM2069, an iron ABC transporter, MM1999 and MM2000 involved in cobalt uptake, MM0893 (CbiM), a cobalt ATP-dependent transporter, and MM0994 (CbiC). Identifications for CbiC and CbiM were based on a single recovered peptide for each, a lower standard of confidence. Numerous other transporters were also observed.

Previously detected Hsp70 analog MM2505 and the membrane-bound ATP-dependent protease LonB were also found in Con A eluate, along with two of three subunits comprising the *M. mazei* thermosome (Bateman et al., 2004), a eukaryotic-type chaperonin complex. Previously, surface-biotinylation with streptavidin affinity chromatography retrieved all 3 *M. mazei* subunits [MM1379 (α), MM0072 (β), and MM1096 (γ)], confirming the proposal (Trent et al., 2003) that a fraction of thermosome (or rosettosome) complexes are membrane-localized. In the present *M. mazei* lectin capture, as well as for our previous *M. acetivorans* surface-tagging and capture efforts (Francoleon et al., 2009), the thermosome γ -subunit was not recovered. Archaeal thermosomes vary in whether their double ring structures are composed of identical subunits, or of two or three different sequences; e.g., the *Methanopyrus kandleri* complex is homomeric (Andrä et al., 1998). Indeed, *M. mazei* proteins most closely related to the *M. kandleri* thermosome, MK1006, are MM1379, MM1096, and MM0072, respectively.

PROTEIN EXPORT AND PROCESSING

Machinery to transport proteins across the membrane is essential to protein secretion. Recent cryo-electron microscopy studies revealed important components of this machinery in yeast, where protein transport across the endoplasmic reticulum begins with the signal peptide of the nascent chain engaging the signal recognition particle (SRP) in the cytoplasm. Co-translational translocation is initiated when the signal peptide is transferred to the protein conducting channel, overlapping 4 binding sites on the large ribosomal subunit. The archaeal analog to translocation across the ER is transport across the cell membrane. The heterotrimeric protein-conducting channel (akin to yeast Sec61 α , β , and γ) consists of integral membrane proteins: MM2147, MM1372, and MM1009, respectively, all of which we observed (Becker et al., 2009; Kampmann and Blobel, 2009), along with accessory factors MM1424 (SecF) and MM1425 (SecD). As in Eukarya and Bacteria, ribosomes contact membranes via Sec-based sites (Ring and Eichler, 2004), consistent with our observations of associated ribosomal proteins. The SecP algorithm predicts secretion of ribosomal proteins MM1760, MM2124, MM2135, and MM2157 (Bendtsen et al., 2004a, 2005), although the relationship

between non-classical secretion predictions, disordered regions, and protein-protein or protein-nucleotide interactions is yet unclear. The mass spectrometry/proteomics community often cites the presence of ribosomal proteins in cell membrane preparations as evidence of poor quality, but it is important to consider that some presence in preparations enriching membrane-associated complexes is legitimate.

Of the total number of *M. mazei* open reading frames detected by our study, genome sequence analysis (Deppenmeier et al., 2002) annotated about 20% as hypothetical, thus highlighting the efficacy of the Con A pull-down approach for discovery. Of the 28 hypothetical proteins, 24 were predicted to be secreted by Exprot, SignalP, SecP, and/or LipoP (Juncker et al., 2003; Bendtsen et al., 2004a,b, 2005; Saleh et al., 2010). In addition, we find that hypothetical proteins MM0716 and MM1364 are glycosylated. Glycosylation of MM1364 correlates with homology to known *Methanosarcinae* S-layer proteins (Francoleon et al., 2009), that also bear glycans.

CONCLUSIONS

LC-MS/MS analyses of proteolytically-digested concanavalin A eluate from *M. mazei* Gö1 cell lysates led to the identification of 154 proteins. Among these, constituents of membrane-bound or membrane-associated complexes known from the literature were well-represented, including all 8 subunits of tetrahydromethanopterin S-methyl transferase (Mtr), seven proteins encoded by the F420H₂ dehydrogenase *fpo* operon, the 3 subunits of methyl coenzyme M reductase (Mcr), the protein products from 8 of 10 A₁A₀ ATPase-related genes (*ahaABCDE-HIK*), and components of the machinery translocating proteins across the cell membrane [protein channel constituents MM2147, MM1372, and MM1009 and accessory factors MM1424 (SecF) and MM1425 (SecD)]. All of these proteins do not bear Con A-interacting saccharides, because lectin binding is performed under non-denaturing conditions. However, the results can be useful in considering strategies to enrich or isolate select membrane complexes from *M. mazei*, and perhaps other *Methanosarcinae*, in order to monitor dynamic changes in protein modifications and/or retrieve complexes from strains not engineered to synthesize tagged proteins for easy retrieval.

Tandem mass spectrometry data associated with protein identifications can be mined to recover novel information that is not automatically provided by the high-throughput analyses. Here it was found that S-layer protein MM1976 was present in multiple forms, including four variants of its N-terminal peptide ADVIEIR. Instances of protein formylation, methyl esterification, methylation, and cyanylation were also found. Knowledge of unanticipated modifications, even if not providing immediate insight, does suggest features to monitor for evidence of dynamic changes. Knowledge gained by data mining can also complement what is obtained from experiments specifically targeting that modification, because the experimental conditions (e.g., chromatography resin and elution conditions) are often different. A high throughput LC-MS/MS run injects only a few hundred nanograms of a peptide mixture. Further effort is underway to characterize unknown glycans, the sites they modify, and

other post-translational modifications, particularly in extensively modified S-layer protein MM1976.

N-termini recovered from a subset of proteins secreted to the membrane or cell surface provide a dataset for comparison to signal peptide algorithms. Disagreement between the number of proteins predicted vs. detected with signal peptide-excised *N*-termini suggests that leaderless secretion is of greater importance than present models imply.

ACKNOWLEDGMENTS

This research was supported by the Department of Energy Biosciences Division grant award DE-FG02-08ER64689 to RPG, by the UCLA-DOE Institute of Genomics and Proteomics (DE-FC03-02ER6342) to JL and RPG, by the Ruth L. Kirschstein National Research Service Award (Grant GM007185) to DL and by the National Institutes of Health (R01 GM085402) to ROL and JL.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <http://www.frontiersin.org/journal/10.3389/fmicb.2015.00149/abstract>

REFERENCES

- Albers, S.-V., Szabó, Z., and Driessens, A. J. M. (2006). Protein secretion in the archaea: multiple paths towards a unique cell surface. *Nat. Rev. Microbiol.* 4, 537–547. doi: 10.1038/nrmicro1440
- Andrä, S., Frey, G., Jaenicke, R., and Stetter, K. O. (1998). The thermosome from *Methanopyrus kandleri* possesses an NH₄⁺-dependent atpase activity. *Eur. J. Biochem.* 255, 93–99. doi: 10.1046/j.1432-1327.1998.2550093.x
- Antelmann, H., Tjalsma, H., Voigt, B., Ohlmeier, S., Bron, S., Van Dijl, J. M., et al. (2001). A proteomic view on genome-based signal peptide predictions. *Genome Res.* 11, 1484–1502. doi: 10.1101/gr.182801
- Armengaud, J. (2009). A perfect genome annotation is within reach with the proteomics and genomics alliance. *Curr. Opin. Microbiol.* 12, 292–300. doi: 10.1016/j.mib.2009.03.005
- Baenziger, J. U., and Fiete, D. (1979). Structural determinants of concanavalin a specificity for oligosaccharides. *J. Biol. Chem.* 254, 2400–2407.
- Bandeira, N., Tsur, D., Frank, A., and Pevzner, P. A. (2007). Protein identification by spectral networks analysis. *Proc. Natl. Acad. Sci. U.S.A.* 104, 6140–6145. doi: 10.1073/pnas.0701130104
- Bateman, A., Coin, L., Durbin, R., Finn, R. D., Hollich, V., Griffiths-Jones, S., et al. (2004). The pfam protein families database. *Nucleic Acids Res.* 32, D138–D141. doi: 10.1093/nar/gkh121
- Bäumer, S., Ide, T., Jacobi, C., Johann, A., Gottschalk, G., and Deppenmeier, U. (2000). The F₄₂₀H₂ Dehydrogenase from *Methanosarcina mazei* is a Redox-Driven proton pump closely related to NADH Dehydrogenases. *J. Biol. Chem.* 275, 17968–17973. doi: 10.1074/jbc.M000650200
- Baumgartner, C., Rejtar, T., Kullolli, M., Akella, L. M., and Karger, B. L. (2008). SeMoP: a new computational strategy for the unrestricted search for modified peptides using LC-MS/MS data. *J. Proteome Res.* 7, 4199–4208. doi: 10.1021/pr800277y
- Becker, T., Bhushan, S., Jarasch, A., Armache, J.-P., Funes, S., Jossinet, F., et al. (2009). Structure of monomeric yeast and mammalian Sec61 complexes interacting with the translating ribosome. *Science* 326, 1369–1373. doi: 10.1126/science.1178535
- Bendtsen, J. D., Jensen, L. J., Blom, N., Von Heijne, G., and Brunak, S. (2004a). Feature based prediction of non-classical and leaderless protein secretion. *Protein Eng. Des. Sel.* 17, 349–356. doi: 10.1093/protein/gzh037
- Bendtsen, J. D., Kiemer, L., Fausbøll, A., and Brunak, S. (2005). Non-classical protein secretion in bacteria. *BMC Bioinformatics* 5:58. doi: 10.1186/1471-2180-5-58
- Bendtsen, J. D., Nielsen, H., Von Heijne, G., and Brunak, S. (2004b). Improved prediction of signal peptides: signalP 3.0. *J. Mol. Biol.* 340, 783–795. doi: 10.1016/j.jmb.2004.05.028
- Blom, N., Ponten, T. S., Gupta, R., Gammeltoft, S., and Brunak, S. (2004). Prediction of post-translational glycosylation and phosphorylation of proteins from the amino acid sequence. *Proteomics* 4, 1633–1649. doi: 10.1002/pmic.200300771
- Bose, A., Pritchett, M. A., Rother, M., and Metcalf, W. W. (2006). Differential regulation of the three methanol methyltransferase isozymes in *Methanosarcina acetivorans* C2A. *J. Bacteriol.* 188, 7274–7283. doi: 10.1128/JB.00535-06
- Cooper, H. J., Häkansson, K., and Marshall, A. G. (2005). The role of electron capture dissociation in biomolecular analysis. *Mass Spectrom. Rev.* 24, 201–222. doi: 10.1002/mas.20014
- Creasy, D. M., and Cottrell, J. C. (2004). Unimod: protein modifications for mass spectrometry. *Proteomics* 4, 1534–1536. doi: 10.1002/pmic.200300744
- Creasy, D. M., and Cottrell, J. S. (2002). Error tolerant searching of uninterpreted tandem mass spectrometry data. *Proteomics* 2, 1426–1434. doi: 10.1002/1615-9861(200210)2:10<1426::AID-PROT1426>3.0.CO;2-5
- Debray, H., Decout, D., Strecker, G., Spik, G., and Montreuil, J. (1981). Specificity of twelve lectins towards oligosaccharides and glycopeptides related to *n*-glycosylproteins. *Eur. J. Biochem.* 117, 41–55. doi: 10.1111/j.1432-1033.1981.tb06300.x
- Degani, Y., and Patchornik, A. (1974). Cyanylation of sulphydryl groups by 2-nitro-5-thiocyanobenzoic acid. High-yield modification and cleavage of peptides at cysteine residues. *Biochemistry* 13, 1–11. doi: 10.1021/bi00698a001
- Deppenmeier, U., Johann, A., Hartsch, T., Merkl, R., Schmitz, R. A., Martinez-Arias, R., et al. (2002). The genome of *Methanosarcina mazei*: evidence for lateral gene transfer between Bacteria and Archaea. *J. Mol. Microbiol. Biotechnol.* 4, 453–461.
- Ding, Y.-H. R., Zhang, S.-P., Tomb, J.-F., and Ferry, J. G. (2002). Genomic and proteomic analyses reveal multiple homologs of genes encoding enzymes of the methanol:coenzyme m methyltransferase system that are differentially expressed in methanol- and acetate-grown *Methanosarcina thermophila*. *FEMS Microbiol. Lett.* 215, 127–132. doi: 10.1111/j.1574-6968.2002.tb11381.x
- Dormeyer, W., Mohammed, S., Van Breukelen, B., Krijgsveld, J., and Heck, A. J. R. (2007). Targeted analysis of protein termini. *J. Proteome Res.* 6, 4364–4465. doi: 10.1021/pr070375k
- Eichler, J. (2003). Facing extremes: archaeal surface-layer (glyco)proteins. *Microbiology* 149, 3347–3351. doi: 10.1099/mic.0.26591-0
- Elzanowski, A., and Ostell, J. (2013). *The Genetic Codes [Online]. National Center for Biotechnology Information (NCBI)*. Available online at: <http://www.ncbi.nlm.nih.gov/Taxonomy/Utils/wprintgc.cgi?mode=c> (Accessed August 30 2014).
- Falb, M., Aivaliotis, M., Garcia-Rizo, C., Bisle, B., Tebbe, A., Kliein, C., et al. (2006). Archaeal N-terminal protein maturation commonly involves N-terminal acetylation: a large-scale proteomics survey. *J. Mol. Biol.* 362, 915–924. doi: 10.1016/j.jmb.2006.07.086
- Falkner, J. A., Falkner, J. W., Yocom, A. K., and Andrews, P. C. (2008). A spectral clustering approach to MS/MS identification of post-translational modifications. *J. Proteome Res.* 7, 4614–4622. doi: 10.1021/pr800226w
- Fischer, R., Gärtner, P., Yeliseev, A., and Thauer, R. K. (1992). N⁵-Methyltetrahydromethanopterin: coenzyme M methyltransferase in methanogenic archaeabacteria is a membrane protein. *Arch. Microbiol.* 158, 208–217. doi: 10.1007/BF00290817
- Fournier, C. T., Cherny, J. J., Truncali, K., Robbins-Pianka, A., Lin, M. S., Krizanc, D., et al. (2012). Amino termini of many yeast proteins map to downstream start codons. *J. Prot. Res.* 11, 5712–5719. doi: 10.1021/pr300538f
- Francoleon, D. R., Boontheung, P., Yang, Y., Kim, U., Ytterberg, A. J., Denny, P. A., et al. (2009). S-layer, surface-accessible, and concanavalin A binding proteins of *Methanosarcina acetivorans* and *Methanosarcina mazei*. *J. Prot. Res.* 8, 1972–1982. doi: 10.1021/pr800923e
- Gallien, S., Perrodou, E., Carapito, C., Deshayes, C., Reyrat, J. M., Van Dorsselaer, A., et al. (2009). Orthoproteogenomics: multiple proteomes investigation through orthology and a new MS-based protocol. *Genome Res.* 19, 128–135. doi: 10.1101/gr.081901.108
- Gevaert, K., Goethals, M., Martens, L., Van Damme, J., Staes, A., Thomas, G. R., et al. (2003). Exploring proteomes and analyzing protein processing by mass spectrometric identification of sorted N-terminal peptides. *Nat. Biotechnol.* 21, 566–569. doi: 10.1038/nbt810
- Grabarse, W., Halhert, F., Shima, S., Thauer, R. K., and Ermler, U. (2000). Comparison of three Methyl-Coenzyme M reductases from phylogenetically

- distant organisms: unusual amino acid modification, conservation and adaptation. *J. Mol. Biol.* 303, 329–344. doi: 10.1006/jmbi.2000.4136
- Hahne, H., Pachi, F., Ruprecht, B., Maier, S. K., Klaeger, S., Helm, D., et al. (2013). DMSO enhances electrospray response, boosting sensitivity of proteomic experiments. *Nat. Methods* 10, 989–991. doi: 10.1038/nmeth.2610
- Harms, U., and Thauer, R. K. (1996). Methylcobalamin:coenzyme M methyltransferase isoenzymes MtaA and MtB from *Methanosarcina barkeri*. *Eur. J. Biochem.* 235, 653–659. doi: 10.1111/j.1432-1033.1996.00653.x
- Hippler, B., and Thauer, R. K. (1999). The energy conserving methyltetrahydromethanopterin:coenzyme M methyltransferase complex from methanogenic archaea: function of the subunit MtrH. *FEBS Lett.* 449, 165–168. doi: 10.1016/S0014-5793(99)00429-9
- Hoelz, D. J., Arnold, R. J., Dobrolecki, L. E., Abdel-Aziz, W., Loehr, A. P., Novotny, M. V., et al. (2006). The discovery of labile methyl esters on proliferating cell nuclear antigen by MS/MS. *Proteomics* 6, 4808–4818. doi: 10.1002/pmic.200600142
- Hoppert, M., and Mayer, F. (1990). Electron microscopy of native and artificial methylreductase high-molecular-weight complexes in strain Göl and *Methanococcus voltae*. *FEBS Lett.* 267, 33–37. doi: 10.1016/0014-5793(90)80281-M
- Hovey, R., Lentes, S., Ehrenreich, A., Salmon, K., Saba, K., Gottschalk, G., et al. (2005). DNA microarray analysis of *Methanosarcina mazei* Göl reveals adaptation to different methanogenic substrates. *Mol. Gen. Genomics* 273, 225–239. doi: 10.1007/s00438-005-1126-9
- Howe, C. J., Auffret, A. D., Doherty, A., Bowman, C. M., Dyer, T. A., and Gray, J. C. (1982). Location and nucleotide sequence of the gene for the proton translocating subunit of wheat chloroplast ATP synthase. *Proc Natl Acad Sci U.S.A.* 79, 6903–6907.
- Jacobson, G. R., Schaffer, M. H., Stark, G. R., and Vanaman, T. C. (1973). Specific chemical cleavage in high yield at the amino peptide bonds of cysteine and cystine residues. *J. Biol. Chem.* 248, 6583–6591.
- Jaipuri, F. A., Collet, B. Y. M., and Pohl, N. L. (2008). Synthesis and quantitative evaluation of glycerol-D-manno-heptose binding to concanavalin A by fluorous-tag assistance. *Angew. Chem. Int. Ed.* 47, 1707–1710. doi: 10.1002/anie.200704262
- Juncker, A. S., Willenbrock, H., Von Heijne, G., Brunak, S., Nielsen, H., and Krogh, A. (2003). Prediction of lipoprotein signal peptides in gram-negative bacteria. *Protein Sci.* 12, 1652–1662. doi: 10.1110/ps.0303703
- Kahnt, J., Buchenau, B., Mahlert, F., Krüger, M., Shima, S., and Thauer, R. K. (2007). Post-translational modifications in the active site region of Methyl-Coenzyme M reductase from methanogenic and methanotrophic archaea. *FEBS J.* 274, 4913–4921. doi: 10.1111/j.1742-4658.2007.06016.x
- Kampmann, M., and Blobel, G. (2009). Nascent proteins caught in the act. *Science* 326, 1352–1353. doi: 10.1126/science.1183690
- Kim, J.-S., Dai, Z., Aryal, U. K., Moore, R. J., Camp, I. D. G., and Qian, W.-J. (2013). Resin-assisted enrichment of n-terminal peptides for characterizing proteolytic processing. *Anal. Chem.* 85, 6826–6832. doi: 10.1021/ac401000q
- Klunker, D., Haas, B., Hirtreiter, A., Figueiredo, L., Naylor, D. J., Pfeifer, G., et al. (2003). Coexistence of group I and group II chaperonins in the archaeon *Methanosarcina mazei*. *J. Biol. Chem.* 278, 33256–33267. doi: 10.1074/jbc.M302018200
- Kornfeld, R., and Ferris, C. (1975). Interaction of immunoglobulin glycoproteins with concanavalin A. *J. Biol. Chem.* 250, 2614–2619.
- Kozak, M. (1997). Interpreting cDNA sequences: some insights from studies on translation. *Mamm. Genome* 7, 563–574.
- Krätsch, C., Carini, P., Hovey, R., and Deppenmeier, U. (2009). Transcriptional profiling of methyltransferase genes during growth of *Methanosarcina mazei* on Trimethylamine. *J. Bacteriol.* 191, 5108–5115. doi: 10.1128/JB.00420-09
- Künkel, A., Vaupel, M., Heim, S., Thauer, R. K., and Hedderich, R. (1997). Heterodisulfide reductase from methanol-grown cells of *Methanosarcina barkeri* is not a flavoenzyme. *Eur. J. Biochem.* 244, 226–234. doi: 10.1111/j.1432-1033.1997.00226.x
- Künkel, A., Vorholt, J. A., Thauer, R. K., and Hedderich, R. (1998). An *Escherichia coli* Hydrogenase-3-type hydrogenase in methanogenic archaea. *FEBS J.* 252, 467–476. doi: 10.1046/j.1432-1327.1998.2520467.x
- Lemker, T., Grüber, G., Schmid, R., and Müller, V. (2003). Defined subcomplexes of the A1 ATPase from the Archaeon *Methanosarcina mazei* Göl: biochemical properties and redox regulation. *FEBS Letters* 544, 206–209. doi: 10.1016/S0014-5793(03)00496-4
- Lemker, T., Ruppert, C., Stöger, H., Wimmers, S., and Müller, V. (2001). Overproduction of a functional A1 ATPase from the Archaeon *Methanosarcina Gö1* in *Escherichia coli*. *Eur. J. Biochem.* 268, 3744–3750. doi: 10.1046/j.1432-1327.2001.02284.x
- Lienard, T., Becher, B., Marschall, M., Bowien, S., and Gottschalk, G. (1996). Sodium ion translocation by N⁵-methyltetrahydromethanopterin : coenzyme M methyltransferase from *Methanosarcina mazei* Göl reconstituted in ether lipid liposomes. *Eur. J. Biochem.* 239, 857–864. doi: 10.1111/j.1432-1033.1996.0857u.x
- Lienard, T., and Gottschalk, G. (1998). Cloning, sequencing and expression of the genes encoding the sodium translocating N⁵-Methyltetrahydromethanopterin:coenzyme M methyltransferase of the methyloptrophic archaeon *Methanosarcina mazei* Göl. *FEBS Lett.* 425, 204–208. doi: 10.1016/S0014-5793(98)00229-4
- Magidovich, H., and Eichler, J. (2009). Glycosyltransferases and oligosaccharyltransferases in archaea: putative components of the N-Glycosylation pathway in the third domain of life. *FEMS Microbiol. Lett.* 300, 122–130. doi: 10.1111/j.1574-6968.2009.01775.x
- Mechref, Y. (2012). Use of CID/ETD mass spectrometry to analyze glycopeptides. *Curr. Protoc. Protein Sci.* 68:12.11.1–12.1.11. doi: 10.1002/0471140864.ps1211s68
- Meinnel, T., and Giglione, C. (2008). Tools for analyzing and predicting N-Terminal protein modifications. *Proteomics* 8, 626–649. doi: 10.1002/pmic.200700592
- Menschaert, G., Vandekerckhove, T. T. M., Landuyt, B., Hayakawa, E., Schoofs, L., Luyten, W., et al. (2009). Spectral clustering in peptidomics studies helps to unravel modification profile of biologically active peptides and enhances peptide identification rate. *Proteomics* 9, 4381–4388. doi: 10.1002/pmic.200900248
- Messner, P. (2009). Prokaryotic protein glycosylation is rapidly expanding from “curiosity” to “ubiquity”. *Chem. Biochem.* 10, 2151–2154. doi: 10.1002/cbic.200900388
- Müller, V., Ruppert, C., and Lemker, T. (1999). Structure and Function of the A₁A₀-ATPases from methanogenic archaea. *J. Bioenerg. Biomembr.* 31, 15–27.
- Nefsky, B., and Bretscher, A. (1989). Landmark mapping: a general method for localizing cysteine residues within a protein. *Proc Natl Acad. Sci. U.S.A.* 86, 3549–3553.
- Ogorzalek Loo, R. R., Du, P., Loo, J. A., and Holler, T. (2002). *In-vivo* labeling: a glimpse of the dynamic proteome and additional constraints for protein identification. *J. Am. Soc. Mass Spectrom.* 13, 804–812. doi: 10.1016/S1044-0305(02)00408-7
- Parker, K. C., Garrels, J. I., Hines, W., Butler, E. M., McKee, A. H. Z., Patterson, D., et al. (1998). Identification of yeast proteins from two-dimensional gels: working out spot cross-contamination. *Electrophoresis* 19, 1920–1932. doi: 10.1002/elps.1150191110
- Patrie, S. M., Ferguson, J. T., Robinson, D. E., Whipple, D., Rother, M., Metcalf, W. W., et al. (2006). Top-Down Mass Spectrometry of <60 kDa proteins from *Methanosarcina acetivorans* using Q-FTMS with automated octopole collisionally activated dissociation (OCAD). *Mol. Cell. Proteomics* 5, 14–36. doi: 10.1074/mcp.M500219-MCP200
- Prats, A. C., Vagner, S., Prats, H., and Amalric, F. (1992). Cis-Acting elements involved in the alternative translation initiation process of human basic fibroblast growth factor mRNA. *Mol. Cell. Biol.* 12, 4796–4805.
- Randaccio, L., Geremia, S., and Wuerges, J. (2007). Crystallography of vitamin B12 proteins. *J. Organometallic Chem.* 692, 1198–1215. doi: 10.1016/j.jorgchem.2006.11.040
- Ring, G., and Eichler, J. (2004). Membrane binding of ribosomes occurs at SecYE-based sites in the archaea *Haloferax volcanii*. *J. Mol. Biol.* 336, 997–1010. doi: 10.1016/j.jmb.2004.01.008
- Rohlin, L., Leon, D. R., Kim, U., Loo, J. A., Ogorzalek Loo, R. R., and Gunsalus, R. P. (2012). Identification of the major expressed S-Layer and cell surface-layer-related proteins in the model methanogenic archaea: *Methanosarcina barkeri* fusaro and *Methanosarcina acetivorans* C2A. *Archaea* 2012:873589. doi: 10.1155/2012/873589
- Running, W., and Reilly, J. P. (2009). Ribosomal proteins of *Deinococcus radiodurans*: their solvent accessibility and reactivity. *J. Proteome Res.* 8, 1228–1246. doi: 10.1021/pr800544y
- Russo, A., Chandramouli, N., Zhang, L., and Deng, H. (2008). Reductive glutaraldehydration of amine groups for identification of protein N-termini. *J. Proteome Res.* 7, 4178–4182. doi: 10.1021/pr800224v

- Saleh, M., Song, C., Nasserulla, S., and Leduc, L. G. (2010). Indicators from archaeal secretomes. *Microbiol. Res.* 165, 1–10. doi: 10.1016/j.micres.2008.03.002
- Sauer, K., Harms, U., and Thauer, R. K. (1997). Methanol:coenzyme M methyltransferase from *Methanoscarcina barkeri*. purification, properties, and encoding genes of the corrinoid protein MT1. *Eur. J. Biochem.* 243, 670–677. doi: 10.1111/j.1432-1033.1997.t01-1-00670.x
- Savitski, M. M., Nielsen, M. L., and Zubarev, R. A. (2005). New data base-independent, sequence tag-based scoring of peptide MS/MS data validates mowse scores, recovers below threshold data, singles out modified peptides, and assesses the quality of MS/MS techniques. *Mol. Cell. Proteomics* 4, 1180–1188. doi: 10.1074/mcp.T500009-MCP200
- Savitski, M. M., Nielsen, M. L., and Zubarev, R. A. (2006). Modifcomb, a new proteomic tool for mapping stoichiometric post-translational modifications, finding novel types of modifications, and fingerprinting complex protein mixtures. *Mol. Cell. Proteomics* 5, 935–948. doi: 10.1074/mcp.T500034-MCP200
- Schmidt, A., Kellermann, J., and Lottspeich, F. (2005). A novel strategy for quantitative proteomics using isotope-coded protein labels. *Proteomics* 5, 4–15. doi: 10.1002/pmic.200400873
- Sebald, W., Graf, T., and Lukins, H. B. (1979). The Dicyclohexylcarbodiimide-binding protein of the mitochondrial ATPase complex from *Neurospora crassa* and *Saccharomyces cerevisiae*: identification and isolation. *Eur. J. Biochem.* 93, 587–599. doi: 10.1111/j.1432-1033.1979.tb12859.x
- Shen, P.-T., Hsu, J.-L., and Chen, S.-H. (2007). Dimethyl isotope-coded affinity selection for the analysis of free and blocked N-Termini of proteins using LC-MS/MS. *Anal. Chem.* 79, 9520–9530. doi: 10.1021/ac701678h
- Silva, J. C., Denny, R., Dorschel, C. A., Gorenstein, M., Kass, I. J., Li, G. Z., et al. (2005). Quantitative proteomic analysis by accurate mass retention time pairs. *Anal. Chem.* 77, 2187–2200. doi: 10.1021/ac048455k
- Sowers, K. R., Boone, J. E., and Gunsalus, R. P. (1993). Disaggregation of *Methanoscarcina* spp. and growth as single cells at elevated osmolarity. *Appl. Environ. Microbiol.* 59, 3832–3839.
- Thauer, R. K. (1998). Biochemistry of methanogenesis: a tribute to marty stephenson. *Microbiology* 144, 2377–2406. doi: 10.1099/00221287-144-9-2377
- Thauer, R. K., Kaster, A.-K., Seedorf, H., Buckel, W., and Hedderich, R. (2008). Methanogenic archaea: ecologically relevant differences in energy conservation. *Nat. Rev. Microbiol.* 6, 579–591. doi: 10.1038/nrmicro1931
- Trent, J. D., Kagawa, H. K., Paavola, C. D., Mcmillan, R. A., Howard, J., Jahnke, L., et al. (2003). Intracellular localization of a group II chaperonin indicates a membrane-related function. *Proc. Natl. Acad. Sci. U.S.A.* 100, 15589–15594. doi: 10.1073/pnas.2136795100
- Venne, A. S., Vögtle, F.-N., Meisinger, C., Sickmann, A., and Zahedi, R. P. (2013). Novel highly sensitive, specific, and straightforward strategy for comprehensive N-terminal proteomics reveals unknown substrates of the mitochondrial peptidase Icp55. *J. Prot. Res.* 12, 3823–3830. doi: 10.1021/pr400435d
- Wilhelm, T., and Jones, A. M. E. (2014). Identification of related peptides through the analysis of fragment ion mass shifts. *J. Prot. Res.* 13, 4002–4011. doi: 10.1021/pr500347e
- Wu, J., and Watson, J. T. (1997). A novel methodology for assignment of disulfide bond pairings in proteins. *Protein Sci.* 6, 391–398. doi: 10.1002/pro.5560060215
- Xing, G., Zhang, J., Chen, Y., and Zhao, Y. (2008). Identification of four novel types of *in vitro* protein modifications. *J. Proteome Res.* 7, 4603–4608. doi: 10.1021/pr800456q
- Xu, G., and Jaffrey, S. R. (2010). N-CLAP: global profiling of N-Termini by chemoselective labeling of the α-amine of proteins. *Cold Spring Harb. Protoc.* 10: pdb.prot5528. doi: 10.1101/pdb.prot5528
- Yalcin, T., Khouw, C., Csizmadia, I. G., Peterson, M. R., and Harrison, A. G. (1995). Why are *b* Ions stable species in peptide spectra? *J. Am. Soc. Mass Spectrom.* 6, 1165–1174. doi: 10.1016/1044-0305(95)00569-2
- Yamaguchi, M., Nakayama, D., Shima, K., Kuyama, H., Ando, E., Okamura, T.-A., et al. (2008). Selective isolation of N-terminal peptides from proteins and their *de novo* sequencing by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry without regard to unblocking of N-terminal amino acids. *Rapid Commun. Mass Spectrom.* 22, 3313–3319. doi: 10.1002/rcm.3735
- Yamazaki, S., Yamazaki, J., Nishijima, K., Otsuka, R., Mise, M., Ishikawa, H., et al. (2006). Proteome analysis of an aerobic hyperthermophilic crenarchaeon, *Aeropyrum pernix* K1. *Mol. Cell. Prot.* 5, 811–823. doi: 10.1074/mcp.M500312-MCP200
- Ytterberg, A. J., Peltier, J. B., and van Wijk, K. J. (2006). Protein profiling of plastoglobules in chloroplasts and chromoplasts. A surprising site for differential accumulation of metabolic enzymes. *Plant Physiol.* 140, 984–997. doi: 10.1104/pp.105.076083
- Yurist-Doutsch, S., Chaban, B., Vandyke, D. J., Jarrell, K. F., and Eichler, J. (2008). Sweet to the extreme: protein glycosylation in archaea. *Mol. Microbiol.* 68, 1079–1084. doi: 10.1111/j.1365-2958.2008.06224.x
- Zubarev, R. A. (2004). Electron-capture dissociation tandem mass spectrometry. *Curr. Opin. Biotechnol.* 15, 12–16. doi: 10.1016/j.copbio.2003.12.002

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

Received: 01 November 2014; accepted: 09 February 2015; published online: 05 March 2015.

Citation: Leon DR, Ytterberg AJ, Boontheung P, Kim U, Loo JA, Gunsalus RP and Ogorzalek Loo RR (2015) Mining proteomic data to expose protein modifications in *Methanoscarcina mazei* strain Gö1. *Front. Microbiol.* 6:149. doi: 10.3389/fmicb.2015.00149

This article was submitted to Microbial Physiology and Metabolism, a section of the journal *Frontiers in Microbiology*.

Copyright © 2015 Leon, Ytterberg, Boontheung, Kim, Loo, Gunsalus and Ogorzalek Loo. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Novel pili-like surface structures of *Halobacterium salinarum* strain R1 are crucial for surface adhesion

Gerald Losensky¹, Lucia Vidakovic¹, Andreas Klingl^{2,3}, Felicitas Pfeifer¹ and Sabrina Fröls^{1*}

¹ Microbiology and Archaea, Department of Biology, Technische Universität Darmstadt, Darmstadt, Germany

² Cell Biology and LOEWE Research Centre for Synthetic Microbiology, Philipps-Universität Marburg, Marburg, Germany

³ Department of Biology I, Biozentrum, University of Munich, Planegg-Martinsried, Germany

Edited by:

Mechthild Pohlschröder, University of Pennsylvania, USA

Reviewed by:

Scott Chimileski, University of Connecticut, USA

Olivera Francetic, Institut Pasteur, France

*Correspondence:

Sabrina Fröls, Microbiology and Archaea, Department of Biology, Technische Universität Darmstadt, Schnittspahnstrasse 10, 64287 Darmstadt, Germany
e-mail: froels@bio.tu-darmstadt.de

It was recently shown that haloarchaeal strains of different genera are able to adhere to surfaces and form surface-attached biofilms. However, the surface structures mediating the adhesion were still unknown. We have identified a novel surface structure with *Halobacterium salinarum* strain R1, crucial for surface adhesion. Electron microscopic studies of surface-attached cells frequently showed pili-like surface structures of two different diameters that were irregularly distributed on the surface. The thinner filaments, 7–8 nm in diameter, represented a so far unobserved novel pili-like structure. Examination of the *Hbt. salinarum* R1 genome identified two putative gene loci (*pil-1* and *pil-2*) encoding type IV pilus biogenesis complexes besides the archaeillum encoding *fla* gene locus. Both *pil-1* and *pil-2* were expressed as transcriptional units, and the transcriptional start of *pil-1* was identified. *In silico* analyses revealed that the *pil-1* locus is present with other euryarchaeal genomes whereas the *pil-2* is restricted to haloarchaea. Comparative real time qRT-PCR studies indicated that the general transcriptional activity was reduced in adherent vs. planktonic cells. In contrast, the transcription of *pilB1* and *pilB2*, encoding putative type IV pilus assembly ATPases, was induced in comparison to the archaealla assembly/motor ATPase (*flaI*) and the ferredoxin gene. Mutant strains were constructed that incurred a *flaI* deletion or *flaI/pilB1* gene deletions. The absence of *flaI* caused the loss of the archaealla while the additional absence of *pilB1* led to loss of the novel pili-like surface structures. The $\Delta flaI/\Delta pilB1$ double mutants showed a 10-fold reduction in surface adhesion compared to the parental strain. Since surface adhesion was not reduced with the non-archaeallated $\Delta flaI$ mutants, the *pil-1* filaments have a distinct function in the adhesion process.

Keywords: haloarchaea, deletion mutant, archaeillum, *Halobacterium salinarum*, surface adhesion, archaeal type IV pili

INTRODUCTION

Various filamentous surface structures have been identified in *Archaea* mediating surface attachment or the formation of cell-cell contacts. These are classified into archaeal type IV pili and non-type IV pili structures. The latter group consists of structurally very diverse representatives (Lassak et al., 2012). The SM1 euryarchaeon forms prickled filaments referred to as “hami” with a distal hook-like structure connecting the cells in a three-dimensional arrangement with regular distances (Moissl et al., 2005; Henneberger et al., 2006). Surface attached cells of *Methanocaldococcus thermoautotrophicum* show networks of thin filaments with diameters of 5.5 nm, the Mth60 fimbriae (Thoma et al., 2008).

Abbreviations: ARF-TSS, Adaptor-and radioactivity-free identification of transcription start site; Fla, Flagella (archaealla) accessory proteins; FlaI, Flagella (archaealla) motor/biogenesis protein; Pil, Type IV pilus biogenesis complex; PilB, Type IV pilus biogenesis complex ATPase subunit; PilC, Type IV pilus biogenesis complex membrane subunit; qRT-PCR, Quantitative reverse transcription polymerase chain reaction; RT-PCR, Reverse transcription polymerase chain reaction; TEM, Transmission electron microscopy.

Type IV pili and type IV pili-like structures including the archaealla (also called archaeal flagella, Jarrell and Albers, 2012) are present in several euryarchaeal and crenarchaeal species. The type IV pilus biogenesis complexes of archaeal pili are related to those of bacteria (Peabody et al., 2003). The structural components of the pili (Pil) and archaealla (Fla) include the assembly/motor-ATPase PilB/FlaI and the multispanning transmembrane protein PilC/FlaJ. These proteins represent in conjunction with the pilins or archaeallins the core components of archaeal type IV pilus systems (Pohlschröder et al., 2011). Archaeal type IV pili mediate versatile functions like motility, adhesion to biotic or abiotic surfaces, cell-cell connections, biofilm formation and architecture, as well as DNA-exchange (Lassak et al., 2012). A common adhesion strategy has not been observed for *Archaea* so far. In the case of *Pyrococcus furiosus*, the multifunctional archaealla are sufficient to mediate surface adhesion and the formation of cell-cell connections (Näther et al., 2006). Surface adhesion of *Haloferax volcanii* solely depends on pili and does not require the presence of the archaealla (Triepel et al., 2010). For *Methanococcus maripaludis* and *Sulfolobus solfataricus* pili and

archaella are both necessary for the attachment to a variety of abiotic surfaces (Zolghadr et al., 2010; Jarrell et al., 2011).

By a screening approach with various haloarchaeal strains we demonstrated that surface adhesion is found with the genera *Halobacterium*, *Haloferax*, *Halorubrum*, and *Halohasta*. Different type strains, their derivatives and natural isolates of *Halobacterium salinarum* are able to adhere to abiotic surfaces, which supports the subsequent formation of biofilms. Initial studies by transmission electron microscopy (TEM) of surface attached *Hbt. salinarum* DSM 3754^T cells showed various filamentous structures on the surface (Fröls et al., 2012). However, the structures mediating surface adhesion are still unknown. In the case of *Hbt. salinarum*, the archaella are the only surface structures described to date (Alam and Oesterhelt, 1984). These filaments are 10 nm in diameter, polarly localized and enable the cells to swim by an ATP-driven rotation (Alam and Oesterhelt, 1984; Cohen-Krausz and Trachtenberg, 2002; Streif et al., 2008).

The aim of the present work was to identify the filamentous structures involved in surface adhesion of *Hbt. salinarum* strain R1 cells *in vitro* and *in vivo*. TEM analyses were used to identify and classify filamentous structures present with surface attached cells of *Hbt. salinarum* R1. A novel type of filamentous structures was observed. The genome sequence was analyzed to search for putative type IV pili gene loci encoding filamentous surface structures other than archaella and two putative type IV pilus biogenesis (*pil*) gene loci were identified. The transcripts were determined and the transcriptional activity of the assembly/motor-ATPases *pilB/flaI* genes was examined by comparative *real time* qRT-PCR analyses in planktonic and surface attached cells. Deletion mutants were constructed to investigate the presence of filamentous structures in dependency of *flaI* and *pilB1* as well as elucidating their role with regard to motility and surface adhesion.

MATERIALS AND METHODS

STRAINS AND GROWTH CONDITIONS

Hbt. salinarum strains R1, DSM 3754^T, PHH1, PHH4, SB3, GN101, and NRC-1 (strain details were listed in Table 1) were grown aerobically at 37°C in complex medium (250 g NaCl, 20 g MgSO₄ × 7H₂O, 2 g KCl, 15 g Oxoid peptone, 50 ml 1 M Tris/HCl pH 7.5 per liter). For cultivation of planktonic and adherent cells an overnight culture with an optical density at 600 nm (OD₆₀₀) of 0.3 was used for inoculation. Before growth the OD₆₀₀ was set to 0.002. Planktonic cells were grown in cultures shaking at 180 rpm and harvested during the exponential growth phase at OD₆₀₀ 0.3 for RNA preparation and OD₆₀₀ 0.5 for DNA preparation. Adherent cells were grown in large Petri dishes (150/20 mm, Sarstedt) as static cultures. After 6 days of growth the supernatant was discarded and the dishes were washed three times with 50 mL salt water (complex medium without peptone) to remove non-adhering cells. Adherent cells were scraped off the dishes using a spatula.

TRANSMISSION ELECTRON MICROSCOPY (TEM)

Planktonic *Hbt. salinarum* R1 cells were grown at 37°C to OD₆₀₀ 0.8, fixed with 1% glutaraldehyde for 30 min at room temperature and applied onto carbon coated copper grids (400 mesh,

Table 1 | Strains and plasmids used for the studies.

Strain	Origin	References
<i>Hbt. salinarum</i> R1 (DSM 671, ATCC 2934)	Derived from DSM670	Pfeiffer et al., 2008b
<i>Hbt. salinarum</i> DSM 3754 ^T (ATCC 33171)	Type strain	Elazari-Volcani 1957 (Gruber et al., 2004).
<i>Hbt. salinarum</i> PHH4	Derived from PHH1	Pfeifer and Blaseio, 1989
<i>Hbt. salinarum</i> SB3	Natural isolate	Ebert et al., 1984
<i>Hbt. salinarum</i> GN101	Natural isolate	Ebert et al., 1984
<i>Hbt. salinarum</i> NRC-1 (ATCC 700922)	Type strain	Pfeiffer et al., 2008b
<i>Hbt. salinarum</i> PHH1	Derived from DSM670	Pfeifer et al., 1988
<i>Hbt. salinarum</i> R1 Δ <i>flaI</i>	In frame deletion of OE2380R (1890 bp)	This study
<i>Hbt. salinarum</i> R1 Δ <i>flaI</i> / Δ <i>pilB1</i>	Derived from <i>Hbt. salinarum</i> R1 Δ <i>flaI</i> In frame deletion of OE2215R (1758 bp)	This study
Plasmid	Relevant characteristics	References
pMKK100	<i>E. coli</i> / <i>Hbt. salinarum</i> shuttle plasmid, mevinoline resistance, <i>bgaH</i> reporter gene	Koch and Oesterhelt, 2005
pMKK100- Δ <i>flaI</i>	Contains fusion fragment of the up- (496 bp) and downstream (440 bp) regions of OE2380R	This study
pMKK100- Δ <i>pilB1</i>	Contains fusion fragment of the up- (492 bp) and downstream (601 bp) regions of OE2215R	This study

Plano GmbH) for 30 s. For the investigation of adherent cells, carbon coated gold grids (400 mesh, Plano GmbH) were placed in a freshly inoculated culture and incubated in static culture at 42°C for 10 days. Adherent cells were fixed with 2% paraformaldehyde w/v and 1% glutaraldehyde w/v over night at 4°C. Samples were washed eight times with double distilled water to avoid the formation of salt crystals. After removing excess fluid using filter paper, samples were contrasted with 2% uranyl acetate (pH 6, containing maleic acid) for 60 s and stored in a desiccator containing silica gel. The software ImageJ (National Institutes of Health, <http://rsb.info.nih.gov/ij/index.html>) was used for measuring the diameter in nm of visible surface appendages.

IN SILICO ANALYSES

The somewhat similar sequences (blastn), position-specific iterated (psi-blast) and protein-protein (blastp) BLAST® alignment search tools were used to analyze gene and protein identities, functions and presence with other genomes (NCBI Resource Coordinators, 2014). Additional analyses were performed using

HaloLex (Pfeiffer et al., 2008a), the UCSC Archaeal Genome Browser (Schneider et al., 2006; Chan et al., 2012) and SMART (Schultz et al., 1998; Letunic et al., 2012). The predictions of transmembrane helices in proteins were performed using the software TMHMM Server v. 2.0 (Krogh et al., 2001), archaeal class III (type IV pilin-like) signal peptides by use of FlaFind 1.2 (Szabo et al., 2007; Esquivel et al., 2013) and the secondary structures of single stranded nucleic acids by the software mfold (Zuker, 2003).

RNA PREPARATION

Total RNA was isolated from planktonic and adherent cells by standard acid guanidinium thiocyanate-phenol-chloroform extraction (Chomczynski and Sacchi, 2006). Genomic DNA was removed by treatment with RNase-free DNaseI (# EN0523, Thermo Fisher Scientific) for 4 h at 37°C. Purified RNA was used to generate complementary DNA (cDNA).

REVERSE TRANSCRIPTION POLYMERASE CHAIN REACTION (RT-PCR)

For transcript mapping 40 µg purified RNA were reversely transcribed into cDNA using Random Hexamer Primers (# SO142, Thermo Fisher Scientific) and RevertAid Reverse Transcriptase (# EP0441, Thermo Fisher Scientific) in a total volume of 160 µL according to the manufacturer's protocol. To investigate co-transcription of neighboring genes, oligonucleotides were designed to amplify fragments encompassing the intergenic region and overlapping adjacent genes (see Table S1 and Figure 2). In case of co-transcription, these primers will lead to PCR products using cDNA as template. RT-PCR analysis of *pil-1* was performed using *Taq/Pfu*-polymerase mix 19:1 (# EP0702 and # EP0502, Thermo Fisher Scientific) (initial step 300 s at 95°C, 35 cycles of 60 s at 95 °C, 90 s at 54°C to 64°C, 135 s at 72°C, end step 300 s at 72°C) according to the manufacturer's protocol. For analysis of *pil-2* the more sensitive Q5-polymerase (# M0491L, New England Biolabs) was used (initial step 300 s at 98°C, 35 cycles of 10 s at 98°C, 30 s at 49°C to 60°C, 40 s at 72°C, end step 120 s at 72°C). Control reactions were performed using a similar RNA sample without reverse transcription to exclude a possible genomic DNA contamination. PCR was performed to validate the amplicon size and specificity of the oligonucleotides using *Hbt. salinarum* R1 genomic DNA as template.

TRANSCRIPTION START SITE DETERMINATION (ARF-TSS)

The transcription start site (TSS) was determined using the Adaptor-and radioactivity-free (ARF-TSS) method described by Wang et al. (2012). Purified RNA was used to generate first strand cDNA using the *pilB1* gene specific oligonucleotide TSS-pil-1-P1-RT (see Table S1 for oligonucleotide sequences) complementary to a sequence located 160 bases downstream of the annotated start codon of OE2215R. The cDNA was circularized using T4 RNA ligase (# EL0021, Thermo Fisher Scientific) to fuse the 3'- and the 5'-end of the cDNA. The circularized cDNA served as template for PCR using the two diverging oligonucleotides TSS-pil-1-P2-PCR and TSS-pil-1-P3-PCR, binding between the sites of the gene specific oligonucleotide and the TSS. PCR products were inserted into pCR® 2.1-TOPO® using TOPO TA Cloning® Kit for Sequencing (# 450641, Invitrogen) following the protocol

of the manufacturer and the resulting constructs were used for sequence analysis with standard M13 oligonucleotides.

DNA PREPARATION AND SOUTHERN ANALYSIS

For preparation of genomic DNA, 2 mL cell culture (OD₆₀₀ 0.5) were sedimented by centrifugation, the cell pellet resuspended in 100 µL salt water and lysed osmotically by the addition of 900 µL TEN-buffer (100 mM NaCl, 1mM EDTA, 20 mM Tris/HCl pH 8.0). Standard phenol/chloroform extraction was performed followed by DNA precipitation using isopropyl alcohol. For Southern analysis 3 µg of genomic DNA cut with *Aat*II were separated on 0.7% agarose gels and blotted on Roti®Nylon membranes (pore size 0.2 µm, Carl Roth GmbH & Co. KG). Southern blots were hybridized in standard hybridization buffer with digoxigenin-labeled DNA-probes and detected by use of Anti-digoxigenin-alkaline phosphatase Fab fragments (# 11093274910, Roche) in combination with the Phototope®-Star Detection Kit (# N7020S, New England Biolabs) according to the manufacturer's protocols. A digoxigenin DNA labeling Kit (# 11277065910, Roche) was used to produce DNA-probes by standard PCR using genomic DNA of *Hbt. salinarum* R1 as template in combination with the following oligonucleotides: pil-1-probe-fwd and pil-1-probe-rev producing a 1541 bp PCR product; pil-2-probe-fwd and pil-2-probe-rev producing a 686 bp PCR product (see Table S1 for oligonucleotide sequences).

QUANTITATIVE REVERSE TRANSCRIPTION POLYMERASE CHAIN REACTION (qRT-PCR)

For qRT-PCR 5 µg RNA supplemented with 1 ng of an external standard RNA in a total volume of 20 µL were used to generate the cDNA. External standard RNA (length 1790 nt) was produced by *in vitro* transcription of the *bgaH* gene, using T7 RNA polymerase (# EP0111, Thermo Fisher Scientific) according to the manufacturer's protocol. qRT-PCR analysis was performed using the StepOne™ Real-Time PCR System (Applied Biosystems) and the SensiFast™ SYBR Hi-ROX Kit (# BIO-92005, Bioline) according to the manufacturer's protocol. Using the StepOne™ software v2.0 the $\Delta\Delta C_T$ -method was applied to calculate relative expression changes of the target genes in adherent cells compared to their expression in planktonic cells (Schmittgen and Livak, 2008). C_T -values were determined by the StepOne™ software (Applied Biosystems). C_T -values of the housekeeping genes *rpoB1* (OE4741R) and *aef2* (OE4729R) were normalized to the external standard *bgaH* to investigate the general transcriptional activity of the cells. C_T -values of the target genes were normalized to the housekeeping gene *rpoB1* (Bleiholder et al., 2012). Samples were examined in triplicates. Control reactions checking for genomic DNA contamination were done as described before.

CONSTRUCTION OF DELETION MUTANTS

The construction of deletion mutants was performed using the pop-in/pop-out strategy (Koch and Oesterhelt, 2005). Approximately 500 bp upstream (US) and downstream (DS) of the gene of interest were amplified from genomic DNA (initial step 300 s at 95°C, 30 cycles of 60 s at 95°C, 60 s at 58°C to 72°C, 30 s to 150 s at 72°C, end step 600 s at 72°C) and fused

by PCR (initial step 300 s at 95°C, 60 s at 60°C, 60 s at 72°C, 10 cycles of 60 s at 95°C, 60 s at 60°C, 30 s at 72°C, end step 600 s at 72°C). Oligonucleotides used are listed in Table S1. The fused US/DS PCR products were cloned into pMKK100 (Koch and Oesterhelt, 2005). Polyethylenglycol-mediated transformation of *Hbt. salinarum* R1 was carried out (Dyall-Smith, 2008), followed by red-blue screening (6 µg/mL of mevinolin and 40 µg/mL of X-gal). Blue transformants were used to inoculate liquid cultures without mevinolin (to induce the pop-out) for three subsequent cultivations with complex medium and plated on agar media containing 40 µg mL⁻¹ X-gal. Red colonies were selected for the absence of the gene of interest and verified at the site of the deletion by PCR using genomic DNA (initial step 300 s at 95°C, 30 cycles of 60 s at 95°C, 60 s at 58°C to 72°C, 30 s to 150 s at 72°C, end step 600 s at 72°C) followed by sequencing with US/DS flanking oligonucleotides (listed in Table S1). The deletion strains and plasmids are listed in **Table 1**.

MOTILITY ASSAY

To investigate swimming motility, *Hbt. salinarum* R1 and the two deletion mutants were grown in semi-solid medium containing 0.3% agar (w/v) (Patenge et al., 2001). 10 µL of a liquid culture (OD₆₀₀ 0.3) were placed in the center of the agar surface and incubated over 96 h in the dark at 42°C. The diameter of each motility halo was measured in cm using the software ImageJ (National Institutes of Health, <http://rsb.info.nih.gov/ij/index.html>). The average motility halo and standard deviation was calculated from 19 replicates per strain.

SURFACE ADHESION ON GLASS

Hbt. salinarum strains were grown in Petri dishes (92/16 mm, Sarstedt) containing 15 mL complex medium inoculated with cells from the exponential growth phase (OD₆₀₀ 0.3–0.5). The starting culture was set to a calculated OD₆₀₀ of 0.002 before the cells were grown at 42°C for 10 days. Coverslips were inserted into the media to allow adherence on glass. Prior the microscopic analyses, overgrown coverslips were washed three times with salt water (complex medium without peptone) to remove all non-adherent cells. Microscopic analyses were performed using a Zeiss Axioskop 2 (camera AxioCam MRm, software AxioVision). The software ImageJ (National Institutes of Health, <http://rsb.info.nih.gov/ij/index.html>) was used to select (using the color-based thresholding function) and measure (using the analyze particles function) the percentage of surface coverage. The quantifications are based on the surface-attached cells of six independent visual fields of separate coverslips, from at least two independent inoculated cultures. Significances (*p*-values) of the percentage of surface coverage between the parental and mutant strains were calculated by an unpaired, two-tailed *t*-test.

RESULTS

IDENTIFICATION OF SURFACE STRUCTURES PRESENT WITH SURFACE ATTACHED CELLS OF *Hbt. salinarum* R1

TEM was used to investigate the presence of filamentous structures with surface attached cells of *Hbt. salinarum* strain R1 grown on carbon coated gold grids for 10 days (**Figure 1A**).

Numerous long (> 50 nm) and flexible pili-like surface structures were observed between the cells on the carbon surface (**Figures 1B,C**). The majority of the surface structures were distributed irregularly as single filaments. Many of the filaments originated from the cell poles, but for most of those no origin could be determined. Differences regarding the diameter of single filaments were noticed at higher magnification (**Figures 1D,E**). The diameters of 100 filaments of at least 10 to 15 photographs were determined using the software ImageJ. The diameters of the filaments were in minimum between 5 and 6 nm and in maximum 14 and 15 nm. The frequency distribution showed two dominant diameters of 7–8 nm and 10–11 nm both with a maximum of 20% (**Figure 1F**). These data suggested the presence of two different filamentous surface structures in *Hbt. salinarum* R1 with average diameters of 7.6 ± 0.7 nm (calculated on the categories 6 to 9 nm, *n* = 30) and 10.3 ± 0.8 nm (calculated on the categories 9 to 12 nm, *n* = 46). The width of the thicker filaments is consistent with the archaella of *Hbt. salinarum* R1 (Alam and Oesterhelt, 1984; Cohen-Krausz and Trachtenberg, 2002), whereas the thinner filaments might represent a novel pili-like surface structure.

IN SILICO IDENTIFICATION AND TRANSCRIPTION OF PUTATIVE TYPE IV PILI GENE LOCI

The assembly/motor-ATPase FlaI (OE2380R) of the archaellum encoding gene cluster was used as starting point to identify further putative type IV pilus biogenesis complex gene loci in the genome sequence of *Hbt. salinarum* R1 (NC_010364.1, Pfeiffer et al., 2008b). By blastp analyses two putative type IV pilus assembly ATPases (PilB) homologs of FlaI were identified which share an amino acid identity of 35 and 28% (query coverage 62 and 48%, *e*-value 6e-75 and 5e-21), respectively. These proteins were termed PilB1 (OE2215R) and PilB2 (OE1347R) based on the amino acid identity compared to FlaI. Both proteins contain a conserved VirB11 ATPase domain found with archaeal type IV pilus secretion systems (arCOG01817, arCOG01818). The genes *pilB1* and *pilB2* (encoding the putative assembly ATPases) were located adjacent to genes coding for putative multispanning transmembrane archaella/pilus assembly proteins (arCOG01808, arCOG01810) possessing 7 and 9 predicted transmembrane helices. These genes were termed *pilC1* (OE2212R) and *pilC2* (OE1344R), with the latter gene having an in-frame stop codon. The encoded protein PilC1 only has a significant amino acid identity of 23% (query coverage 24%, *e*-value 2e-04) to the archaeal transmembrane protein FlaJ (OE2379R).

Reverse transcriptase polymerase chain reaction analyses (RT-PCR) were performed to determine whether the corresponding genes of the putative type IV pilus assembly ATPases *pilB1* (OE2215R), *pilB2* (OE1347R) and putative transmembrane proteins *pilC1* (OE2212R), *pilC2* (OE1344/42R) are transcribed. DNaseI-treated total RNA of *Hbt. salinarum* R1 was used for the generation of the cDNA. No amplification products were observed with RNA, which was not reversely transcribed as template, confirming the absence of DNA contaminations in the sample. Gene specific oligonucleotides were used for the RT-PCR reaction (listed in Table S1) to amplify gene-to-gene overlapping fragments of adjacent genes. The RT-PCR studies confirmed

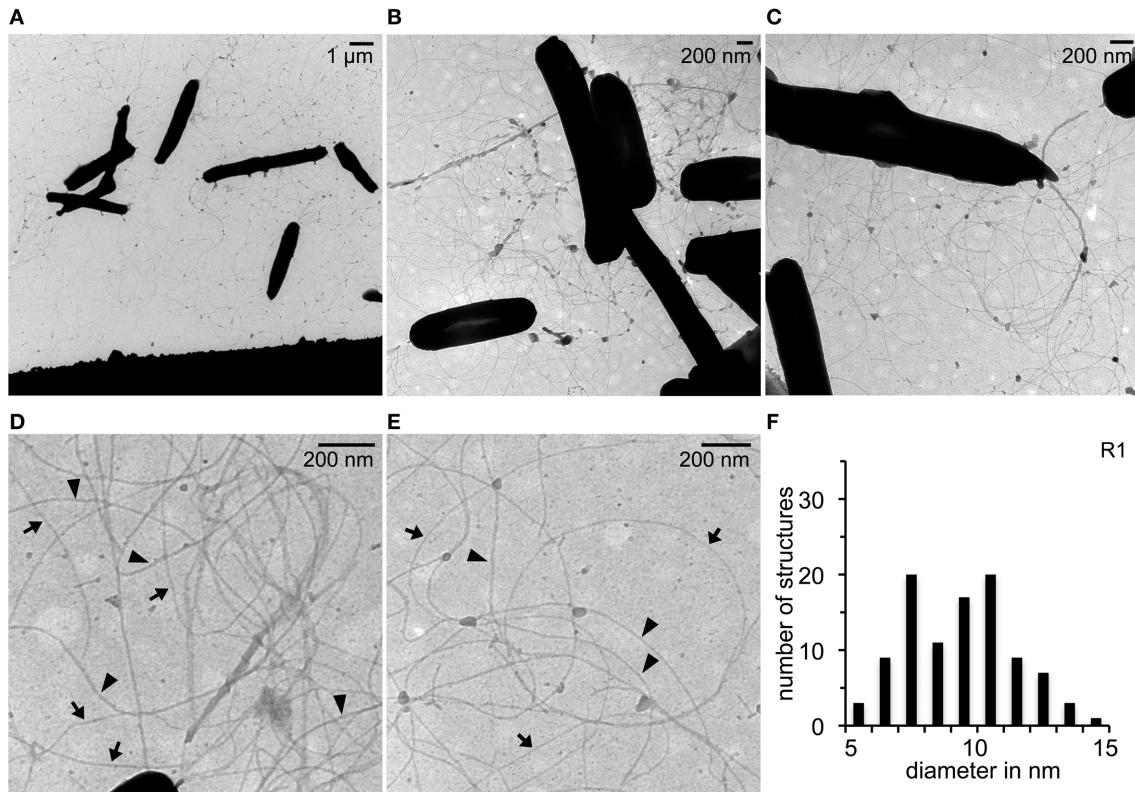


FIGURE 1 | TEM analyses of *Hbt. salinarum* R1 attached to surfaces.

(A) Surface attached cells grown for 10 d at 42°C on carbon coated gold grids. (B,C) Pili-like surface structures observed on the carbon surface. (D,E)

Determination of different diameters of pili-like structures. Arrow heads, wide diameters; arrows, thinner structures. (F) Frequency distribution observed with different diameters determined for 100 surface structures.

the transcriptional activity of the identified *pilB1*, *pilC1*, and *pilB2*, *pilC2* genes (Figures 2A,C). RT-PCR analyses of the gene regions upstream of *pilB1/2* and downstream of *pilC1/2* yielded the transcriptional unit of the putative *pil-1/2* gene loci. The transcript of the *pil-1* locus (4.4 kbp) spans the three open reading frames OE2215R, OE2212R, and OE2210R (Figure 2A). Minor or not detectable amplification products were observed for the oligonucleotides combination used to amplify the overlapping gene region of OE2217R and OE2215R. Therefore, the adaptor-and radioactivity-free (ARF) method by Wang et al. (2012) was used to determine the transcriptional start site of the *pil-1* locus (Wang et al., 2012). The sequence analyses of four PCR-products identified the guanine, 4 nt downstream of the predicted AUG as start site of the *pilB1* transcript. A GUG motif as alternative translational start codon is present at position +76 nt, implying the presence of a 5'-untranslated region (5'-UTR) with *pilB1* (Figure 2B). With regard to the *pil-2* locus (6.9 kpb), co-transcription of seven open reading frames, OE1347R through OE1332R, was determined (Figure 2C). These include three genes encoding potential prepilins (OE1340R, OE1336R, OE1334R) containing type IV pilin-like signal peptides as predicted by the software FlaFind 1.2. No potential prepilin encoding genes were found with the *pil-1* locus or within the surrounding genomic region of about 100 kbp. A schematic illustration summarizing the results for the *pil-1* and *pil-2* loci is given in Figure S1.

PRESENCE OF THE *pil-1* AND *pil-2* GENE LOCI IN OTHER HALOARCHAEAL GENOMES

Southern analyses using gene specific digoxigenin labeled *pilB1/C1* and *pilB2* probes were carried out to investigate the occurrence of the *pil-1* and *pil-2* loci in seven different *Hbt. salinarum* strains. The type strain DSM3754^T, the three closely related wild type strains R1, NRC-1, PHH1, and the PHH1 derivative PHH4 (strain information is given in Table 1) as well as two natural isolates derived from salt flats in San Francisco (SB3) or Guerrero Negro (GN101) sharing 16S rRNA sequence identities of 98 to 99% to *Hbt. salinarum* R1 but possessing distinct plasmid populations different from the three wild type strains (Ebert et al., 1984). All these species differ in their ability to adhere to surfaces. *Hbt. salinarum* R1 and DSM 3754^T show strong, *Hbt. salinarum* PHH4 and SB3 moderate adhesion to a plastic surface, whereas no significant adhesion is observed for *Hbt. salinarum* NRC-1, PHH1, and GN101 (Fröls et al., 2012).

To investigate the presence of the *pil-1* and *pil-2* loci in the genomes of these *Hbt. salinarum* strains, total DNA was isolated and hydrolyzed with the restriction enzyme *AatII* for Southern analyses (Figure S2A). The *pilB1/C1* probe hybridized with DNA of all strains but strain specific variations were observed (Figure S2B). In five of the seven strains tested (R1, NRC-1, PHH1, PHH4, and SB3) the *pilB1/C1* probe hybridized with two fragments of 4.1 kbp and 3.3 kbp, corresponding to

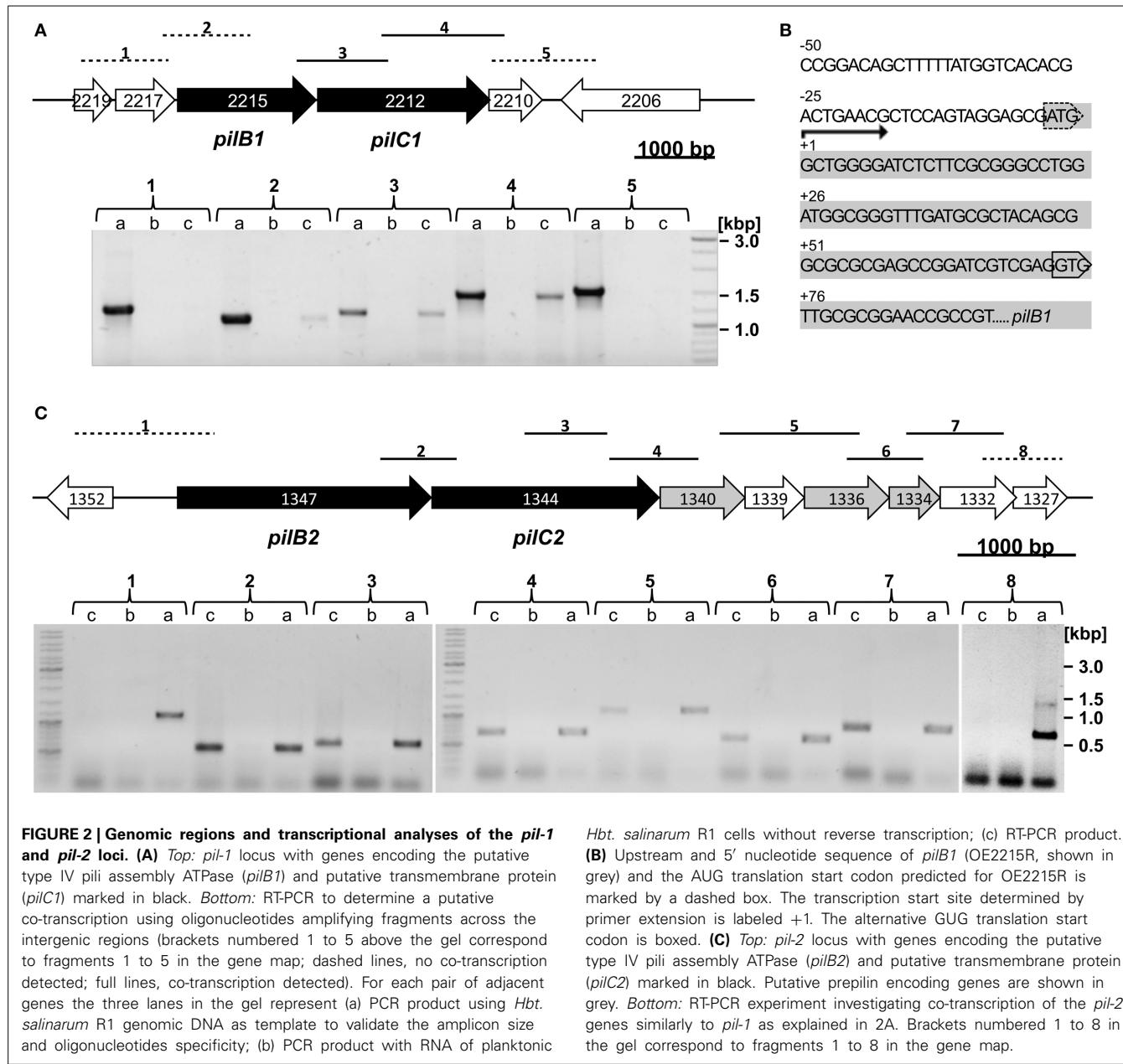


FIGURE 2 | Genomic regions and transcriptional analyses of the *pil-1* and *pil-2* loci. (A) Top: *pil-1* locus with genes encoding the putative type IV pilus assembly ATPase (*pilB1*) and putative transmembrane protein (*pilC1*) marked in black. Bottom: RT-PCR to determine a putative co-transcription using oligonucleotides amplifying fragments across the intergenic regions (brackets numbered 1 to 5 above the gel correspond to fragments 1 to 5 in the gene map; dashed lines, no co-transcription detected; full lines, co-transcription detected). For each pair of adjacent genes the three lanes in the gel represent (a) PCR product using *Hbt. salinarum* R1 genomic DNA as template to validate the amplicon size and oligonucleotides specificity; (b) PCR product with RNA of planktonic

Hbt. salinarum R1 cells without reverse transcription; (c) RT-PCR product. **(B)** Upstream and 5' nucleotide sequence of *pilB1* (OE2215R, shown in grey) and the AUG translation start codon predicted for OE2215R is marked by a dashed box. The transcription start site determined by primer extension is labeled +1. The alternative GUG translation start codon is boxed. **(C)** Top: *pil-2* locus with genes encoding the putative type IV pilus assembly ATPase (*pilB2*) and putative transmembrane protein (*pilC2*) marked in black. Putative prepeptidase genes are shown in grey. Bottom: RT-PCR experiment investigating co-transcription of the *pil-2* genes similarly to *pil-1* as explained in 2A. Brackets numbered 1 to 8 in the gel correspond to fragments 1 to 8 in the gene map.

the theoretical sizes calculated from the genome sequence of *Hbt. salinarum* R1 (NC_010364.1). However, only one restriction fragment (3.3 kbp) was detected for *Hbt. salinarum* DSM 3754^T and GN101. The *pilB2* probe was expected to label a single fragment with a theoretical size of 4.3 kbp. This fragment was detected in *Hbt. salinarum* R1, DSM 3754^T, and the natural isolate GN101. *Hbt. salinarum* PHH4 contained a 7 kb fragment, and PHH1 and NRC-1 fragments larger than 10 kbp (Figure S2C). Inspection of the genomic region encoding the *pil-2* locus in NRC-1 (NC_002607) identified a 10 kbp insertion in *pilB2*. The 10 kb insert contains the insertion elements ISH11 (993 bp) and ISH2 (204 bp), and sequence similarities to genes encoding halophage proteins (CopG protein, phage terminase, primase, integrase). The insert inactivates the *pilB2* gene in NRC-1.

Blastn analyses using gene sequences of the transcriptional unit of *pil-1* and *pil-2* were performed to investigate the occurrence of these gene clusters in other archaeal genomes. The analyses indicated that the *pil-1* gene locus (4.4 kbp) is present in a broad range of other representatives of the *Halobacteriaceae* (Table S2). Regarding *Hfx. volcanii* D2 *pil-1* is related to the gene locus encompassing the *pilB3* and *pilC3* genes, required for the PilA pilus biosynthesis (Esquivel and Pohlschröder, 2014). The core unit *pilB1-C1* of the *pil-1* locus is also present in the genomes of methanogenic and hyperthermophilic euryarchaeota but not in the genomes of crenarchaeota or other archaeal phyla. The *pil-2* locus (6.9 kbp) is exclusively found in the genomes of other haloarchaeal strains not with other euryarchaeota (see Table S3). Low identities (query coverage 2% to 7%, *e*-value 8e-05 to 3e-04)

were found compared to high GC Gram⁺ actinobacteria, like the biofilm forming *Microbacterium xylanilyticum* (Kim et al., 2005).

COMPARATIVE qRT-PCR OF PLANKTONIC AND SURFACE ATTACHED CELLS

The expression of the assembly/motor-ATPase encoding genes (*flaI*, *pilB1*, and *pilB2*) was investigated by quantitative reverse transcription polymerase chain reaction (qRT-PCR) in surface attached cells as well as planktonic cells of *Hbt. salinarum* R1. Total RNA was isolated from planktonic cells during the exponential growth phase (OD₆₀₀ 0.3) and from surface attached cells (grown for 6 days). As control for the cDNA synthesis efficiency *bgaH* RNA (β -D-galactosidase, HVO_A0326, *Hfx. volcanii* DS2), was added prior to the cDNA generation and used as an external standard. The oligonucleotides used for qRT-PCR are listed in Table S1. To investigate the general transcriptional activity in surface attached cells two “housekeeping” genes, *rpoB1* (DNA-directed RNA polymerase subunit B', OE4741R) and *aef2* (translation elongation factor, OE4729R), were analyzed by qRT-PCR and normalized to the external control *bgaH* (Figure 3A). For *rpoB1* and *aef2* a 10-fold reduction of the expression was observed in adherent cells compared to planktonic cells. Thus, the overall transcriptional activity is reduced in surface-attached cells of *Hbt. salinarum* R1 after 6 days of incubation. The relative expression of the *pilB1* (OE2215R), *pilB2* (OE1347R), and *flaI* (OE2380R) genes was determined by qRT-PCR to investigate the transcriptional activity of the corresponding gene loci in planktonic vs. surface attached cells, the *rpoB1* was used as internal standard. An induced transcriptional activity was observed for *pilB1*, *pilB2*, and *flaI* in surface attached cells compared to

planktonic cells (Figure 3B). The relative expressions of *pilB1* and *pilB2* were 5.2-fold respectively 8.5-fold enhanced in adherent cells compared to planktonic cells. In contrast, *flaI* showed a 2.9-fold induced transcription in adherent cells compared to planktonic cells which was similar to the gene *fdx* (ferredoxin, OE4217R) with a 2.8-fold induced transcription. Ferredoxin is involved in cellular electron transfer and reported to be constitutively expressed (Twellmeyer et al., 2007). These data indicated a higher transcriptional activity of the *pil-1* and *pil-2* gene loci in surface attached cells. However, the number of cycles for the fluorescence signal to cross the threshold (*C_T*-value) was for *pilB1* 10 cycles lower (*C_T* 21 in surface attached cells) than for *pilB2* (*C_T* 31 in surface attached cells), indicating that the transcriptional activity of *pilB1* is higher compared to *pilB2*.

CONSTRUCTION AND PHENOTYPIC CHARACTERIZATION OF Δ *flaI* AND Δ *flaI*/ Δ *pilB1* MUTANTS

Two deletion mutant strains were constructed from *Hbt. salinarum* R1 to investigate a possible connection between the observed pili-like structures and the induced *pilB1* expression in surface attached cells. The first target for a deletion was *flaI* (OE2380R) to generate cells lacking the archaella, analogous to studies on the archaella encoding *fla*-gene cluster of *Hbt. salinarum* strain S9 showing that the deletion of the assembly ATPase corresponding gene *flaI* is sufficient to obtain non-archaellated cells (Patenge et al., 2001). In the second step the Δ *flaI* mutant strain was used to construct a Δ *flaI*/ Δ *pilB1* double mutant (OE2380R/OE2215R) to investigate whether the *pil-1* locus is involved in the biogenesis of the novel pili-like structures. A two step recombination method was used to construct markerless in-frame gene deletions in *Hbt. salinarum* R1 as described in Koch and Oesterhelt (2005). The construction of a Δ *pilB1* single mutant was not successful so far. The deletions of *flaI* and *flaI/pilB1* were identified by PCR using oligonucleotides flanking the genes of interest (Figure S3). The successful deletions were confirmed by sequencing of PCR amplified fragments derived from the deletion mutants, including the regions used for homologous recombination.

The presence of archaella with the parental strain and the deletion mutants was analyzed by a swimming motility assay. Semi solid plates with 0.3% (w/v) agar were inoculated in the center with 10 μ L of an exponential culture. The motility halo was measured in cm after 3 days of incubation (Figure 4). Motility was

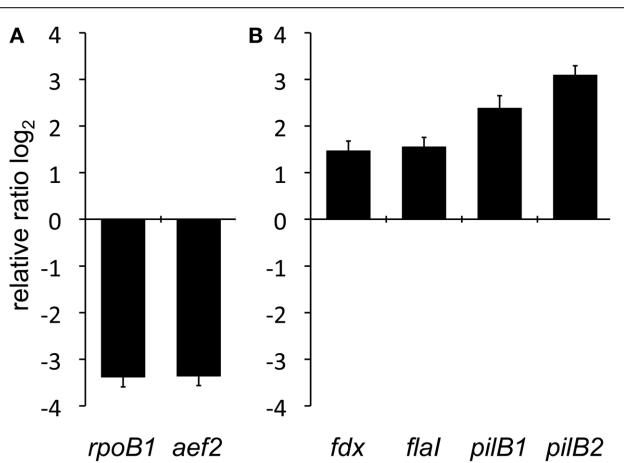


FIGURE 3 | Comparative qRT-PCR analyses of planktonic and surface attached cells. (A) Investigation of two representative housekeeping genes encoding RNA polymerase subunit B' (*rpoB1*) and the translation elongation factor 2 (*aef2*). Relative expression was normalized to external standard *bgaH* RNA. (B) Relative transcriptional quantification of the assembly ATPase encoding genes of the archaellum (*flaI*) and the type IV pilus biogenesis complexes *pil-1* (*pilB1*) and *pil-2* (*pilB2*) as well as the constitutively expressed ferredoxin gene (*fdx*). The bars represent the fold change of gene expression shown in base 2 logarithmic scale in adherent cells compared to the planktonic state, which is defined by the baseline.

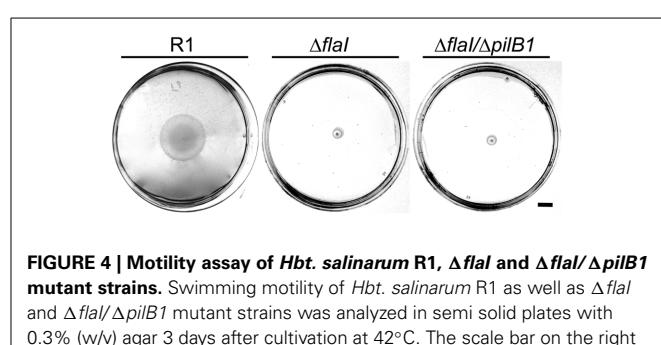


FIGURE 4 | Motility assay of *Hbt. salinarum* R1, Δ *flaI* and Δ *flaI*/ Δ *pilB1* mutant strains. Swimming motility of *Hbt. salinarum* R1 as well as Δ *flaI* and Δ *flaI*/ Δ *pilB1* mutant strains was analyzed in semi solid plates with 0.3% (w/v) agar 3 days after cultivation at 42°C. The scale bar on the right is 1 cm.

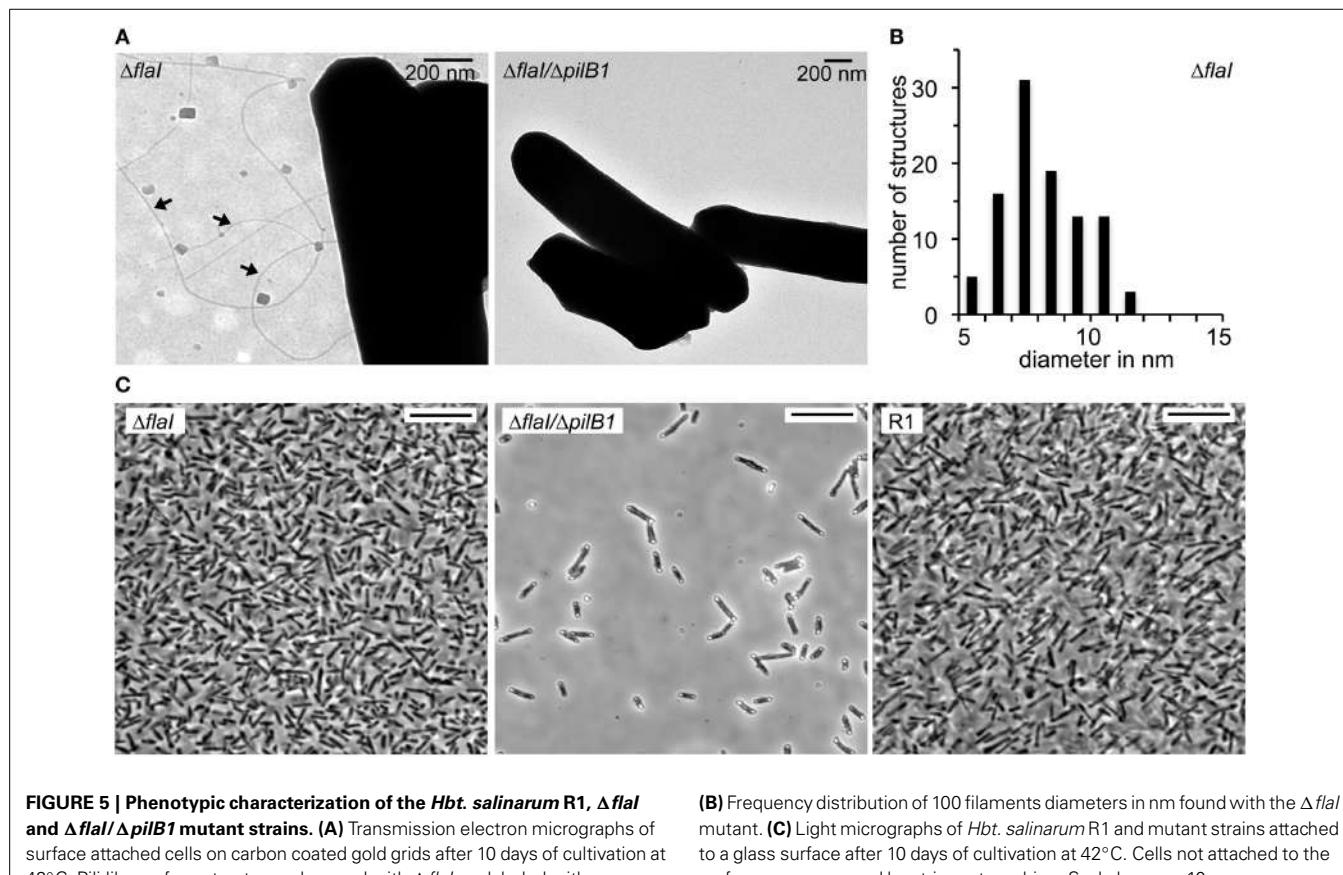
observed with the *Hbt. salinarum* R1 parental strain (motility halo 4.2 ± 1.2 cm, $n = 19$). In contrast, both deletion mutant strains $\Delta flaI/\Delta pilB1$ (0.7 ± 0.1 cm, $n = 19$) and $\Delta flaI$ (0.7 ± 0.2 cm, $n = 19$) did not show any motility, demonstrating that *flaI* is required for swimming. This is in accordance with the data by Patenge et al. (2001). However, with surface attached $\Delta flaI$ cells pili-like surface structures were observed by electron microscopy (Figure 5A). In comparison to the *Hbt. salinarum* R1 parental strain (see Figure 1) the number of surface structures observed with the cells of the $\Delta flaI$ mutant were reduced to 10–30%. For planktonic cells of the non-archaellated $\Delta flaI$ mutant a maximum of 2 to 3 structures were visible per cell (data not shown). The diameters of 100 surface structures found with the $\Delta flaI$ mutant were determined. The frequency distribution indicated a predominant diameter of 7 to 8 nm found with 31% of the pili-like structures (Figure 5B). The average diameter was calculated to be 7.6 ± 0.8 nm (calculated on the categories 6 to 9 nm, $n = 66$), which is identical to the average diameter determined previously for the novel pili-like structures (Figure 1) of the parental strain. No pili-like structures were observed with any cells of the $\Delta flaI/\Delta pilB1$ mutant, neither with planktonic nor surface attached cells of the entire culture (Figure 5A). These results demonstrated that *pilB1* is required for the assembly of the novel pili-like structures.

To determine a possible role of the novel pili-like structures in surface adhesion, cells of *Hbt. salinarum* R1 parental strain,

$\Delta flaI$ and $\Delta flaI/\Delta pilB1$ mutants were grown on cover slips for 10 days. Prior to microscopic analyses the overgrown cover slips were stringently washed to remove all non-surface attached cells. A monolayer of surface attached cells was observed for *Hbt. salinarum* R1 and the $\Delta flaI$ mutant strains. In contrast, the number of surface attached cells was strongly reduced with the $\Delta flaI/\Delta pilB1$ mutant where only a few cells attached to the surface (Figure 5C). To quantify these findings the surface coverage percentage of six photographs was determined for the parental and the mutant strains using ImageJ. The surface coverage of the non-archaellated $\Delta flaI$ mutants added up to $44\% \pm 3.6$ and was therefore higher [$t_{(12)} = 3.3, p < 0.01$] than the value determined for the parental strain *Hbt. salinarum* R1 ($36\% \pm 4.0$). The surface coverage observed with the non-archaellated and non-piliated cells of the $\Delta flaI/\Delta pilB1$ mutant was 10-fold reduced [$4\% \pm 1.3, t_{(12)} = 16.5, p < 0.001$].

DISCUSSION

Surface adhesion of the moderate halophile *Hfx. volcanii* DS2 is solely dependent on pili while archaella are not required (Tripepi et al., 2010; Esquivel et al., 2013). For the extremely halophilic archaeon *Hbt. salinarum* structures mediating adhesion have not been described yet. Cells of *Hbt. salinarum* possess polar archaella mediating the swimming motility (Alam and Oesterhelt, 1984). Their possible role in surface adhesion as well as the presence of further surface structures besides archaella were unclear. A first



hint for the existence of various filamentous structures came from TEM analyses of surface attached cells of *Hbt. salinarum* DSM 3754^T (Fröls et al., 2012).

In this report we determined that *Hbt. salinarum* R1 possesses additional surface structures crucial for surface adhesion. The pili-like structures of 7 to 8 nm are thinner compared to the Aap (archaeal adhesive pili) of *Sulfolobus acidocaldarius* with 10 nm and the type IV pili of *M. maripaludis* with 8.5 nm in diameter (Wang et al., 2008; Henche et al., 2012). Comparison of filamentous surface structures found with surface attached cells of the non-archaellated $\Delta flaI$ deletion mutant and the parental strain showed that these novel pili-like structures represent approximately one-third of the total cell appendages, illustrating that the archaella are the major surface structures in *Hbt. salinarum* R1. For planktonic cells the total number of these pili-like surface structures is reduced to a maximum of 2 to 3 structures per cell, which explains why the novel structures were not observed before.

Within the genome sequence of *Hbt. salinarum* R1 two putative type IV pili gene loci (*pil-1* and *pil-2*) were present. Transcriptional analyses of the *pil-1* locus identified a 5'-UTR with an alternative GUG translation start codon, which is found in 16% of the genes in *Hbt. salinarum* NRC-1 (Torarinsson et al., 2005). In addition, a putative Shine-Dalgarno element cGAGccGg [consensus sequence for *Hbt. salinarum* GGAGGUGA (Mankin et al., 1985)] was found located 9 nt upstream of the alternative GUG translation start codon, which is in good agreement with the distance for Shine-Dalgarno sequences present in *Hbt. salinarum* PHH1 (Sartorius-Neef and Pfeifer, 2004). Reporter gene studies with the gene encoding the *Hbt. salinarum* PHH1 gas vesicle protein H showed that this SD sequence influences the translation efficiency (Sartorius-Neef and Pfeifer, 2004). In the case of *pil-1* a putative stem-loop structure was predicted for the 5'-UTR, masking the Shine-Dalgarno sequence and the GUG translation start codon, which might effect the *pil-1* translation efficiency, too.

Comparative genome analyses showed that the core unit of the *pil-1* locus is also present in genomes of other euryarcheota and presumably represents a general system for surface adhesion. In contrast, the *pil-2* locus was solely present in haloarchaeal species. It is possible that this region was received from bacteria by lateral gene transfer, since the *pil-2* locus exhibits weak identities with genome sequences of actinobacteria. Comparative analyses of archaeal and bacterial genomes identified 2264 bacterial gene acquisitions by lateral gene transfer in *Archaea* and over 1000 in Haloarchaea, most of them derived from Actinobacteria (Nelson-Sathi et al., 2012, 2014). Representatives of these tolerate the high salt concentrations occurring in saline environments (Hamed et al., 2013).

Inspection of the particular genomic region in *Hbt. salinarum* NRC-1 identified an additional 10 kbp insert in the *pilB2* gene inactivating the *pil-2* locus. This region represents one of the 12 differences between the chromosomes of *Hbt. salinarum* R1 and NRC-1 (Pfeiffer et al., 2008b) and contains ISH elements and phage-specific sequences. It is flanked by an 8 bp duplication in NRC-1, that is only present as one copy in the *Hbt. salinarum* R1, DL-1, and *Hfx. volcanii* D2, but not with the genomic *pilB2* regions of *Natronomonas pharaonis*, or *Salinarchaeum* sp.,

suggesting a re-deletion of the region in some of the haloarchaeal strains and the possible reconstitution of *pilB2*.

The results of the quantitative reverse transcription PCR analyses implied a specific transcriptional up-regulation of the *pil-1* and *pil-2* loci in surface attached cells of *Hbt. salinarum* R1. Whether this is regulated directly by the surface contact or indirectly by changing growth conditions (for instance limited oxygen or nutrient supply) is not yet clear. Motility in *Hfx. volcanii* can be altered under certain medium conditions. Archaela-dependent swimming motility is observed with complex or defined medium containing casamino acids but not with defined medium (Triepi et al., 2010). For *Hbt. salinarum* R1 no significant changes in motility or surface adhesion were observed so far under anaerobic conditions or in response to the reduction of the nutrient source (Völkel and Fröls, unpublished results).

The phenotypic characterization of the $\Delta flaI$ and $\Delta flaI/\Delta pilB1$ deletion mutants indicated that a functional *pil-1* locus is essential for the formation of the novel pili-like filamentous structures. However, additional studies like the constructions of a *pilB1* single deletion strain and its phenotypic characterizations are required, as well as the preparation of the pili structures to analyze their pilin composition. It is questionable which pilins form the surface structure since no putative pilins are encoded in the surrounding genomic region of the *pil-1* locus. Within the genome sequence of *Hbt. salinarum* R1 more than 30 putative pilin encoding genes were predicted (Losensky and Fröls, unpublished results). Which ones of these putative pilins are expressed and assembled into pili structures is not known so far. For *Hfx. volcanii* D2 six pilins (PilA1-6) encoded in different genomic regions were identified. Each of the six pilins alone or in combination is sufficient for the biosynthesis of pili mediating surface adhesion (Esquivel et al., 2013).

The absence of pili-like structures and archaela leads to an impaired surface adhesion phenotype while in the absence of the archaela surface adhesion was increased. We conclude that the pili-like structures and not the archaela are crucial for surface adhesion of *Hbt. salinarum* R1. Nevertheless, we cannot exclude that the archaela are involved in the initial attachment to the surface, as assumed for *S. acidocaldarius* (Henche et al., 2012). In bacteria the flagella are important to mediate the initial reversible attachment to overcome the hydrodynamic boundary layer and repulsive forces. Moreover, a reduced flagellar rotation triggered the transcriptional activation of extracellular polymeric substances by modulation of the second messenger cyclic diguanylate (Petrova and Sauer, 2012).

On the contrary, in the absence of the archaela and pili-like structures low amounts of cells still attached to the surface, indicating that additional surface structures or mechanisms were present. Additional pili-like structures are possibly expressed in minor amounts by the *pil-2* locus, which encompasses three putative pilins. However, due to the internal stop codon within *pilC2* it is questionable whether the *pil-2* locus is functional. Surface adhesion might also be mediated by amyloid adhesins (i.e., short fibers formed by an extracellular protein) present with several bacterial biofilm forming species (Larsen et al., 2007). In *Archaea* amyloid proteins are present in the large dense cell clusters and

tower like biofilm structures of *Hfx. volcanii* D2 (Chimileski et al., 2014).

Overall the phenotypic characterization of the deletion mutants indicated that the novel pili-like structures are crucial for surface adhesion of *Hbt. salinarum* R1, but the structures themselves still remain to be characterized on the genetic and protein levels.

ACKNOWLEDGMENTS

Arnulf Kletzin and Friedhelm Pfeiffer are thanked for fruitful discussions. We thank the reviewers for valuable comments on the manuscript.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <http://www.frontiersin.org/journal/10.3389/fmicb.2014.00755/abstract>

REFERENCES

- Alam, M., and Oesterhelt, D. (1984). Morphology, function and isolation of halobacterial flagella. *J. Mol. Biol.* 176, 459–475. doi: 10.1016/0022-2836(84)90172-4
- Bleiholder, A., Frommherz, R., Teufel, K., and Pfeifer, F. (2012). Expression of multiple tfb genes in different *Halobacterium salinarum* strains and interaction of TFB with transcriptional activator GvpE. *Arch. Microbiol.* 194, 269–279. doi: 10.1007/s00203-011-0756-z
- Chan, P. P., Holmes, A. D., Smith, A. M., Tran, D., and Lowe, T. M. (2012). The UCSC Archaeal Genome Browser: 2012 update. *Nucleic Acids Res.* 40, D646–D652. doi: 10.1093/nar/gkr990
- Chimileski, S., Franklin, M. J., and Papke, R. T. (2014). Biofilms formed by the archaeon *Haloferax volcanii* exhibit cellular differentiation and social motility, and facilitate horizontal gene transfer. *BMC Biol.* 12:65. doi: 10.1186/s12915-014-0065-5
- Chomczynski, P., and Sacchi, N. (2006). The single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction: twenty-something years on. *Nat. Protoc.* 1, 581–585. doi: 10.1038/nprot.2006.83
- Cohen-Krausz, S., and Trachtenberg, S. (2002). The structure of the archeabacterial flagellar filament of the extreme halophile *Halobacterium salinarum* RIM1 and its relation to eubacterial flagellar filaments and type IV pili. *J. Mol. Biol.* 321, 383–395. doi: 10.1016/S0022-2836(02)00616-2
- Dyall-Smith, M. L. (2008). *The Halohandbook: Protocols for Halobacterial Genetics*. Available online at: http://www.haloarchaea.com/resources/halohandbook_Halohandbook_2008_v7.pdf
- Ebert, K., Goebel, W., and Pfeifer, F. (1984). Homologies between heterogeneous extrachromosomal DNA populations of *Halobacterium halobium* and four new isolates. *Mol. Gen. Genet.* 194, 91–97. doi: 10.1007/BF00383502
- Esquivel, R. N., and Pohlschröder, M. (2014). A conserved type IV pilin signal peptide H-domain is critical for the post-translational regulation of flagella-dependent motility. *Mol. Microbiol.* 93, 494–504. doi: 10.1111/mmi.12673
- Esquivel, R. N., Xu, R., and Pohlschröder, M. (2013). Novel archaeal adhesion pilins with a conserved N terminus. *J. Bacteriol.* 195, 3808–3818. doi: 10.1128/JB.00572-13
- Fröls, S., Dyall-Smith, M., and Pfeifer, F. (2012). Biofilm formation by haloarchaea. *Environ. Microbiol.* 14, 3159–3174. doi: 10.1111/j.1462-2920.2012.02895.x
- Gruber, C., Legat, A., Pfaffenhuemer, M., Radax, C., Weidler, G., Busse, H. J., et al. (2004). *Halobacterium noricense* sp. nov., an archaeal isolate from a bore core of an alpine Permian salt deposit, classification of *Halobacterium* sp. NRC-1 as a strain of *Hbt. salinarum* and emended description of *Hbt. salinarum*. *Extremophiles* 8, 431–439. doi: 10.1007/s00792-004-0403-6
- Hamedí, J., Mohammadi Panah, F., and Ventosa, A. (2013). Systematic and biotechnological aspects of halophilic and halotolerant actinomycetes. *Extremophiles* 17, 1–13. doi: 10.1007/s00792-012-0493-5
- Henché, A. L., Ghosh, A., Yu, X., Jeske, T., Egelman, E., and Albers, S. V. (2012). Structure and function of the adhesive type IV pilus of *Sulfolobus acidocaldarius*. *Environ. Microbiol.* 14, 3188–3202. doi: 10.1111/j.1462-2920.2012.02898.x
- Henneberger, R., Moissl, C., Amann, T., Rudolph, C., and Huber, R. (2006). New insights into the lifestyle of the cold-loving SM1 euryarchaeon: natural growth as a monospecies biofilm in the subsurface. *Appl. Environ. Microbiol.* 72, 192–199. doi: 10.1128/AEM.72.1.192-199.2006
- Jarrell, K. F., and Albers, S. V. (2012). The archaellum: an old motility structure with a new name. *Trends Microbiol.* 20, 307–312. doi: 10.1016/j.tim.2012.04.007
- Jarrell, K. F., Stark, M., Nair, D. B., and Chong, J. P. (2011). Flagella and pili are both necessary for efficient attachment of *Methanococcus maripaludis* to surfaces. *FEMS Microbiol. Lett.* 319, 44–50. doi: 10.1111/j.1574-6968.2011.02264.x
- Kim, K. K., Park, H. Y., Park, W., Kim, I. S., and Lee, S. T. (2005). *Microbacterium xylanilyticum* sp. nov., a xylan-degrading bacterium isolated from a biofilm. *Int. J. Syst. Evol. Microbiol.* 55, 2075–2079. doi: 10.1099/ijss.0.63706-0
- Koch, M. K., and Oesterhelt, D. (2005). MpcT is the transducer for membrane potential changes in *Halobacterium salinarum*. *Mol. Microbiol.* 55, 1681–1694. doi: 10.1111/j.1365-2958.2005.04516.x
- Krogh, A., Larsson, B., Von Heijne, G., and Sonnhammer, E. L. (2001). Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J. Mol. Biol.* 305, 567–580. doi: 10.1006/jmbi.2000.4315
- Larsen, P., Nielsen, J. L., Dueholm, M. S., Wetzel, R., Otzen, D., and Nielsen, P. H. (2007). Amyloid adhesins are abundant in natural biofilms. *Environ. Microbiol.* 9, 3077–3090. doi: 10.1111/j.1462-2920.2007.01418.x
- Lassak, K., Ghosh, A., and Albers, S. V. (2012). Diversity, assembly and regulation of archaeal type IV pili-like and non-type-IV pili-like surface structures. *Res. Microbiol.* 163, 630–644. doi: 10.1016/j.resmic.2012.10.024
- Letunic, I., Doerks, T., and Bork, P. (2012). SMART 7: recent updates to the protein domain annotation resource. *Nucleic Acids Res.* 40, D302–D305. doi: 10.1093/nar/gkr931
- Mankin, A. S., Kagramanova, V. K., Teterina, N. L., Rubtsov, P. M., Belova, E. N., Kopylov, A. M., et al. (1985). The nucleotide sequence of the gene coding for the 16S rRNA from the archaeabacterium *Halobacterium halobium*. *Gene* 37, 181–189. doi: 10.1016/0378-1119(85)90271-9
- Moissl, C., Rachel, R., Briegel, A., Engelhardt, H., and Huber, R. (2005). The unique structure of archaeal ‘hami’: highly complex cell appendages with nano-grappling hooks. *Mol. Microbiol.* 56, 361–370. doi: 10.1111/j.1365-2958.2005.04294.x
- NCBI Resource Coordinators (2014). Database resources of the National Center for Biotechnology Information. *Nucleic Acids Res.* 42, D7–D17. doi: 10.1093/nar/gkt1146
- Näther, D. J., Rachel, R., Wanner, G., and Wirth, R. (2006). Flagella of *Pyrococcus furiosus*: multifunctional organelles, made for swimming, adhesion to various surfaces, and cell-cell contacts. *J. Bacteriol.* 188, 6915–6923. doi: 10.1128/JB.00527-06
- Nelson-Sathi, S., Dagan, T., Landan, G., Janssen, A., Steel, M., McInerney, J. O., et al. (2012). Acquisition of 1,000 eubacterial genes physiologically transformed a methanogen at the origin of Haloarchaea. *Proc. Natl. Acad. Sci. U.S.A.* 109, 20537–20542. doi: 10.1073/pnas.1209119109
- Nelson-Sathi, S., Sousa, F. L., Roettger, M., Lozada-Chavez, N., Thiergart, T., Janssen, A., et al. (2014). Origins of major archaeal clades correspond to gene acquisitions from bacteria. *Nature* 513:60. doi: 10.1038/nature13805. [Epub ahead of print].
- Patenge, N., Berendes, A., Engelhardt, H., Schuster, S. C., and Oesterhelt, D. (2001). The fla gene cluster is involved in the biogenesis of flagella in *Halobacterium salinarum*. *Mol. Microbiol.* 41, 653–663. doi: 10.1046/j.1365-2958.2001.02542.x
- Peabody, C. R., Chung, Y. J., Yen, M. R., Vidal-Inigliardi, D., Pugsley, A. P., and Saier, M. H. Jr. (2003). Type II protein secretion and its relationship to bacterial type IV pili and archaeal flagella. *Microbiology* 149, 3051–3072. doi: 10.1099/mic.0.26364-0
- Petrova, O. E., and Sauer, K. (2012). Sticky situations: key components that control bacterial surface attachment. *J. Bacteriol.* 194, 2413–2425. doi: 10.1128/JB.00003-12
- Pfeifer, F., and Blasie, U. (1989). Insertion elements and deletion formation in a halophilic archaeabacterium. *J. Bacteriol.* 171, 5135–5140.
- Pfeifer, F., Blasie, U., and Ghahraman, P. (1988). Dynamic plasmid populations in *Halobacterium halobium*. *J. Bacteriol.* 170, 3718–3724.
- Pfeiffer, F., Broicher, A., Gillich, T., Klee, K., Mejia, J., Rampp, M., et al. (2008a). Genome information management and integrated data analysis with HaloLex. *Arch. Microbiol.* 190, 281–299. doi: 10.1007/s00203-008-0389-z
- Pfeiffer, F., Schuster, S. C., Broicher, A., Falb, M., Palm, P., Rodewald, K., et al. (2008b). Evolution in the laboratory: the genome of *Halobacterium salinarum*.

- strain R1 compared to that of strain NRC-1. *Genomics* 91, 335–346. doi: 10.1016/j.ygeno.2008.01.001
- Pohlschröder, M., Ghosh, A., Tripepi, M., and Albers, S. V. (2011). Archaeal type IV pilus-like structures—evolutionarily conserved prokaryotic surface organelles. *Curr. Opin. Microbiol.* 14, 357–363. doi: 10.1016/j.mib.2011.03.002
- Sartorius-Neef, S., and Pfeifer, F. (2004). *In vivo* studies on putative Shine-Dalgarno sequences of the halophilic archaeon *Halobacterium salinarum*. *Mol. Microbiol.* 51, 579–588. doi: 10.1046/j.1365-2958.2003.03858.x
- Schmittgen, T. D., and Livak, K. J. (2008). Analyzing real-time PCR data by the comparative C(T) method. *Nat. Protoc.* 3, 1101–1108. doi: 10.1038/nprot.2008.73
- Schneider, K. L., Pollard, K. S., Baertsch, R., Pohl, A., and Lowe, T. M. (2006). The UCSC Archaeal Genome Browser. *Nucleic Acids Res.* 34, D407–D410. doi: 10.1093/nar/gkj134
- Schultz, J., Milpetz, F., Bork, P., and Ponting, C. P. (1998). SMART, a simple modular architecture research tool: identification of signaling domains. *Proc. Natl. Acad. Sci. U.S.A.* 95, 5857–5864. doi: 10.1073/pnas.95.11.5857
- Streif, S., Staudinger, W. F., Marwan, W., and Oesterhelt, D. (2008). Flagellar rotation in the archaeon *Halobacterium salinarum* depends on ATP. *J. Mol. Biol.* 384, 1–8. doi: 10.1016/j.jmb.2008.08.057
- Szabo, Z., Stahl, A. O., Albers, S. V., Kissinger, J. C., Driessen, A. J., and Pohlschröder, M. (2007). Identification of diverse archaeal proteins with class III signal peptides cleaved by distinct archaeal preprotein peptidases. *J. Bacteriol.* 189, 772–778. doi: 10.1128/JB.01547-06
- Thoma, C., Frank, M., Rachel, R., Schmid, S., Näther, D., Wanner, G., et al. (2008). The Mth60 fimbriae of *Methanothermobacter thermoautotrophicus* are functional adhesins. *Environ. Microbiol.* 10, 2785–2795. doi: 10.1111/j.1462-2920.2008.01698.x
- Torarinsson, E., Klenk, H. P., and Garrett, R. A. (2005). Divergent transcriptional and translational signals in Archaea. *Environ. Microbiol.* 7, 47–54. doi: 10.1111/j.1462-2920.2004.00674.x
- Tripepi, M., Imam, S., and Pohlschröder, M. (2010). *Haloferax volcanii* flagella are required for motility but are not involved in PilD-dependent surface adhesion. *J. Bacteriol.* 192, 3093–3102. doi: 10.1128/JB.00133-10
- Twellmeyer, J., Wende, A., Wolfertz, J., Pfeiffer, F., Panhuysen, M., Zaigler, A., et al. (2007). Microarray analysis in the archaeon *Halobacterium salinarum* strain R1. *PLoS ONE* 2:e1064. doi: 10.1371/journal.pone.0001064
- Wang, C., Lee, J., Deng, Y., Tao, F., and Zhang, L. (2012). ARF-TSS: an alternative method for identification of transcription start site in bacteria. *Biotechniques* 1–3. doi: 10.2144/000113858. [Epub ahead of print].
- Wang, Y. A., Yu, X., Ng, S. Y., Jarrell, K. F., and Egelman, E. H. (2008). The structure of an archaeal pilus. *J. Mol. Biol.* 381, 456–466. doi: 10.1016/j.jmb.2008.06.017
- Zolghadr, B., Klingl, A., Koerdt, A., Driessen, A. J., Rachel, R., and Albers, S. V. (2010). Appendage-mediated surface adherence of *Sulfolobus solfataricus*. *J. Bacteriol.* 192, 104–110. doi: 10.1128/JB.01061-09
- Zuker, M. (2003). Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res.* 31, 3406–3415. doi: 10.1093/nar/gkg595

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

Received: 30 September 2014; accepted: 11 December 2014; published online: 13 January 2015.

Citation: Losensky G, Vidakovic L, Klingl A, Pfeifer F and Fröls S (2015) Novel pilus-like surface structures of *Halobacterium salinarum* strain R1 are crucial for surface adhesion. *Front. Microbiol.* 5:755. doi: 10.3389/fmicb.2014.00755

This article was submitted to Microbial Physiology and Metabolism, a section of the journal *Frontiers in Microbiology*.

Copyright © 2015 Losensky, Vidakovic, Klingl, Pfeifer and Fröls. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Pyrococcus furiosus flagella: biochemical and transcriptional analyses identify the newly detected *flaB0* gene to encode the major flagellin

Daniela J. Näther-Schindler^{1,2*}, Simone Schopf^{1,3}, Annett Bellack^{1*}, Reinhard Rachel¹ and Reinhard Wirth¹

¹ Institute of Microbiology and Archaea Center, University of Regensburg, Regensburg, Germany

² Plant Development, Department of Biology I, Biocenter of the Ludwig Maximilian University of Munich, Planegg-Martinsried, Germany

³ Department of Biology - Section Environmental Microbiology, Technical University Freiberg, Freiberg, Germany

Edited by:

Mechthild Pohlschroder, University of Pennsylvania, USA

Reviewed by:

Dirk Linke, Max Planck Society, Germany

Friedhelm Pfeiffer,
Max-Planck-Institute of Biochemistry, Germany

***Correspondence:**

Daniela J. Näther-Schindler, Plant Development, Department of Biology I, Biocenter of the Ludwig Maximilian University of Munich, Großhadernerstr. 2-4, 82152 Planegg-Martinsried, Germany
e-mail: daniela.naether@gmx.de;
Annett Bellack, Institute of Microbiology and Archaea Center, University of Regensburg, Universitätsstr. 31, 93053 Regensburg, Germany
e-mail: annett.bellack@ur.de

We have described previously that the flagella of the Euryarchaeon *Pyrococcus furiosus* are multifunctional cell appendages used for swimming, adhesion to surfaces and formation of cell-cell connections. Here, we characterize these organelles with respect to their biochemistry and transcription. Flagella were purified by shearing from cells followed by CsCl-gradient centrifugation and were found to consist mainly of a ca. 30 kDa glycoprotein. Polymerization studies of denatured flagella resulted in an ATP-independent formation of flagella-like filaments. The N-terminal sequence of the main flagellin was determined by Edman degradation, but none of the genes in the complete genome code for a protein with that N-terminus. Therefore, we resequenced the respective region of the genome, thereby discovering that the published genome sequence is not correct. A total of 771 bp are missing in the data base, resulting in the correction of the previously unusual N-terminal sequence of flagellin FlaB1 and in the identification of a third flagellin. To keep in line with the earlier nomenclature we call this *flaB0*. Very interestingly, the previously not identified *flaB0* codes for the major flagellin. Transcriptional analyses of the revised flagellar operon identified various different cotranscripts encoding only a single protein in case of FlaB0 and FlaJ or up to five proteins (FlaB0-FlaD). Analysing the RNA of cells from different growth phases, we found that the length and number of detected cotranscript increased over time suggesting that the flagellar operon is transcribed mostly in late exponential and stationary growth phase.

Keywords: archaeal flagella, *Pyrococcus furiosus*, Fla proteins, major flagellin, FlaB0, transcriptional analyses

INTRODUCTION

Archaea have been shown to possess various distinct types of cell surface appendages (reviewed e.g., by Ng et al., 2008 or Jarrell et al., 2013) of which flagella are the best characterized ones. Superficially, these structures seem to be very similar to bacterial flagella; however, analyses of the ultrastructure, the involved proteins and the biosynthesis machinery identified fundamental differences (see e.g., Thomas et al., 2001 or Ghosh and Albers, 2011 for reviews on archaeal flagella). Based on these findings a renaming of archaeal flagella into archaella has been suggested (Jarrell and Albers, 2012) but has been questioned because of serious flaws as consequences of such a nomenclature (Wirth, 2012). With respect to the ongoing discussion about a name reflecting the function and uniqueness of these cell surface structures, we decided to retain the term flagella.

Archaeal flagella are built in their part emanating from the cell from mostly more than one protein, the so-called flagellins. *In silico* analyses of many different archaeal genomes found that the genes encoding flagellins (*flaA* and/or *flaB*) are arranged in an operon together with additional proteins assumed

to be motor and anchoring components. In Euryarchaeota, the operon comprises the *fla*-associated genes *flaC* to *flaJ*, whereas in Crenarchaeota, *flaCDE* are missing and *flaX* is present (which is absent from Euryarchaeota). Interestingly, neither these genes nor the corresponding proteins show any similarities to their bacterial counterparts (Jarrell et al., 2013). Hence, our current knowledge of the assembly of archaeal flagella is based on genetic analyses. Deletion studies in *Halobacterium salinarum*, *Methanococcus maripaludis*, and *Sulfolobus acidocaldarius* have shown that all of the *fla*-associated genes are necessary for proper assembly and function of flagella (Patenge et al., 2001; Chaban et al., 2007; Lassak et al., 2012). Flagellins are synthesized as preproteins; their signal peptide is removed by FlaK and an N-linked glycan is attached by the oligosaccharyltransferase AglB. These posttranslationally modified subunits are supposed to be incorporated at the base of the growing non-tubular structure involving the activity of the ATPase FlaI and the conserved membrane protein FlaJ (Jarrell et al., 2013).

In addition to the mentioned mesophilic and thermophilic species, the Euryarchaeon *Pyrococcus furiosus* is a model organism

for hyperthermophilic Archaea. Despite the availability of a genetic system (Waege et al., 2010; Lipscomb et al., 2011) and numerous –omics-based approaches (for a summary see Bridger et al., 2012), data on its flagella are restricted to a publication of our group (Näther et al., 2006). We have shown that *P. furiosus* uses its flagella not only for swimming, but is able to adhere with these cell surface organelles to specific surfaces including cells of its own species, thereby forming biofilms. In addition, also the formation of cell-cell connections via cable-like aggregated flagella was observed (Näther et al., 2006). In further studies we have demonstrated that also the flagella of the fastest organisms on earth (Herzog and Wirth, 2012), namely the Euryarchaeon *Methanocaldococcus villosovorus*, can be used for adhesion to various surfaces; again, also formation of cell-cell connections by flagella was described (Bellack et al., 2011). Beside the functional studies, we have proven that the flagella of *P. furiosus* consist of mainly one glycoprotein (Näther et al., 2006), but the N-terminal sequence we identified did not match perfectly to any protein annotated in the published genome sequence (Robb et al., 2001). Therefore, we resequenced the flagellar operon in this study and discovered that a 771 bp segment was missing previously in the genome sequence. On this segment, we identified an in-frame start codon for the *flaB1* gene and a new gene, *flaB0*, which encodes the major flagellin. In addition we performed *in vitro* polymerization studies of flagellin monomers and analyzed transcription of the revised flagellar operon of *P. furiosus*.

MATERIALS AND METHODS

GROWTH OF *P. FURIOSUS*, FLAGELLA PREPARATION, AND REPOLYMERIZATION OF DENATURATED FLAGELLINS

Growth of cells and preparation of flagella therefrom by shearing followed by CsCl-gradient centrifugation was as described (Näther et al., 2006). For repolymerization studies, flagella were isolated as follows: cells were lysed by osmotic shock; membranes were then harvested by differential centrifugation and solubilized overnight at room temperature by addition of 0.5% n-dodecyl β-D-maltopyranoside (DDM). After purification by CsCl-gradient centrifugation (Näther et al., 2006), flagella were denatured by addition of SDS to a final concentration of 1% and heating at 100°C for 30 min. The samples were dialyzed four times for 1 h each against 5 mM HEPES buffer (pH = 7.0), followed by an overnight dialysis. These samples were incubated in tightly closed vials at various temperatures (8°C, 37°C, 60°C, and 90°C) with/without addition of 1 mM ATP. To avoid evaporation samples incubated at 60°C and 90°C were overlaid with chill-out liquid wax (*Biorad Laboratories GmbH; Munich, Germany*). Aliquots were analyzed by SDS-PAGE after 1, 2, and 6 days, without heating prior to loading.

N-terminal sequencing by Edman degradation was performed by the protein analytic facility of the Biochemistry Department of the University of Regensburg.

DNA ISOLATION AND SEQUENCING

For DNA isolation cells were collected from 40 ml cultures by centrifugation and DNA was isolated according to Bellack et al. (2011). Alternatively, cells were resuspended in 0.8 ml TNE buffer (100 mM Tris/Cl; 50 mM NaCl; 50 mM EDTA; pH = 8.0). Lysis

of cells was by addition of 0.1 ml of 10% SDS plus 0.1 ml of 10% N-lauroylsarcosine and cautious mixing. After addition of 10 µl RNase (10 mg/ml) and incubation for 15 min at room-temperature 50 µl proteinase K (20 mg/ml) was added and the sample heated for 1 h to 55°C. After repeated phenol extractions DNA was precipitated from the water phase with 800 µl 2-propanol, the pellet was washed with 70% ice-cold ethanol and dissolved in water.

For resequencing of the genomic region around the *flaB2* gene, primer walking analyses were performed using primers 353420_f (5'-ATGGAAAAACTAGAGAACCGTTG-3'), 352920_f (5'-TGGCTCAGCTTCACCAGC-3'), 352542_f (5'-AATATTAG ATGAGGGATTCAAGTTAA-3'), 352509_f (5'-GGATTATG GAAAGGCAATTCTTCTC-3'), 353159_r (5'-TATTGCCATCTT AACTATGGTCCC-3'), and 351761_r (5'-ATCACATTATACTCAA ATGTTGGGG-3'). Primer numbers refer to the binding position in the original genome sequence (Robb et al., 2001).

PCR reactions using primers 353483_f (5'-GGATTATGGAAA GGCAATTCTTCTC-3') and 351761_r were used to analyze genomic DNA from various *P. furiosus* strains for the presence of the *flaB0* gene.

GENERATION OF ANTIBODIES

To raise specific antibodies against each flagellin, the respective central region (Figure 2, gray sequences) was amplified via PCR using primers FlaB0-MT_f (5'-GGATCCGAGAAAACAGCATATCACAAAGGA-3'), FlaB0-MT_r (5'-AAGCTTACCGAAAACCTCCATTCCCT-3'), FlaB1-MT_f (5'-GGATCCAGTGGAGAACTGTACACTGGAAAGA-3'), FlaB1-MT_r (5'-AAGCTTGCTCTATAATTAAAGACATCATCCGT-3'), FlaB2-MT_f (5'-GCAGCCATATGAGGTATTACGATCCA-5'), and FlaB2-MT_r (5'-GAAGGGGATCCTCAGTAGAGGTTCCA-5'). Fragments were cloned into the low-copy number plasmid pQE30 (QIAGEN; Hilden, Germany) to avoid instable clones. The plasmid was transformed into the *E. coli* expression strain BL21 Star(DE3)pLysS; the corresponding ~6 kDa peptides could be purified after induction with IPTG via Ni-chelate chromatography and were used to immunize rabbits (Davids Biotechnologie; Regensburg, Germany).

ISOLATION OF RNA, REVERSE TRANSCRIPTION PCR, AND NORTHERN BLOT EXPERIMENTS

500 ml of exponentially growing cells (~1 × 10⁷ cells/ml and 4–5 × 10⁷ cells/ml; direct cell counting using a Thoma counting chamber) or stationary cells (~2 × 10⁸ cells/ml) were collected by centrifugation and resuspended in 1 ml of Trizol™ each. After incubation for 10 min at room temperature 0.2 ml chloroform was added and the lysate was cautiously mixed. After centrifugation (12,000 × g, 15 min, 4°C) the water phase was collected, 0.5 ml ice-cold 2-propanol was added and precipitation was for 30 min or overnight at –20°C. RNA was collected by centrifugation as above and the pellet was washed with 1 ml of ice-cold 70% ethanol. The pellet was air dried, resuspended in 100 µl H₂O and 90 µl *DNase I Incubation Buffer* plus 10 µl *DNase I* (both from *High Pure RNA Isolation Kit*, Roche Diagnostics GmbH; Mannheim, Germany) were added. After 15 min incubation at room temperature further processing,

including a phenol/chloroform extraction and RNA precipitation, was as recommended in the *High Pure RNA Isolation Kit*.

To detect specific mRNA transcripts, mRNA was transcribed into cDNA using the *Super Script II reverse Transcriptase* protocol as suggested by the supplier (*Invitrogen GmbH; Karlsruhe, Germany*). The various cDNAs were amplified via PCR using different combination of primers which were designed for each gene of the flagellar operon; primers are given in Supplementary Figure S1. In each case a negative control without addition of cDNA was included; the positive control included PCR reactions using genomic DNA.

For Northern Blotting, RNA probes were labeled with digoxigenin using the *DIG Northern Starter Kit* (*Roche Diagnostics GmbH; Mannheim, Germany*). Gel electrophoresis, northern blot, hybridization and detection was as recommended in the manufacturer's instructions.

TEM ANALYSES

Preparation of specimens by negative staining and for immunolabeling was as described earlier (Näther et al., 2006; Rachel et al., 2010).

RESULTS

IDENTIFICATION OF *flaB0*, A THIRD FLAGELLAR GENE, IN THE GENOME OF *P. FURIOSUS*

Purification of flagella via isopycnic CsCl-gradient centrifugation and analysis by SDS-PAGE identified one major glycoprotein, whose N-terminal amino acid sequence was determined to be AVGIGTLIVF (Näther et al., 2006). Sequence alignments illustrated that this N-terminal sequence did not perfectly match to the annotated flagellins of *P. furiosus* or to any of the other

proteins translated from the published genome (Robb et al., 2001). More precisely, the N-terminus of protein FlaB2 should read AIGIGTLIVE, but Edman degradation of the major flagellin never indicates any heterogeneity at position 2. In case of protein FlaB1 we found that the published sequence lacks the motif AIGIGTLIVFIAM, which is very highly conserved in all flagellins annotated in the publically available genomes of the genus *Pyrococcus*. However, this motif is encoded directly in front of the annotated *flaB1* gene but misses an upstream in-frame start codon. Based on these findings we decided to resequence the genome region that codes for the flagellins. Indeed, we identified a major mistake in that part of the published *P. furiosus* genome sequence: a total of 771 bp are missing. By combining this new sequence with the published genome (Robb et al., 2001), the sequence of the *flaB1* gene now contains a proper start codon and its N-terminus becomes highly similar to other flagellins. In addition, we detected another ORF coding for a third flagellin which we call *flaB0* to keep in line with the existing nomenclature. As a consequence, the flagellar operon of *P. furiosus* was revised (Figure 1). The missing genomic sequence containing the annotation of the *flaB0* gene/FlaB0 protein was submitted to NCBI BankIt; the corresponding GenBank number is KM892551.

FlaB0, THE MAJOR FLAGELLIN OF *P. FURIOSUS*

All flagella isolated from *P. furiosus* by different methods (shearing, DDM treatment, Triton X-114 treatment according to Kalmokoff et al., 1988) over a period of nearly 10 years were composed of one major flagellin as indicated by SDS-PAGE. The finding that the N-terminal amino acid sequence of this protein unambiguously was AVGIGTLIVF suggests that the newly

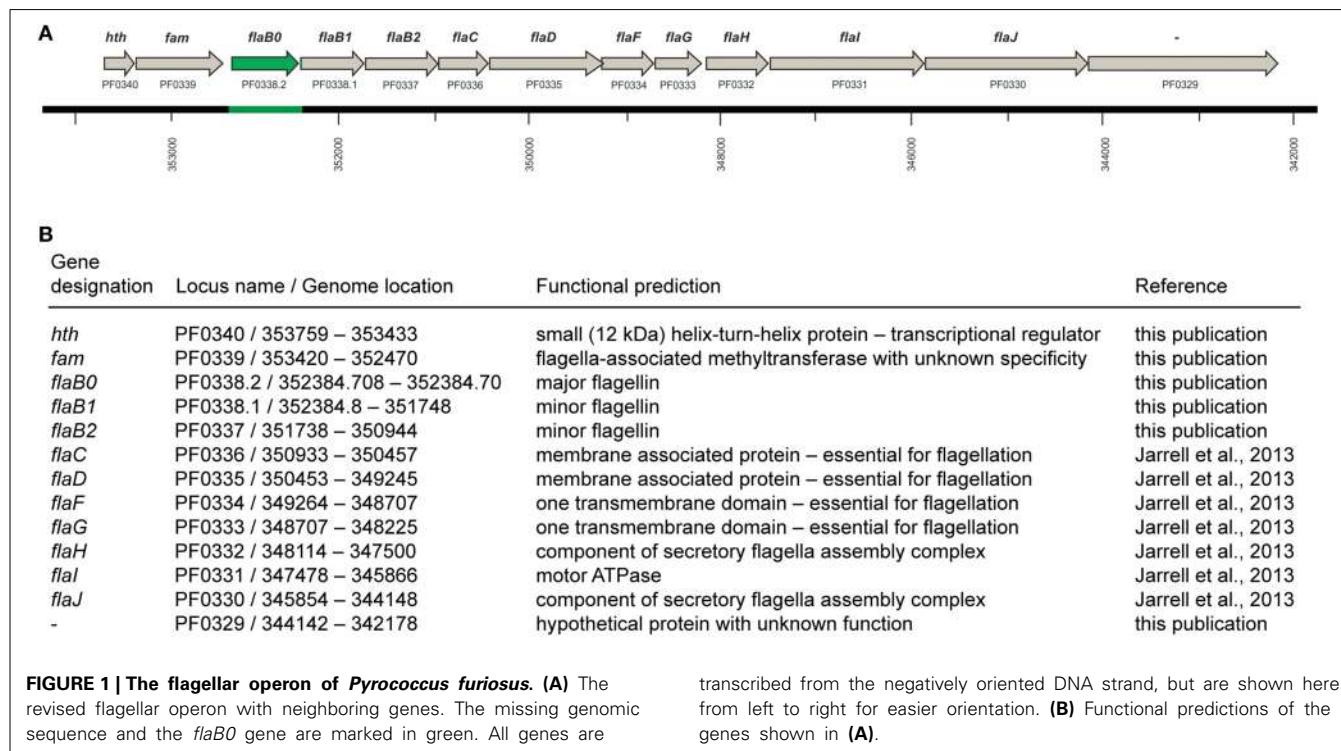


FIGURE 1 | The flagellar operon of *Pyrococcus furiosus*. **(A)** The revised flagellar operon with neighboring genes. The missing genomic sequence and the *flaB0* gene are marked in green. All genes are

transcribed from the negatively oriented DNA strand, but are shown here from left to right for easier orientation. **(B)** Functional predictions of the genes shown in **(A)**.

detected FlaB0 is the major flagellin of *P. furiosus* whereas FlaB1 and FlaB2 are only minor flagellins.

To ask for the presence of the two minor flagellins in our flagella preparations, we raised specific antibodies against all three flagellins. Because of the highly conserved N- and C-terminal part of FlaB0, FlaB1, and FlaB2, we subcloned the unique central part of each flagellin (gray sequences in **Figure 2**) and used the peptides for immunization of rabbits. The resulting antisera had a low titer, especially for the FlaB2-peptide. Western blots (data not shown) using these antisera proved that all three flagellins are present in the protein band at around 30 kDa. In addition, purified antibodies were used to immuno-label flagella preparations and cells adherent to carbon-coated gold grids for TEM. Again, we could show that antibodies against sheared *P. furiosus* flagella detach adherent cells from their solid support as described earlier (Näther et al., 2006). Some single cells, however, remained on the grid and their flagella were clearly labeled over their whole length. In contrast, no signals were detected using any of the antibodies against the recombinant flagellin middle parts. Specific antibodies against FlaB1 and FlaB2 reacted mostly with the ends of purified flagella (data not shown).

DENATURATED FLAGELLA CAN REPOLYMERIZE INTO FILAMENTOUS STRUCTURES

We furthermore asked if “native flagella” could be repolymerized from denatured flagellins, spontaneously without energy and

without a template. For depolymerization a flagella preparation was denatured by addition of SDS and incubation at 100°C, followed by extensive dialysis. We found that heat treatment is necessary for complete denaturation, otherwise SDS-PAGE shows the presence of minor protein bands with molecular weights of ca. 60, 90, and >100 kDa. Two dimensional SDS-PAGE clearly proved that these bands could be dissociated into the ~30 kDa flagellin monomers (data not shown).

Flagellins from denatured flagella were incubated at different temperatures to analyze their potential to repolymerize. Incubation for 1 day or longer at temperatures higher than 60°C resulted in aggregation of the ~30 kDa flagellins into high-molecular weight polymers, forming in part also filamentous structures as proven by TEM (**Figure 3**). Comparing these filaments to native flagella, we found the diameter to be smaller, and no helical ultrastructure was present. Addition of ATP to the samples had no influence on the formation of aggregates or filaments (data not shown).

CONSERVATION OF *flaB0* IN VARIOUS *P. FURIOSUS* STRAINS

The genome of *P. furiosus* has been reported to be dynamic (Bridger et al., 2012) — a feature of this hyperthermophile we experienced also in our Regensburg labs. Over the years, we have identified at least 2 strains differing from the original *P. furiosus* isolate whose origin/history is shown in **Figure 4**. The original strain named Vc1 was deposited as type strain DSM3638^T

↓		
FlaB0	MAKKGAVGIGTLIVFIAMVLVAAVAAGVLIQTSGLQQKSQATGRETTQEVASGIKVLSV	60
FlaB1	-MRKGAIIGIGTLIVFIAMVLVAAVAAGVIIGTAGYLQQKAQAAGRQTTQEVASGIKIVNV	59
FlaB2	-MKGKGAIIIGIGTLIVFIAMVLVAAVAAGVLIATSGYLQQKAMATGRQTTQEVASGIKVTGV	59
	:*****:*****.***.*: * :*****: * :*****:*****: *	
FlaB0	VGKTD---SNKTYVEKLAIYIISPNAGSEGIDLNNTRVVLNSNGTVQAVLKYEKTAY----	112
FlaB1	FGYINATPPSNGTIVKMAIFVTPNAGSSGIDLSNVKVVLSDGKLVVYNYSGEL----	113
FlaB2	FGYINGTPPGASNISRIVIYVAPNAGSSGIDLRYVKIVLSDGKRMAYRYYDPKEDGSSD	119
	.*: : : : .*:*****.**** .:*****:*. .* .*	
FlaB0	-----HKGAVGDVFNA-----STA-WN-----LSNT	132
FlaB1	-----YTGKILDLFNLPVWNN-----TKNG	133
FlaB2	LKPEYIHYKGDIKPNIFAYGEWEPYYKNKKPQISGEYITDNINVSAVWWNLYSAYNNSSKL	179
	:* : : : * .: :	
FlaB0	NFGIIVLQDADNSV-DQNYPTLNKGDIVVITVKV-----GEGNGVFGKGIPPRTKITGKVI	187
FlaB1	TFSIAVVNDVG-SKMEDTHPTLEWGDTVALLRT-----DDVFNYKSKNGIGPSTRIIGKVI	189
FlaB2	LFGIAVVQDGDNSLSDPQHPTLSWGDALALMIWTFFDDNNISNGFLRPGTKIIGKVI	239
	. *: * * : :***. ** ..: . * *: * *: * ***	
FlaB0	PEFGAPGVIEFTTPSTYTSEVIELQ	212
FlaB1	PDAGAAGVIDFTTPPTFEYNVIELQ	214
FlaB2	PESGAAGVIDFTTPSTYTQNLMEQ	264
	*: *** ***:*** *: :***	

FIGURE 2 | Sequence alignment of the three *P. furiosus* flagellins.

Amino acid identities for the three proteins are indicated by asterisks (*); conservative amino acid exchanges are indicated by colons (:), and semi-conservative amino acid exchanges are indicated by dots (.). The arrow shows the signal peptidase processing site. Bold ladders

represent the FlaB1 sequence correction resulting from the resequencing performed in this study. Regions indicated by gray sequences identify the least conserved central part of the proteins used to raise flagellin-specific antibodies (primers used for cloning are given in Materials and Methods).

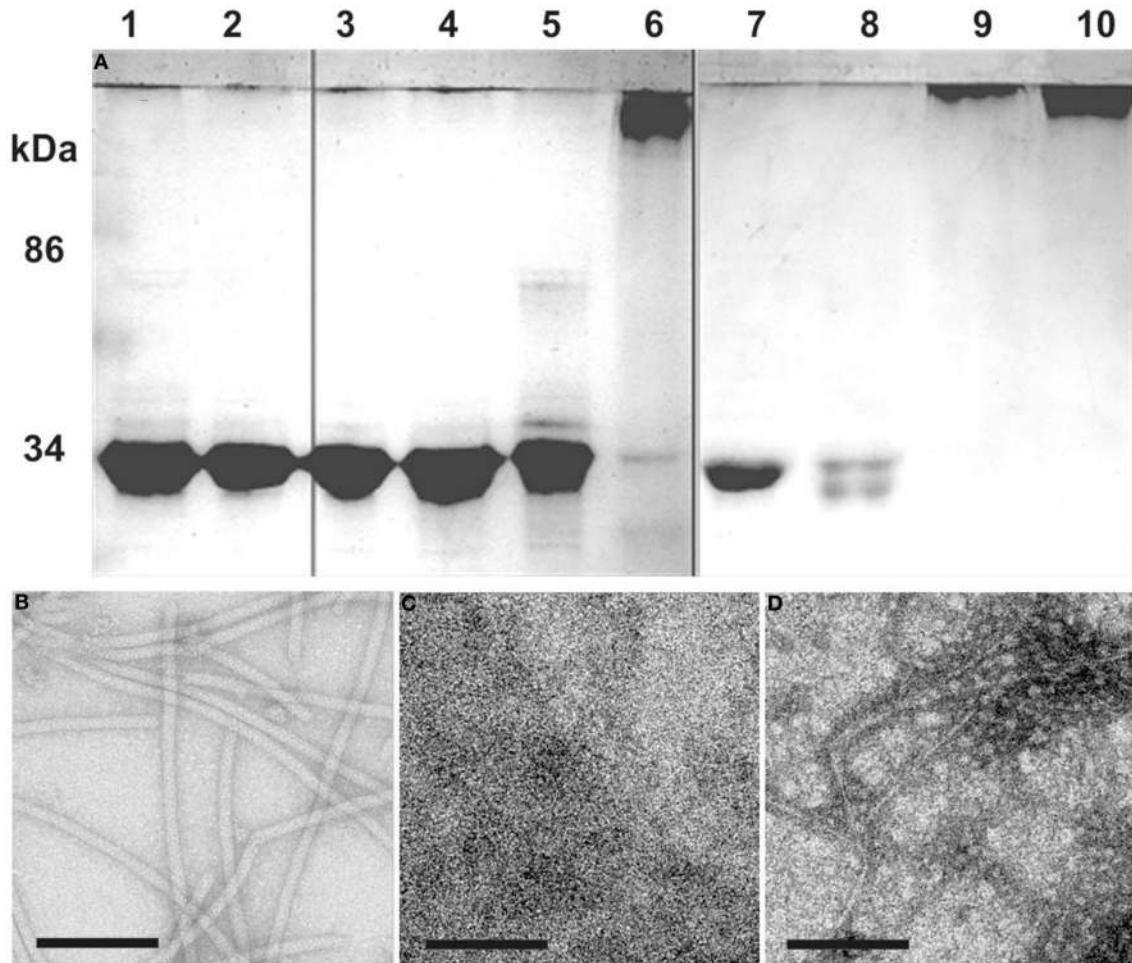


FIGURE 3 | Repolymerization of denatured flagella. (A) SDS-PAGE: Flagella purified by CsCl-gradient centrifugation were denatured into monomeric flagellins by SDS and heat denaturation (lane 1). After extensive dialysis against 5 mM HEPES buffer only single flagellins were observed (lane 2). The denatured flagellins were used for polymerization assays at 8°C

(lanes 3 and 7), 37°C (lanes 4 and 8), 60°C (lanes 5 and 9), and 90°C (lanes 6 and 10). Analysis was done after 1 (lane 3–6) or 6 days (lane 7–10) of incubation. (B–D) show TEM analyses of: (B) the flagella preparation; (C) denatured flagellins (lane 2); (D) the result from a 90°C repolymerization after 1 day (lane 6). Size bars are 100 nm, each.

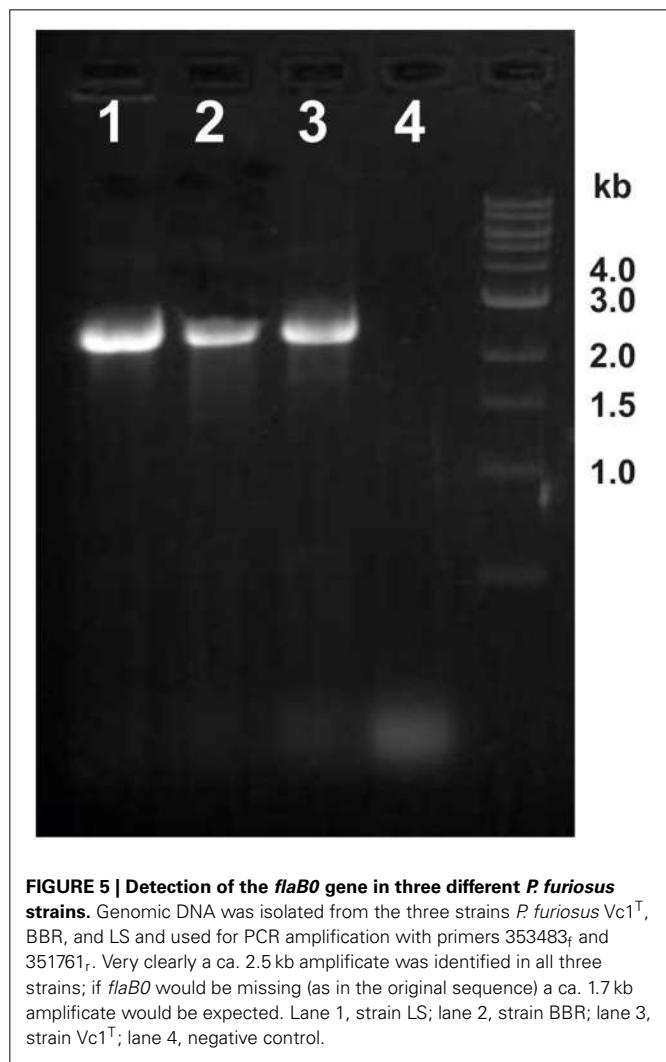
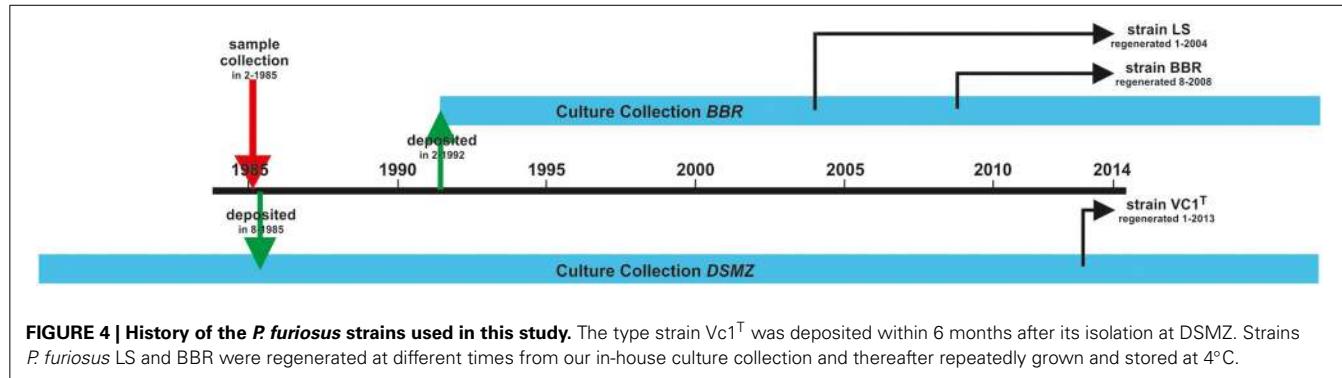
at the German Culture Collection (Deutsche Sammlung für Mikroorganismen und Zellkulturen, DSMZ) ca. 6 months after its isolation. The same isolate was repeatedly regrown (for ca. 7 years) from stocks stored at 4°C and deposited in 1992 in our Regensburg Culture Collection, (Bakterienbank Regensburg, BBR). Therefrom strain LS was regenerated in 2004 and was repeatedly regrown from stocks stored at 4°C. Another derivate, strain BBR was regenerated from our in-house culture collection in 2008 and repeatedly regrown from stocks stored at 4°C. The three strains of *P. furiosus*, namely Vc1^T, LS, and BBR, differ with respect to their binding behavior to various surfaces if tested as described (Näther et al., 2006), they express different amounts of flagella and their cell morphology differs drastically (Bellack, 2011; data will be described in detail elsewhere).

We therefore asked if the newly discovered flagellin gene *flaB0* is conserved not only in the type strain but also in the two lab derivatives. Hence, genomic DNA was isolated and primers 353483f

and 351761r were used to amplify the region around *flaB0*. For all three strains a 2.5 kb fragment was amplified as expected for the presence of *flaB0* (Figure 5). As the primer numbers refer to the binding position in the public genome of *P. furiosus* (Robb et al., 2001), the fragment should be only 1.7 kb in length when no *flaB0* would be present. Genomic sequencing of the *flaB0* region confirmed the sequence we determined earlier for the missing 771 bp segment in all three strains (data not shown).

TRANSCRIPTIONAL ANALYSES OF THE *P. FURIOSUS* FLAGELLAR OPERON

We asked if all flagella-related genes of *P. furiosus* would be transcribed together, or if various smaller transcripts might exist. Transcripts were analyzed by PCR after reverse transcription of mRNA into cDNA, the positive control used genomic DNA instead. A negative control without addition of cDNA proved that in all cases only transcripts from mRNA were analyzed (data not shown).



We detected different length cotranscripts for each of the genes of the flagellar operon with exception of *flaJ* were only the single gene transcript was found (Figure 6A). The original data using the different primers are shown exemplarily for *flaB0* in Figures 6B,C, all other data are given in Supplementary Figure S1. Several transcripts including *flaB0* were found whereof the largest with ca. 3.1 kb contained all three flagellins, *flaC*, and

flaD. Besides, various transcripts for the genes *flaF-flaI* were detected. Interestingly, we found a transcript containing *hth* and *fam* whereas *flaJ* and PF0329 were never part of a cotranscript. Analyses of RNA of cells from different growth phases showed that the transcripts changed over time; the original data are shown exemplarily in Figure 7. In early exponential phase, only few short transcripts were present compared to late exponential and stationary phase indicating that the flagellar operon is transcribed only to a limited degree in early exponential growth phase.

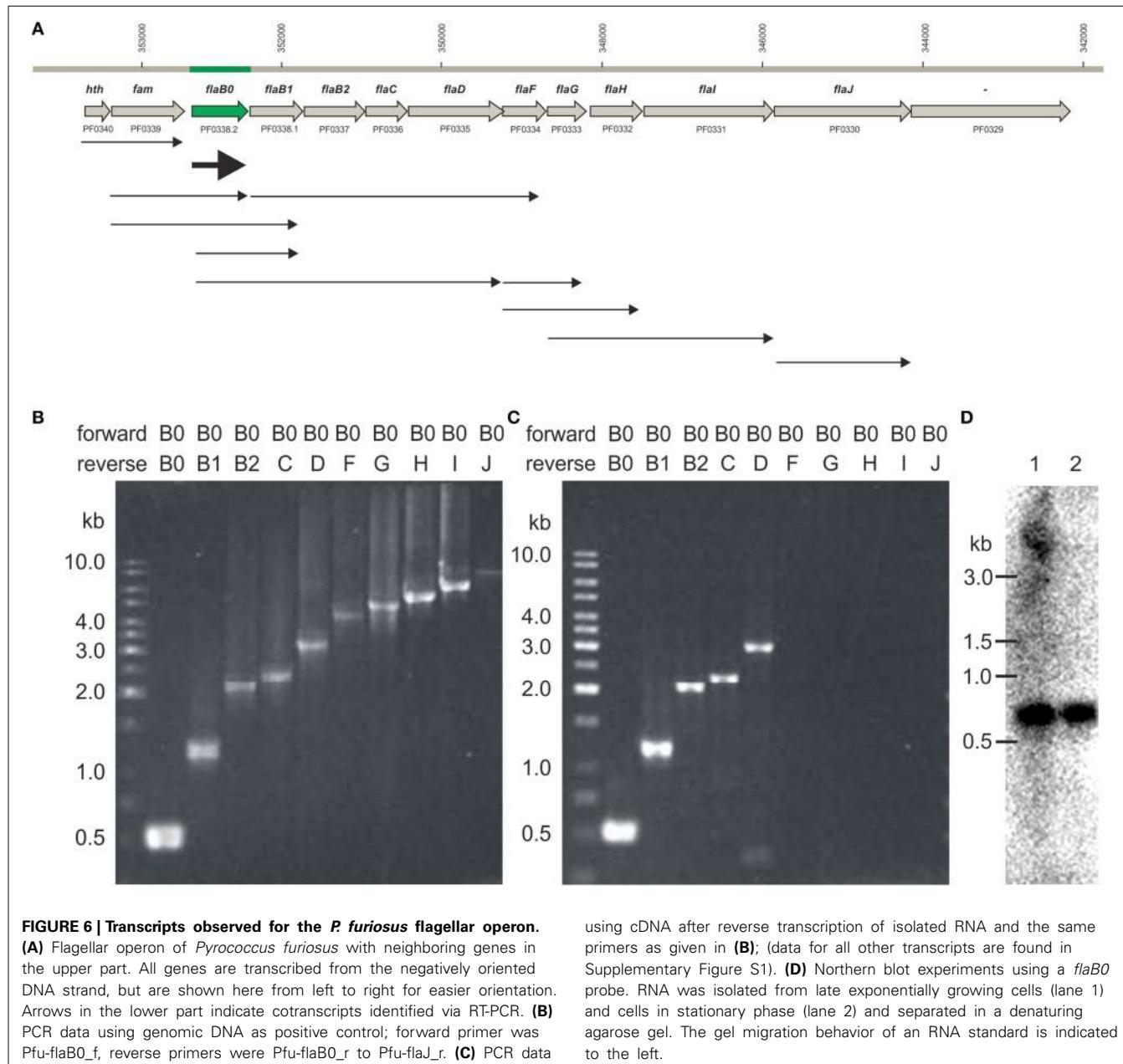
Northern blot experiments using RNA isolated from late exponentially growing cells showed the existence of a prominent ca. 600 bp long transcript using a probe for *flaB0*. In addition a much less prominent smear above ca. 4 kb was detected (Figure 6D). For cells in the stationary growth phase, only the ~600 bp long *flaB0* transcript was detected.

DISCUSSION

DIFFICULTIES TO CLONE *flaB0* MIGHT HAVE PREVENTED ITS IDENTIFICATION IN GENOME SEQUENCING

The genome of *P. furiosus* was one of the first archaeal genomes to be sequenced (Robb et al., 2001). In those “old days” of genome sequencing the shotgun cloning and sequencing approach — first used by the Venter lab to determine the *Haemophilus influenzae* genome (Fleischmann et al., 1995) — was the only way to obtain reliable data. A general problem with this approach is the fact that some genes are difficult to clone or might be even toxic for the host, normally *Escherichia coli*. In our studies, we found that *flaB0* could not be cloned into *E. coli* using standard approaches. Cloning in a vector system with expression under the strong T7 polymerase promoter as used e.g., for cloning of flagellins of *Methanococcus voltae* (Bayley and Jarrell, 1999) or in the IMPACT system (*intein-mediated purification with an affinity binding tag*) failed since the protein turned out to be toxic. We furthermore experienced problems with subcloning parts of *P. furiosus* flagellin genes, which was especially true for *flaB0*. Only the middle part of *flaB0* could be cloned, but not the N- and C-terminal regions. Hence, we suggest that these problems in cloning might also have happened during the original genome sequencing.

The only way we could obtain the *flaB0* sequence was to sequence directly from genomic DNA via primer walking. Since reading lengths of >600 bp were very difficult to obtain in the early days of genomic sequencing it is not too surprising in



retrospect that *flaB0* was not detected in the original published *P. furiosus* genome sequence (Robb et al., 2001).

FlaB0 IS THE MAJOR FLAGELLIN OF *P. FURIOSUS* FLAGELLA

The here newly described gene *flaB0* codes for the major flagellin of *P. furiosus*, whereas FlaB1 and FlaB2 are only minor flagellins. This statement is supported by the fact that one major glycoprotein of ca. 30 kDa was identified in all flagellar preparations isolated over a period of nearly 10 years; the N-terminal sequence of this protein read as AVGIGTLIVF which can be matched clearly to FlaB0. There was no heterogeneity at position 2 and the presence of FlaB1 and FlaB2 was proven only by use of specific antibodies. In further experiments we were able to show that antisera raised against the least conserved part of the flagellins FlaB1

and FlaB2 labeled CsCl-gradient purified flagella mostly on their ends, whilst an antiserum raised against purified flagella reacted much stronger over the whole length of flagella.

REPOLYMERIZATION OF DENATURED FLAGELLA

Flagellins derived from SDS- plus heat-denatured flagella could clearly be repolymerized into smaller aggregates and fibrillar structures via simple heat treatment. The ultrastructure and diameter of such fibrils differs obviously from that of purified flagella. This, however, is not too surprising if one takes into account that for flagella assembly most likely a platform containing (at least) the proteins FlaC, FlaD, FlaF, and FlaG is necessary *in vivo* (see Jarrell et al., 2013 and references therein). In addition this process is supposed to require ATP; in our hands repeated

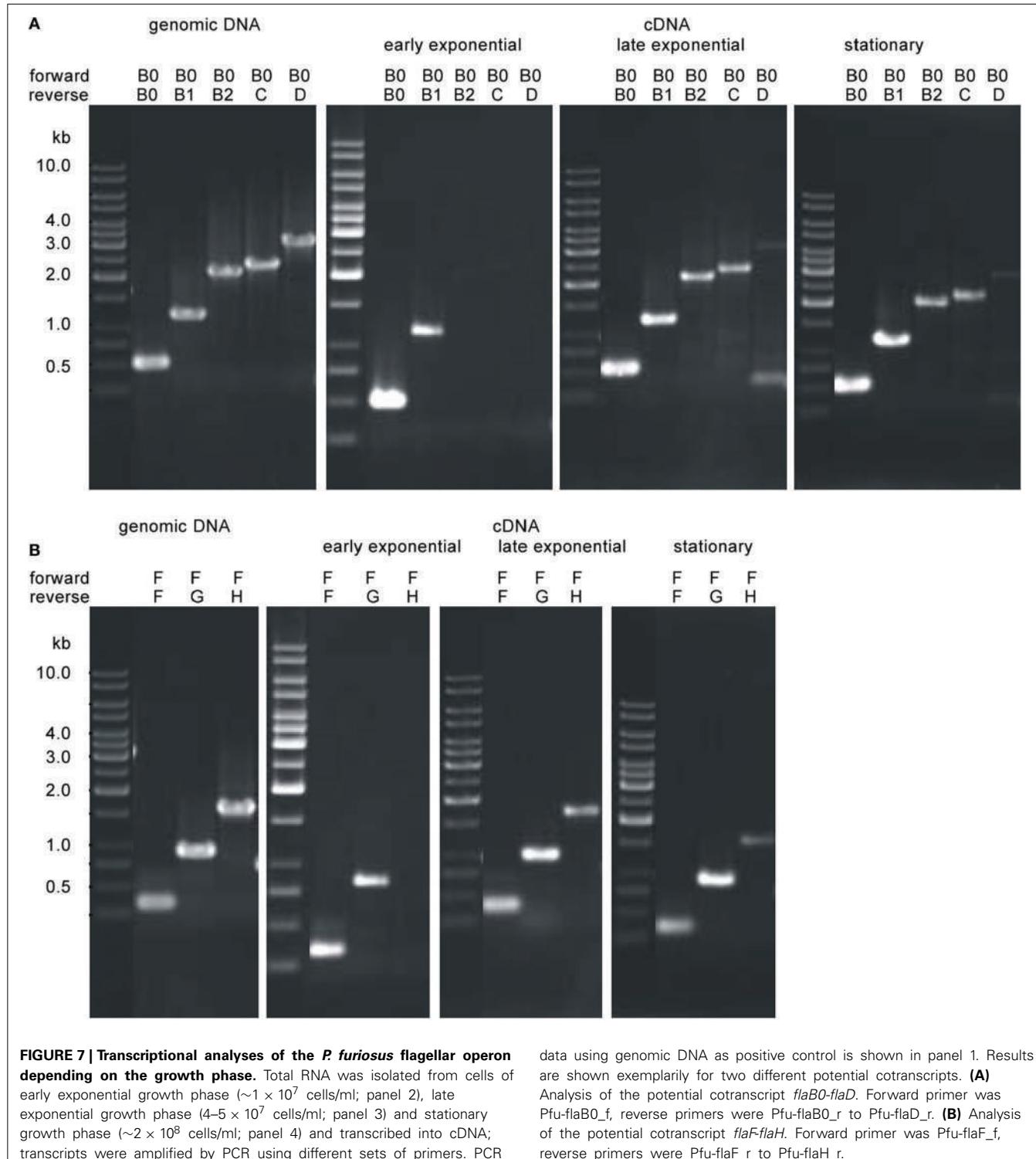


FIGURE 7 | Transcriptional analyses of the *P. furiosus* flagellar operon depending on the growth phase. Total RNA was isolated from cells of early exponential growth phase ($\sim 1 \times 10^7$ cells/ml; panel 2), late exponential growth phase ($4\text{--}5 \times 10^7$ cells/ml; panel 3) and stationary growth phase ($\sim 2 \times 10^8$ cells/ml; panel 4) and transcribed into cDNA; transcripts were amplified by PCR using different sets of primers. PCR

data using genomic DNA as positive control is shown in panel 1. Results are shown exemplarily for two different potential cotranscripts. **(A)** Analysis of the potential cotranscript *flaB0-flaD*. Forward primer was Pfu-flaB0_f, reverse primers were Pfu-flaB0_r to Pfu-flaD_r. **(B)** Analysis of the potential cotranscript *flaF-flaH*. Forward primer was Pfu-flaF_f, reverse primers were Pfu-flaF_r to Pfu-flaH_r.

ATP addition to the *in vitro* repolymerization assays, however had no effect.

THE FLAGELLAR OPERON OF VARIOUS *P. FURIOSUS* STRAINS

Transcription of genes *pf0329* to *pf0340* is from the negatively oriented DNA strand, whilst the neighboring genes are transcribed

form the positively oriented strand. Operon prediction using the Prokaryotic Operon DataBase (ProOpDB; Taboada et al., 2012) reveals two operons in this region encoding PF0340-PF0339 and PF0338-PF0329, respectively. Therefore, we analyzed this part of the genome for transcription including the flagellar operon neighboring genes *pf0329* and *pf0340*. Both, our RT-PCR

experiments and Northern Blot analyses show that there is not a single cotranscript detectable for the genes inside the flagellar operon of *P. furiosus*, starting at *flaB0* (or even *pf0340*) and ending with *flaJ* (or *pf0329*). Rather, various cotranscripts were detected. Single transcripts were observed for *flaB0* and *flaJ*. From these results we conclude that the flagellar operon is composed of genes *pf0340-pf0330*, whereas the gene encoding the hypothetical protein PF0329 is not a part of the operon.

Possible explanations for the existence of various cotranscripts we identified are as follows. *fam-flaB0*, *fam-flaB1*: the specificity of the postulated methyltransferase is unknown; it potentially could modify FlaB0 and FlaB1 (and also FlaB2). Analyzing the publically available genomes of the genus *Pyrococcus* for the presence of an ortholog of the *P. furiosus* methyltransferase, we found the respective gene in all species directly upstream of the flagellin genes, supporting our hypothesis that the enzyme acts on flagellins. *flaB0-flaD*, *flaB1-flaF*: FlaC, FlaD (and FlaE, which is not present in *P. furiosus*) are argued to be necessary for flagella assembly (see e.g., Schlesner et al., 2009); therefore, a coexpression with the flagellins would be expected. *flaF-flaG*; *flaF-flaH*; *flaG-flaI*: FlaF and FlaG have been argued to be essential for expression of flagella (Jarrell et al., 2013). Since no direct data for the function of those proteins are available, any argument about cotranscription or direct interaction of encoded proteins would be pure speculation. FlaH, FlaI, and FlaJ are probably part of the secretion and motor complex of archaeal flagella (Jarrell et al., 2013). Structural and genetic studies of the ATPase FlaI of *S. acidocaldarius* revealed that the protein forms a hexameric crown-like ring; its conformational changes and interactions with membrane lipids and binding partners (mostly FlaJ) regulate assembly and rotation of flagella (Reindl et al., 2013). Hence we expected cotranscription of the corresponding genes. However, our results showed that *flaJ* is transcribed only as a single gene regardless of the growth phase. In contrast, the cotranscripts described herein differed depending on the growth phase; these findings are in line with our electron microscopic studies of *P. furiosus* cells showing that flagella are assembled particularly in late logarithmic phase (early exponential cells possess no or only few flagella; data not shown).

The absence of one single flagellar transcript is supported by data for other Archaea: a very good overview was given by Thomas et al. (2001) (see especially Figure 4, therein). In all cases, one or more major transcripts encoding the flagellins — FlaA and/or FlaB proteins — have been identified by Northern blots. Minor transcripts, argued to code for additional proteins FlaC to FlaJ have been found in all of these cases; transcripts not starting with *flaA* or *flaB*, however have not been analyzed. Also for *Sulfolobus solfataricus* a major transcript, encoding the flagellin FlaB has been found (Szabo et al., 2007), but again transcripts encoding further genes in the flagellar operon have not been characterized.

Because the three *P. furiosus* strains presented in this study show differences in the number of flagella and adhesion properties the question arose if their flagellar operons might be different and, most notably, if the *flaB0* gene is conserved. The genome of *P. furiosus* is known to be highly dynamic as proven for the genetically tractable strain COM1 (Bridger et al., 2012) and different

Pyrococcus strains originated from environmental samples (e.g., Escobar-Paramo et al., 2005; White et al., 2008). COM1 is derived from strain DSM3638^T by targeted gene disruption of the *pyrF* locus (Lipscomb et al., 2011), it possesses 45 full or partial insertion sequences compared to 35 in strain DSM3638^T, resulting in inactivation of 13 genes. In addition alterations in 102 of 2134 predicted genes were observed, together with major chromosomal rearrangements, deletions etc. (Bridger et al., 2012). Despite these proven changes in the genome, we found that *flaB0* is well-conserved in all three *P. furiosus* strains described in this study supporting our data that *flaB0* encodes the major *P. furiosus* flagellin. Differences in flagellation and adherence, therefore, might be caused by alterations in other regions of the genome and/or regulatory effects. In this context, we note to have shown earlier that flagella contribute to adhesion (Näther et al., 2006), but we also are aware of the fact that other archaeal cell surface appendages like pili (Jarrell et al., 2011), fibers (Müller et al., 2009), fimbriae (Thoma et al., 2008), and ham (Moissl et al., 2005) at least can contribute to adhesion to various surfaces.

AUTHOR CONTRIBUTIONS

Daniela J. Näther-Schindler, Simone Schopf, Annett Bellack, Reinhard Rachel, and Reinhard Wirth designed the study and analyzed the data. Research was performed by Daniela J. Näther-Schindler, Simone Schopf, and Annett Bellack. Annett Bellack and Reinhard Wirth wrote the paper; all authors agreed to the final version.

ACKNOWLEDGMENTS

The expert technical help by E. Papst, Y. Bilek, T. Hader, and K. Eichinger is gratefully acknowledged. We thank Nadin Wimmer for cloning of FlaB2. This study was supported in its very early stages by a grant from the Deutsche Forschungsgemeinschaft (DFG WI731/10-1) to Reinhard Rachel and Reinhard Wirth.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <http://www.frontiersin.org/journal/10.3389/fmicb.2014.00695/abstract>

REFERENCES

- Bayley, D. P., and Jarrell, K. F. (1999). Overexpression of *Methanococcus voltae* flagellin subunits in *Escherichia coli* and *Pseudomonas aeruginosa*: a source of archaeal preflagellin. *J. Bacteriol.* 181, 4146–4153.
- Bellack, A. (2011). *Cell Architecture and Flagella of Hyperthermophilic Archaea*. Ph.D. thesis, University of Regensburg.
- Bellack, A., Huber, H., Rachel, R., Wanner, G., and Wirth, R. (2011). *Methanocaldococcus villosum* sp. nov., a heavily flagellated archaeon adhering to surfaces and forming cell-cell contacts. *Int. J. Syst. Evol. Microbiol.* 61, 1239–1245. doi: 10.1099/ijss.0.023663-0
- Bridger, S. L., Lancaster, W. A., Poole, F. L., Schut, G. J., and Adams, M. W. W. (2012). Genome sequencing of a genetically tractable *Pyrococcus furiosus* strain reveals a highly dynamic genome. *J. Bacteriol.* 194, 4097–4106. doi: 10.1128/JB.00439-12
- Chaban, B., Ng, S. Y. M., Kanbe, M., Saltzman, I., Nimmo, G., Aizawa, S.-I., et al. (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*. *Mol. Microbiol.* 66, 596–609. doi: 10.1111/j.1365-2958.2007.05913.x
- Escobar-Paramo, P., Ghosh, S., and DiRuggiero, J. (2005). Evidence for genetic drift in the diversification of a geographically isolated population of the

- hyperthermophilic archaeon *Pyrococcus*. *Mol. Biol. Evol.* 22, 2297–2304. doi: 10.1093/molbev/msi227
- Fleischmann, R. D., Adams, M. D., White, O., Clayton, R. A., Kirkness, E. F., Kerlavage, A. R., et al. (1995). Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd. *Science* 269, 496–512. doi: 10.1126/science.7542800
- Ghosh, A., and Albers, S.-V. (2011). Assembly and function of the archaeal flagellum. *Biochem. Soc. Trans.* 39, 64–69. doi: 10.1042/BST0390064
- Herzog, B., and Wirth, R. (2012). Swimming behavior of selected species of Archaea. *Appl. Environ. Microbiol.* 78, 1670–1674. doi: 10.1128/AEM.06723-11
- Jarrell, K. F., and Albers, S.-V. (2012). The archaellum: an old motility structure with a new name. *Trends Microbiol.* 20, 307–312. doi: 10.1016/j.tim.2012.04.007
- Jarrell, K. F., Ding, Y., Nair, D. B., and Siu, S. (2013). Surface appendages of Archaea: structure, function, genetics and assembly. *Life* 3, 86–117. doi: 10.3390/life3010086
- Jarrell, K. F., Stark, M., Nair, D. B., and Chong, J. P. J. (2011). Flagella and pili are both necessary for efficient attachment of *Methanococcus maripaludis* to surfaces. *FEMS Microbiol. Lett.* 319, 44–50. doi: 10.1111/j.1574-6968.2011.02264.x
- Kalmokoff, M. L., Jarrell, K. F., and Koval, S. F. (1988). Isolation of flagella from the Archaeabacterium *Methanococcus voltae* by phase separation with Triton X-114. *J. Bacteriol.* 170, 1752–1758.
- Lassak, K., Neiner, T., Ghosh, A., Klingl, A., Wirth, R., and Albers, S.-V. (2012). Molecular analysis of the crenarchaeal flagellum. *Mol. Microbiol.* 83, 110–124. doi: 10.1111/j.1365-2958.2011.07916.x
- Lipscomb, G. L., Stirret, K., Schut, G. J., Yang, F., Jenney, F. E. Jr., Scott, R. A., et al. (2011). Natural competence in the hyperthermophilic archaeon *Pyrococcus furiosus* facilitates genetic manipulation: construction of markerless deletions of genes encoding the two cytoplasmic hydrogenases. *Appl. Environ. Microbiol.* 77, 2232–2238. doi: 10.1128/AEM.02624-10
- Moissl, C., Rachel, R., Briegel, A., Engelhardt, H., and Huber, R. (2005). The unique structure of archaeal hami, highly complex cell appendages with nano-grappling hooks. *Mol. Microbiol.* 56, 361–370. doi: 10.1111/j.1365-2958.2005.04294.x
- Müller, D. W., Meyer, C., Gürster, S., Küper, U., Huber, H., Rachel, R., et al. (2009). The Ih0670 fibers of *Ignicoccus hospitalis*: a new type of archaeal cell surface appendage. *J. Bacteriol.* 191, 6465–6468. doi: 10.1128/JB.00858-09
- Näther, D. J., Rachel, R., Wanner, G., and Wirth, R. (2006). Flagella of *Pyrococcus furiosus*: multifunctional organelles, made for swimming, adhesion to various surfaces, and cell-cell contacts. *J. Bacteriol.* 188, 6915–6923. doi: 10.1128/JB.00527-06
- Ng, S. Y. M., Zolghadr, B., Driessen, A. J. M., Albers, S.-V., and Jarrell, K. J. (2008). Cell surface structures of Archaea. *J. Bacteriol.* 190, 6039–6047. doi: 10.1128/JB.00546-08
- Patenge, N., Berendes, A., Engelhardt, H., Schuster, S. C., and Oesterhelt, D. (2001). The *fla* gene cluster is involved in the biogenesis of flagella in *Halobacterium salinarum*. *Mol. Microbiol.* 41, 653–663. doi: 10.1046/j.1365-2958.2001.02542.x
- Rachel, R., Meyer, C., Klingl, A., Gürster, S., Heimerl, T., Wasserburger, N., et al. (2010). “Analysis of the ultrastructure of Archaea by electron microscopy,” in *Methods in Cell Biology*, Vol. 96, ed T. Müller-Reichert (Philadelphia, PA: Elsevier), 47–69.
- Reindl, S., Ghosh, A., Williams, G. J., Lassak, K., Neiner, T., Henche, A.-L., et al. (2013). Insights into FlaI functions in archaeal motor assembly and motility from structures, conformations, and genetics. *Mol. Cell* 49, 1069–1082. doi: 10.1016/j.molcel.2013.01.014
- Robb, F. T., Maeder, D. L., Brown, J. R., DiRuggiero, J., Stump, M. D., Yeh, R. K., et al. (2001). Genomic sequence of the hyperthermophile *Pyrococcus furiosus*: implications for physiology and enzymology. *Methods Enzymol.* 330, 134–157. doi: 10.1016/S0076-6879(01)30372-5
- Schlesner, M., Miller, A., Streif, S., Staudinger, W., Müller, J., Scheffer, B., et al. (2009). Identification of Archaea-specific chemotaxis proteins which interact with the flagellar apparatus. *BMC Microbiol.* 9:56. doi: 10.1186/1471-2180-9-56
- Szabo, Z., Sani, M., Groeneveld, M., Zolghadr, B., Schelert, J., Albers, S.-V., et al. (2007). Flagellar motility and structure in the hyperthermophilic archaeon *Sulfolobus solfataricus*. *J. Bacteriol.* 189, 4305–4309. doi: 10.1128/JB.00042-07
- Taboada, B., Ciria, R., Martinez-Guerrero, C. E., and Merino, E. (2012). ProOpDB: prokaryotic operon database. *Nucleic Acids Res.* 40, D627–D631. doi: 10.1093/nar/gkr1020
- Thoma, C., Frank, M., Rachel, R., Schmid, S., Näther, D., Wanner, G., et al. (2008). The Mth60-fimbriae of *Methanothermobacter thermoautotrophicus* are functional adhesins. *Environ. Microbiol.* 10, 2785–2795. doi: 10.1111/j.1462-2920.2008.01698.x
- Thomas, N. A., Bardy, S. L., and Jarrell, K. F. (2001). The archaeal flagellum: a different kind of prokaryotic motility structure. *FEMS Microbiol. Rev.* 25, 147–174. doi: 10.1111/j.1574-6976.2001.tb00575.x
- Waege, I., Schmid, G., Thumann, S., Thomm, M., and Hausner, W. (2010). Shuttle vector-based transformation system for *Pyrococcus furiosus*. *Appl. Environ. Microbiol.* 76, 3308–3313. doi: 10.1128/AEM.01951-09
- White, J. R., Escobar-Paramo, P., Mongodin, E. F., Nelson, K. E., and DiRuggiero, J. (2008). Extensive genome rearrangements and multiple horizontal gene transfer in a population of *Pyrococcus* isolates from Vulcano Island, Italy. *Appl. Environ. Microbiol.* 74, 6447–6451. doi: 10.1128/AEM.01024-08
- Wirth, R. (2012). Response to Jarrell and Albers: seven letters less does not say more. *Trends Microbiol.* 20, 511–512. doi: 10.1016/j.tim.2012.07.007

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

Received: 30 September 2014; accepted: 24 November 2014; published online: 11 December 2014.

*Citation: Näther-Schindler DJ, Schopf S, Bellack A, Rachel R and Wirth R (2014) *Pyrococcus furiosus* flagella: biochemical and transcriptional analyses identify the newly detected *flabO* gene to encode the major flagellin. Front. Microbiol. 5:695. doi: 10.3389/fmicb.2014.00695*

This article was submitted to Microbial Physiology and Metabolism, a section of the journal Frontiers in Microbiology.

Copyright © 2014 Näther-Schindler, Schopf, Bellack, Rachel and Wirth. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

OPEN ACCESS

Edited by:

Sonja-Verena Albers,
University of Freiburg, Germany

Reviewed by:

Sonja-Verena Albers,
University of Freiburg, Germany

Luis Raul Comolli,

ALS-Molecular Biology Consortium
and Lawrence Berkeley National
Laboratory, USA

Ariane Briegel,

California Institute of Technology, USA

***Correspondence:**

Christine Moissl-Eichinger,
Department of Internal Medicine,
Medical University Graz,
Auenbruggerplatz 15, 8036 Graz,
Austria

christine.moissl-eichinger@
meduni.graz.at

Specialty section:

This article was submitted to
Microbial Physiology and Metabolism,
a section of the journal
Frontiers in Microbiology

Received: 07 January 2015

Accepted: 17 May 2015

Published: 09 June 2015

Citation:

Perras AK, Daum B, Ziegler C,
Takahashi LK, Ahmed M, Wanner G,
Klingl A, Leitinger G, Kolb-Lenz D,
Gribaldo S, Auerbach A, Mora M,
Probst AJ, Bellack A and
Moissl-Eichinger C (2015) S-layers at
second glance? Altarchaeal grappling
hooks (hami) resemble archaeal
S-layer proteins in structure and
sequence. *Front. Microbiol.* 6:543.
doi: 10.3389/fmicb.2015.00543

S-layers at second glance? Altarchaeal grappling hooks (hami) resemble archaeal S-layer proteins in structure and sequence

Alexandra K. Perras^{1,2}, Bertram Daum³, Christine Ziegler⁴, Lynelle K. Takahashi⁵,
Musahid Ahmed⁵, Gerhard Wanner⁶, Andreas Klingl⁶, Gerd Leitinger⁷,
Dagmar Kolb-Lenz^{8,9}, Simonetta Gribaldo¹⁰, Anna Auerbach², Maximilian Mora¹,
Alexander J. Probst¹¹, Annett Bellack² and Christine Moissl-Eichinger^{1,2,12*}

¹ Department of Internal Medicine, Medical University of Graz, Graz, Austria, ² Department of Microbiology and Archaea Center, University of Regensburg, Regensburg, Germany, ³ Department of Structural Biology, Max Planck Institute of Biophysics, Frankfurt, Germany, ⁴ Department of Biophysics, University of Regensburg, Regensburg, Germany, ⁵ Chemical Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, CA, USA, ⁶ Faculty of Biology, Ludwig-Maximilians-University of Munich, Munich, Germany, ⁷ Research Unit Electron Microscopic Techniques, Institute of Cell Biology, Histology and Embryology, Medical University of Graz, Graz, Austria, ⁸ Institute of Cell Biology, Histology and Embryology, Medical University of Graz, Graz, Austria, ⁹ Core Facility Ultrastructure, Analysis, Center for Medical Research Institute, Medical University of Graz, Graz, Austria, ¹⁰ Unité Biologie Moléculaire du Gène chez les Extrémophiles, Département de Microbiologie, Institut Pasteur, Paris, France, ¹¹ Department of Earth and Planetary Science, University of California, Berkeley, Berkeley, CA, USA, ¹² BioTechMed-Graz, Graz, Austria

The uncultivated “*Candidatus Altarchaeum hamiconexum*” (formerly known as SM1 Euryarchaeon) carries highly specialized nano-grappling hooks (“hami”) on its cell surface. Until now little is known about the major protein forming these structured fibrous cell surface appendages, the genes involved or membrane anchoring of these filaments. These aspects were analyzed in depth in this study using environmental transcriptomics combined with imaging methods. Since a laboratory culture of this archaeon is not yet available, natural biofilm samples with high *C. A. hamiconexum* abundance were used for the entire analyses. The filamentous surface appendages spanned both membranes of the cell, which are composed of glycosyl-archaeol. The hami consisted of multiple copies of the same protein, the corresponding gene of which was identified via metagenome-mapped transcriptome analysis. The hamus subunit proteins, which are likely to self-assemble due to their predicted beta sheet topology, revealed no similarity to known microbial flagella-, archaella-, fimbriae- or pili-proteins, but a high similarity to known S-layer proteins of the archaeal domain at their N-terminal region (44–47% identity). Our results provide new insights into the structure of the unique hami and their major protein and indicate their divergent evolution with S-layer proteins.

Keywords: archaea, S-layers, archaeal cell surface appendages, hami, nano-grappling hooks, double-membrane, environmental transcriptomics, electron cryo-tomography

Introduction

In the course of evolution, nature has developed simple and fascinating solutions for various challenges. Particularly microbial life seems thus to harbor an enormous potential of exploitable biomaterial, such as enzymes and other biomolecules. These compounds are thought to prove very useful for diverse applications, in e.g., medicine, pharmacy, or industry (e.g., Beg et al., 2001; Hasan et al., 2006; Dutta and Kundu, 2014). However, the majority of naturally occurring exploitable biomaterial remains to be explored, because a substantial amount of microorganisms resist cultivation in the laboratory and thus escape detailed characterization of their metabolic potential and functional traits.

Cultivation-independent methods such as metagenomics enable scientists to directly access the genetic information of (entire) microbial communities. The sequence information retrieved can be used for assembly of near complete to complete genomes from key or underrepresented members of the communities (Tyson et al., 2004; Sharon and Banfield, 2013; Sharon et al., 2013). This information thus provides the basis for functional annotation of these novel microbial genomes. However, annotation of genes from lineages with only distant representatives is sometimes limited. Some cases have been reported in which approximately 50% of the predicted proteins could not be assigned a function (Baker et al., 2010; Kantor et al., 2013). Consequently, linking metagenomic data from uncultivated microorganisms with information retrieved by other molecular methods and/or imaging techniques in order to characterize such unknown predicted proteins is a promising approach. Imaging techniques, however, can currently not be conducted for highly complex microbial communities (e.g., those from soil) without substantial loss of information. Nevertheless, populations with low and simple diversity and uneven abundance of its members, such as the uncultivated acid mine drainage microbial community, can be studied in detail using a variety of these techniques, enabling researchers to link metagenomics to cellular characteristics (Comolli et al., 2009; Baker et al., 2010; Yelton et al., 2013; Comolli and Banfield, 2014).

For instance, genomes of ARMAN cells have been linked to their ultrastructure (Baker et al., 2006, 2010; Comolli et al., 2009); the latter also revealed that their cell wall is composed of a double membrane—a highly unusual feature in the domain of Archaea (Klingl, 2014). This feature was described to be shared only with a few other members in this domain, which are represented by the genus *Ignicoccus* (Rachel et al., 2002; Junglas et al., 2008), the *Methanomassiliicoccus* species (Dridi et al., 2012; Iino et al., 2013; Borrel et al., 2014) and the recently described “*Candidatus Altarchaeum hamiconexum*” (formerly known as SM1 Euryarchaeon, Rudolph et al., 2001; Probst et al., 2014).

Ca. A. hamiconexum is a representative of the recently introduced euryarchaeal order “*Candidatus Altarchaeales*,” a widespread group of uncultivated archaea thriving in subsurface aquifers (Probst et al., 2014). The tiny coccoid *Ca. A. hamiconexum* cells are washed up from the subsurface in extraordinarily pure biofilms (Henneberger et al., 2006; Probst et al., 2013). In a very recent publication, metagenomic

sequencing data of *Ca. A. hamiconexum* biofilms were combined with isotopic-based lipidomics to reveal its autotrophic metabolism, which may be the basis for substantial carbon dioxide fixation in the subsurface. Lipidomics has further shown that the archaeon’s double membrane is composed of core diether lipids with either two phytanyl chains or a combination of one phytanyl and one sesterpanyl chain (Probst et al., 2014). Anchored in this membrane, unique cell surface appendages called “hami” (singular: “hamus”; Moissl et al., 2005) were identified via ultrastructural analyses. Hundreds of these hami protrude from each cell and interlink with those of neighboring cells in order to form a biofilm (Probst et al., 2014). Each filament is assembled from three protein sub-filaments that are wound up to a barbed-wire-like structure and a distal grappling hook. This unique structure is supposed to be composed of one major protein species (Moissl et al., 2005; Probst et al., 2014)—similar to surface layer proteins (S-layer), which usually also consist of one or two protein species that assemble into a 2-dimensional array on cell surfaces (Sleytr et al., 1988; Eichler, 2003; Veith et al., 2009; Klingl et al., 2011; for a detailed review on archaeal S-layers see Klingl, 2014, same Research Topic). It was proposed that individual hami subunits are expressed in the cytoplasm, transferred through the inner membrane by the Sec-pathway and then assembled in the periplasmic space between inner and outer membrane (Probst et al., 2014). Although six genes were annotated as putative hamus subunit-encoding genes, the actual gene that is expressed for hamus formation *in vivo* has not been identified so far (Probst et al., 2014).

In this study a combination of -omic techniques with electron microscopy was applied in order to identify the *bona fide* gene sequence of the hamus subunit, shed light onto its phylogenetic evolution and further analyze its structure and the membrane in which it is anchored. Due to its barbed-wire-like structure and in particular its distal nano-hook, the hamus is considered an exploitable biomaterial and thus a tool for nanobiotechnology (Moissl-Eichinger et al., 2012), for which we provide the basis in this communication.

Materials and Methods

Sampling and Sampling Processing

Archaeal biofilm samples were taken from the cold sulfidic spring, called Mühlbacher Schwefelquelle Isling (MSI), which is located in close vicinity to Regensburg, Germany (Rudolph et al., 2004). The biofilms were harvested from double-opened flasks, which were incubated for 2–3 days in the water-flow of the spring, approx. 30 cm below the surface. The flasks were equipped with a polyethylene net, which proved useful to catch upwelling biofilm-pieces from the spring-water. After the incubation period, the entire flask was closed under water using rubber stoppers and transported on ice to the laboratory, where the samples were immediately processed (Probst et al., 2013).

TOF-SIMS (Time of Flight Secondary Ion Mass Spectrometry) Chemical Imaging

Archaeal biofilms were washed from the nets, and free-floating biofilm pieces were collected onto gold-plated screens (hole

100 µm, G225G1, Plano GmbH, Wetzlar, Germany). Samples were immediately dried and the gold-coated aperture disks were placed onto silicon wafers and affixed along the edges with adhesive tape, with care to avoid contact with the biofilm.

Chemical imaging was performed with a modified commercial reflectron-type time-of-flight secondary ion mass spectrometer (TOF-SIMS V; IonTOF, Germany). Mass-selected Bi₃⁺ ions with 25 keV kinetic energy impacted the sample surface at 45° with respect to the surface normal. Ejected cationic and anionic chemical species were collected in separate analyses. Time-of-flight secondary ion mass spectrometry (TOF-SIMS) spectra were acquired with Bi₃⁺ pulses in high current bunched mode, over an area of 500 µm × 500 µm, with a 256 pixel × 256 pixel raster scan at a repetition rate of 2.5 kHz, and secondary ions were extracted with a 10 µs long extraction –2000 V (positive ion mode) or +2000 V (negative ion mode) pulse. Electron charge compensation was not used.

Electron Cryo-Tomography

Biofilm samples were centrifuged at 16,000 × g using an Eppendorf 5415 C table top centrifuge, and cell pellets were resuspended in an equal volume of KPH buffer (0.7 mM NaCl, 0.1 mM MgCl₂, 1.6 mM CaSO₄, 1 mM HEPES; the pH was adjusted with NaOH to 6.5; Moissl et al., 2005). Cellular suspensions were mixed with an equal amount of 10 nm colloidal protein-A gold (Aurion, Wageningen, The Netherlands), and 3 µl of this mixture were added to a glow-discharged Quantifoil (Quantifoil, Großlobichau, Germany) grid, blotted for 3–5 s and plunged into liquid ethane.

Samples were transferred frozen into a Polara G2 Tecnai transmission electron microscope (FEI, Hillsboro, USA) operated at 300 kV. The TEM was equipped with a Gatan 4 × 4 k charge-coupled device (CCD) camera or a K2-Summit direct electron detector as well as a Tridiem energy filter (Gatan Inc., Pleasanton, USA). Images were recorded using a magnification of 34,000 × on the CCD and 41,000 × on the K2 summit, corresponding to a pixel size of 0.6 or 0.54 in the final image, respectively. Zero loss filtered tomographic tilt series were collected in a range of –60° to +60°, at increments of 2°–2.5° and a defocus of 6–10 µm using the Gatan Digital Micrograph Latitude software (Gatan Inc., Pleasanton, USA). The maximum cumulative dose was 150 e[–]/Å².

For electron tomography at room temperature, samples were high-pressure frozen and embedded in epon resin as described in Perras et al. (2014). Tilt series were recorded at 120 kV on a JEOL JEM-2100, equipped with a LaB₆ cathode and a 2 × 2 k F214 fast scan CCD camera (TVIPS, Gauting, Germany) in a range of –60° to +60° in steps of 2°, a defocus of 3–6 µm and a magnification of 14,000 ×, which equals a pixel size of 1 nm in the final micrograph. Tilt series were reconstructed into tomograms with the IMOD software (Kremer et al., 1996), using weighted backprojection or SIRT and displayed in 3 dmod (IMOD).

Ultrastructural Analysis Using Transmission Electron Microscopy (TEM) And Scanning Electron Microscopy (SEM)

For analysing the cell surface appendages, the unfixed, purified hamis were deposited on a carbon-coated copper grid and

negatively stained with uranylacetate [2% (w/v), pH 4.5]. The samples were examined with a CM12 transmission electron microscope (FEI Co., Eindhoven, The Netherlands) operated at 120 kV. All images were digitally recorded using a slow-scan CCD camera that was connected to a computer with TVIPS software (TVIPS GmbH, Gauting, Germany). Scanning electron microscopy was performed as described in Probst et al. (2014).

Particle Characterization within the Cells

For element analysis of the particles within coccoid cells, *Ca. A. hamiconexum* biofilm flocks were embedded in TAAB embedding resin (TAAB, Aldermaston, UK) and thin sectioned as described in Milić et al. (2015) followed by staining with platinum blue and lead citrate. Energy Filtered TEM (EFTEM) was performed with a Gatan GIF Quantum 963 energy filter using an FEI Tecnai 20 microscope at 120 kV acceleration voltage. To visualize the elemental distributions, elemental maps were made using the three window method at the standard losses provided by Gatan Digital Micrograph software (see also Teubl et al., 2014).

Energy Dispersive X-Ray spectroscopy was performed using an Edax silicum type ultrathin unit (SUTW) detector, as described in Milić et al. (2015); the corresponding images were made with scanning transmission EM using a High Angle Annular Dark Field detector (HAADF).

Purification of Hamus Filaments and Antibody Generation

For the production of hamus-specific antibodies for protein analyses and structural investigations, hamus filaments were released from the archaeal cell surface. The purification procedure, as well as the production of hamus-specific antibodies, has been described elsewhere (Probst et al., 2014). In brief, the archaeal biofilm cells were lysed using 0.1% (w/v) sodium dodecyl sulfate (SDS) and cell debris was removed via subsequent centrifugation and sucrose-gradient centrifugation steps.

Denaturing SDS-PAGE Analysis and Western Blot

Hamus filaments were purified as described above and mixed with protein loading dye [Tris/HCl pH 7.5, 60 Mm; Glycerol 10% (v/v); SDS 2% (w/v); bromphenol blue 0.01% (w/v); 2-mercaptoethanol 5% (w/v)] and heated for 30 min in boiling water. Afterwards, proteins were separated via SDS-PAGE (Laemmli, 1970) using a Mini Protean 3 Cell [Bio-Rad Laboratories Inc., Munich; 12.5% (w/v) polyacrylamide linear gradient gel], at 15 mA followed by a higher current of 30 mA until the dye front reached the bottom of the gel.

The separated proteins were afterwards semi-dry-blotted onto a Roti®-PVDF (polyvinylidene fluoride) membrane (Carl Roth GmbH + Co. KG, Karlsruhe, Germany) using a semidry transfer cell instrument (Bio-Rad, Munich, Germany) operated at 16 V for 1 h. Blocking was performed by incubation of the membrane in Tris buffered saline [including Tween 200.01% (v/v), 3% milk powder (w/v); TBST-B] overnight. After a washing step with Tris buffered saline [including Tween 200.01% (v/v); TBST], the primary antibody (anti-hamus) was applied (1:5,000 dilution in

TBST-B) and incubated for 3 h under agitation. The membrane was washed using TBST and the secondary antibody [anti-chicken coupled with horseradish peroxidase (1:1,000 in TBST-B; Sigma-Aldrich Chemie GmbH, Munich, Germany)], was applied for 2 h]. The reaction was visualized by applying a 3-amino-9-ethylcarbazole solution [20 mg of 3-amino-9-ethylcarbazole dissolved in 1 ml ethanol p.a., followed by mixing with 50 ml of potassium acetate, pH 5, 20 mM, 100 µl of triton X-100, 10% (v/v) and 10 µl of H₂O₂] after another washing step.

Liquid Chromatography Mass Spectrometry of Peptides

For identification of peptides in the band showing positive reaction in the western blot analysis, the corresponding band in the SDS-PAGE was cut out and trypsin digested. Obtained peptides were then subjected to liquid chromatography coupled to tandem mass spectrometry (LC-MS/MS). HPLC was carried out using a Ultimate3000 RSLC nano-HPLC System (Thermo Fisher Scientific; at the facilities of Prof. Dr. R. Deutzmann, University of Regensburg) with a reversed phase chromatography analytical column (ReproSil Pur 120 C18-AQ, 75 µm × 25 cm). The mobile phase consisted of a linear gradient containing 0.1% (v/v) formic acid (eluent A) and 80% (v/v) acetonitrile, 0.1% (v/v) formic acid (eluent B). HPLC was coupled online to a maXis plus UHR-QTOF system (Bruker Daltonics) via nano-electrospray source and up to five most abundant precursors selected for fragmentation by collisional induced dissociation (CID). Identification of the obtained peptide mass fingerprints was performed by genome database searching using the PeptideMass software (Wilkins et al., 1997).

Fluorescence Immuno-Labeling

For immuno-staining, the archaeal biofilms were fixed with paraformaldehyde (5%; v/v) at room temperature (1 h) and washed three times with PBS (phosphate buffered saline). The cell suspension was applied into a well of a gelatine-coated slide [P. Marienfeld KG, Lauda-Koenigshofen, Germany; slides were dipped into a 0.01% (w/v) gelatine solution and dust-free air-dried]. The dried cells were afterwards incubated in 16 µl of PBST [+0.1% (w/v) SDS] at 37°C and the PBST buffer was replaced with 16 µl PBST buffer containing the anti-hamus IgG (Davids biotechnology, Regensburg, Germany; dilution 1:2,000; 1 h, 37°C). Subsequently, the slide was washed 15 min in PBST [+0.1% (v/v) SDS], rinsed with H₂O and air dried. The secondary antibody (goat anti-chicken, Cy3-labeled; 0.64 mg/ml, Sigma Aldrich, Germany; dilution 1:500) was added and incubated for 1 h at 37°C. After washing two times with PBST [+0.1% (v/v) SDS], the slide was rinsed with H₂O again, DAPI counterstained and analyzed by fluorescence microscopy (Olympus BX53F, Hamburg, Germany) with epifluorescence equipment and the respective imaging software (cellSens, Olympus).

Transcriptomic Analysis of Hamus Gene(S)

The presence or absence of specific hamus subunit transcripts was tested via specific mRNA detection in archaeal biofilm samples (see Supplementary Table S1, containing list of genes and primers). Three hamus subunits (i.e., MSIBFv1_A2980002, MSIBFv1_A2020020, MSIBFv1_A3210004; deposited in the

European Nucleotide Archive; accession code: PRJEB6121; Probst et al., 2014) were examined in detail.

RNA was isolated using the PowerBiofilm™ RNA Isolation Kit (Mobio Laboratories Inc., Carlsbad, USA), following the manufacturers' instructions (DNA digestion was performed for 30 min). DNase treatment was repeated after precipitation of nucleic acids. Contamination of the RNA extraction by residual DNA was excluded by PCR amplification, using a universal archaeal forward primer combined with a reverse primer binding in a non-transcribed region (e.g., primer: 344aF and 64R-23S, sequences in Supplementary Table S1). Samples showing no positive PCR signals were assumed to be free of contaminating DNA and were further processed. RNA was reverse-transcribed into cDNA (QuantiTect Rev. Transcription Kit, Qiagen, Hilden, Germany). Transcripts were amplified using specific primers, which were designed using the web tool Primer3v.0.4.0 software (http://biotools.umassmed.edu/bioapps/primer3_www.cgi; parameters: product size: optimum full length of gene, GC% 40–60%, annealing temperature: 60°C optimum). To confirm the specificity, primers were searched (Altschul et al., 1990) against NCBI NR and against the *Ca. A. hamiconexum* metagenome (Probst et al., 2014). The designed primer pairs (see Supplementary Table S1) were used individually using cDNA as a template and tested for amplification success (denaturation time: 5 min 95°C; 30 cycles: 45 s 94°C, 45 s 60°C, 130 s 72°C; final elongation: 10 min 72°C). Positive PCR products were purified (HiYield® Gel/PCR DNA Fragments Extraction Kit; Süd-Laborbedarf GmbH, Gauting, Germany) and Sanger-sequenced (LGC Genomics GmbH, Berlin, Germany). Experiments were carried out in duplicates.

Hamus Protein Analysis, Structure Determination, and Phylogenetic Tree Reconstruction

The trans-membrane region of the hamus protein was predicted by TMHMM v2.0 (<http://www.cbs.dtu.dk/services/TMHMM/>). The protein characteristics were analyzed using GenScript's Peptide Property Calculator (https://www.genscript.com/ssl-bin/site2/peptide_calculation.cgi) and by NetNGlyc 1.0 (<http://www.cbs.dtu.dk/services/NetNGlyc>).

Secondary structure was predicted using PSIPRED (<http://bioinf.cs.ucl.ac.uk/psipred/>), alignment of the sequences was performed using the multi-sequence alignment program MAFFT (<http://mafft.cbrc.jp/alignment/server/>). PSIPRED prediction was combined with a fold recognition search using pGenThreader (Lobley et al., 2009).

Hamus protein region 5–81 was searched against the NCBI database (blastp, Altschul et al., 1990). The 50 mostly related protein sequences were retrieved and used for tree (Neighbor Joining and Maximum Likelihood) reconstruction via MEGA 6 (Tamura et al., 2013).

Results

The *Ca. A. Hamiconexum* Double Membrane Is Composed of Glycosyl-Archaeol Species

In positive ion TOF-SIMS spectra, several notable mass spectral peaks were observed in the high mass range [700–1,200 atomic

mass units (amu) and around 2,000 amu]. Most prominent of these mass peaks is a cluster of peaks centered about \sim 1,000 amu, which was assigned to sodiated diglycosyl archaeol (Figure S1). 16 mass units lower and higher of the sodiated lipid peaks were additional clusters of peaks which could either reflect lithiated and potassiated variants, or, for the m/z \sim 1,016 peaks contribution from a core hydroxyarchaeol with sodiated diglycosyl polar head group. Minor contributions of sodiated triglycosyl archaeol and monoglycosyl archaeol were also detected at m/z 1,162 and 838, respectively.

In addition to the salt adducts of the mono-, di- and tri-glycosyl lipids, a significant contribution from m/z \sim 1,070 was observed. From LC/MS/MS data of additional samples of the SM1 biofilm, this lipid was assigned to a diglycosyl diether lipid with one C_{20} hydrocarbon chain and one C_{25} hydrocarbon chain.

In the mass range where dimers or tetraether species would occur, several groups of peaks could also be observed, albeit at relatively low intensity (Figure S2). Prominent among these species was a cluster of peaks with a maximum peak intensity observed at m/z 1,974.5 and 2,006.5 amu. M/z 1,974.5 was found too low in mass to correspond to a simple sodiated dimer of the diglycosyl lipid (strongest isotope peak would appear at m/z 1,977.6), and too high in mass to correspond to a sodiated diglycosyl dialkyl tetraether (which would have its strongest isotope peak at m/z 1,973.5). Based on the strong sodiated lipid contribution in the diglycosyl diether lipid-related peaks, it was assumed that the peaks represented a sodiated species. With this assumption, one possible structure could be a sodiated *trialkyl* tetraether lipid with four total glycosyl groups on the ends (Figure S2). It is not clear whether this species is native to the biofilm or a result of a dimerization process that occurs during the ion sputtering event.

Particles Localized inside and outside of the *Ca.*

A. Hamiconexum Cells Reveal Different Elemental Composition, Indicating the Presence of Sulfur and Phosphor Deposits within the Cells

TOF-SIMS analysis revealed inorganic species embedded in the biofilm, containing Na, K, Ca, Mg, and Fe (positive ion

mode TOF-SIMS; Figure S3). Ca^{2+} and Mg^{2+} appeared to be concentrated within particles in the biofilm, which may indicate the formation of insoluble mineral carbonates. As revealed by scanning electron microscopy, such particles were easily visible in the preparations (Figure 1A), and many particles were located in close vicinity or even touched by the hamis of highly actively dividing cells (Figure 1A).

Electron microscopic images of ultrathin-sectioned cells showed that most of the cells exhibited dark inclusions within their cells (Figure 1B). Such dense areas were examined using energy dispersive X-ray spectroscopy and were identified containing most likely phosphorus and sulfur (in some cells) (data not shown). This result was further confirmed using EFTEM (Figure S4).

Hamus-Filaments Are Three-Dimensional Structures and Span Both Membranes

Preliminary results of electron tomography from entire *Ca. A. hamiconexum* cells after high-pressure-freezing, freeze substitution and epon embedding resulted in a cell wall model (Figure 2). Both the inner and the outer membrane were visualized in a 50 nm section through the cell (Figure 2; Supplementary Movies 1, 2). Filamentous structures (8 nm in diameter), most likely representing the hamis, passed through both membranes. At the cytoplasmic end of some of the filaments, an electron dense structure could be detected, representing a potential anchorage-associated structure of the filaments.

Electron cryo-tomography of hamis several μ m away from the cell surface (Figure 3) revealed the typical barbed-wire structure as seen in micrographs of negatively stained samples (Figure 4A, see also: Moissl et al., 2005; Perras et al., 2014; Probst et al., 2014). In the tomograms, the hamis showed a repeating pattern of prickle triplets at an average center-to-center distance of 47 nm. Each prickle revealed an average length of 36 nm and emanated from the backbone of the filament. However, hamis missing the typical barbed-wire pattern were also observed (Figures 4A,B). They were partially plain filaments and could be identified as hamis by their dimensions and typical hook (Figure 4A).

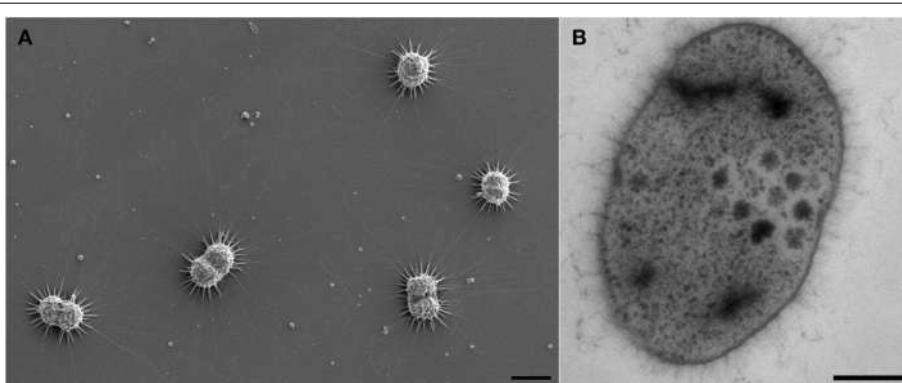


FIGURE 1 | Electron micrographs of *Ca. A. hamiconexum* cells within the biofilm. **(A)** Scanning electron micrograph of dividing cells in the biofilm, showing particles close to the hamis and the cells. Bar: 1 μ m. **(B)** Thin-section of one cell, showing dark-stained areas. Bar. 200 nm.

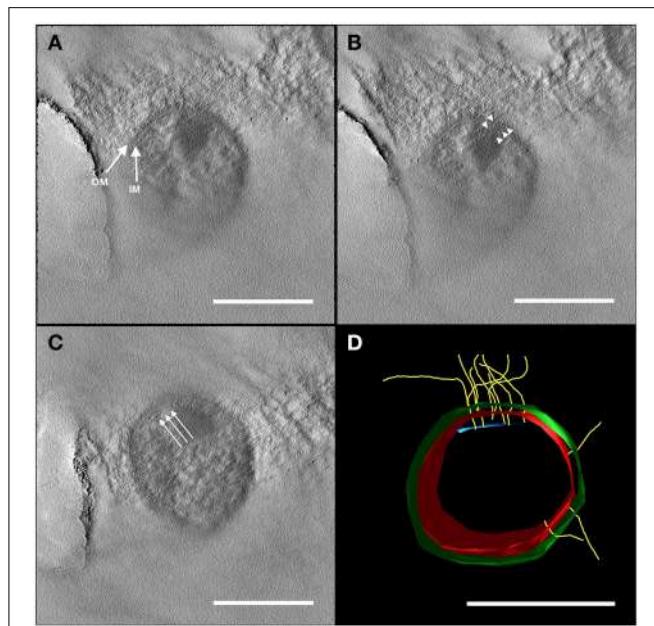


FIGURE 2 | Electron tomography of a 50 nm section of high-pressure frozen and freeze substituted cells. The tilt series was carried out at 120 kV from -60 to $+60^\circ$ with an increment of 2° and a final magnification of 14,000 \times . The reconstruction of the tomogram was performed with IMOD. In selected virtual slices, several structures were indicated. **(A)** The final tomogram revealed the inner (IM) and outer membrane (OM) with a thickness of about 6 nm in each case. The overall thickness of inner and OM together with the periplasm was 44 nm. Therefore, the periplasm has a width of about 32 nm. **(B)** Further, two hami filaments with a diameter of approximately 8 nm were indicated spanning both the IM and OM (arrow heads). **(C)** Underneath the IM, another layer could be recognized (arrows). As the resolution of the tomogram is quite low, this could either be kind of an anchoring structure/mechano sensor or a preparation/reconstruction artifact as it was just visible in a very low number of the virtual slices of the tomogram. With this tomogram, a model (segmentation) was constructed **(D)** illustrating the IM (red) and OM (green), the hami filaments (yellow) and the supposed anchoring-associated structure (blue). For simplicity, the membranes were illustrated as a monolayer. Scale bars: 500 nm.

Apart from individual hamus filaments, bundles of such filaments were found (Figure 4C), suggesting that they are capable of forming super-filaments that are interlinked by the prickle. Tomograms of the cell body showed a plethora of filaments emanating from the cells (Supplementary Movies 3, 4). Most of these filaments showed the barbed wire-like structure typical for hami. Due to the thickness of the cell body, filaments passing the membranes could not be clearly resolved.

The Major Hamus Subunit Protein Is Encoded by One of Six Homolog Genes

Purified hami were found to be composed of one major protein (“120 kDa protein”; Moissl et al., 2005). The antibodies generated against purified hamus filaments showed a strong and specific reaction against the surfaces of coccoid *Ca. A. hamiconexum* cells within the biofilms (Figure 5). No signal was obtained from (filamentous) bacteria, occasionally enclosed in the biofilm. The same antibodies were used for a Western

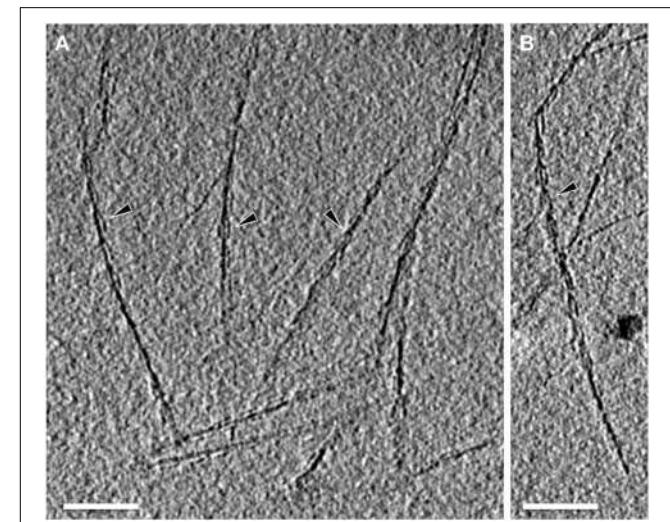


FIGURE 3 | Electron cryo-tomography of the hami. Tomographic slices through hami emanating from the cell surface. Hami occur as single barbed wire-like filaments **(A,B;** black arrowheads). Scale bar: 100 nm.

blot immuno-assay, showing a clear, strong signal appearing on the gel—the previously identified “120 kDa protein” (Figure S5; Moissl et al., 2005). The LC-MS/MS fingerprint from this protein was compared with the metagenomic information of *Ca. A. hamiconexum* retrieved during a recent study (Probst et al., 2014). Six homologous open reading frames (ORFs) potentially encoding for the major hamus protein were identified within the metagenomic data set. Supported by transcriptomic data (the retrieved RNA sequence revealed 100% identity in sequence and length), ORF MSIBFv1_A321004 was identified to code for the transcribed hamus subunit protein. This ORF was located on a relatively small contig (\sim 5,600 nucleotides), containing only five ORFs in total. Transcription of the other homologous hamus genes could not be proven.

The Hamus Protein Carries Glycosylation Sites and an S-Layer-Like N-Terminal Region

The identified hamus ORF MSIBFv1_A321004 coded for an acidic, soluble 97 kDa protein (hydrophobicity: -0.28 , PI 5.18) with major components Gly (9.71%) and Thr (9.49%). A sec-signal peptide was predicted [amino acid (aa) 1-27; Probst et al., 2014], and in the same region of the gene, in congruence, TMHMM predicted a transmembrane helix (aa: 7-29). This sequence was predicted to be hydrophobic and charged positively (pi: 10.51). At least seven potential glycosylation sites were proposed for the hamus subunit protein. Thus, the discrepancy between the protein mass estimated from SDS-PAGE (120 kDa) and the gene-predicted mass (97 kDa) resulted most likely from post-translational glycosylation of the protein (Moissl et al., 2005).

The hamus subunit gene (ORF MSIBFv1_A321004) shared the contig with four additional ORFs (Figure S6), encoding for two proteins of unknown function, a glutamate-tRNA-ligase (closest related protein from *Methanobacterium*

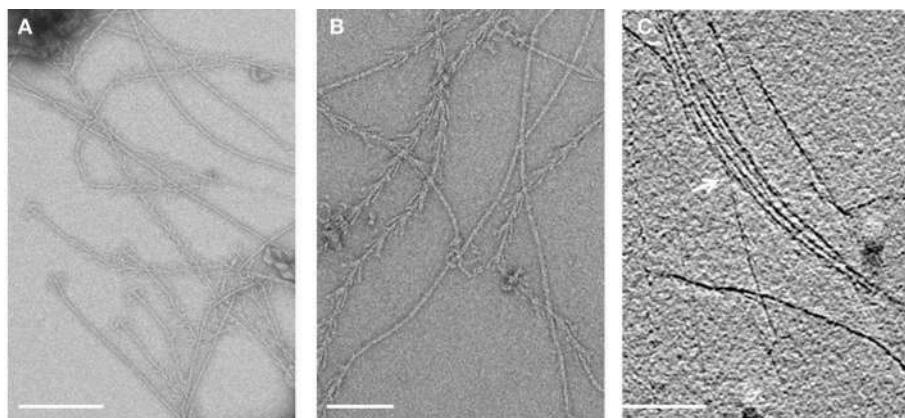


FIGURE 4 | Electron micrographs of isolated hamis. (A,B) Overview of negatively stained hamis with visible triple grappling hooks. The typical barbed-wire structure of the filament is shown, next to plain (probably) stretched filaments. Electron cryo-tomography shows the bundling of the hamis (**C**, white arrow). Scale bar: 500 nm.

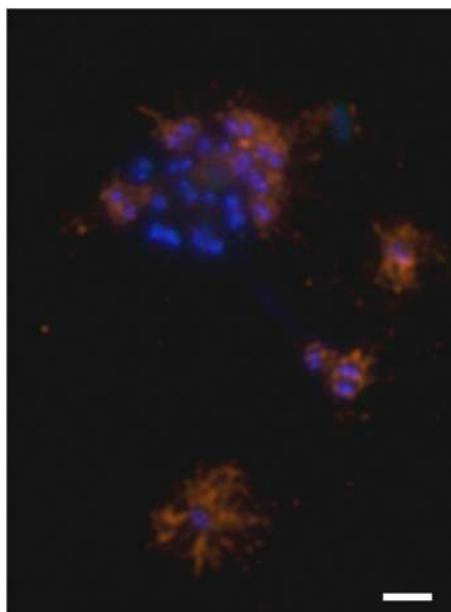


FIGURE 5 | Immuno-staining of cells embedded in the archaeal biofilm. DNA is stained blue (DAPI) and hamis were visualized using a CY3-labeled anti-hamis antibody (orange). The hamis formed a halo around the cell. Scale bar: 2 μ m.

formicicum, DSM3637) and an acylphosphatase, with highest similarity to *Korarchaeum cryptofilum* (strain OPF8) acylphosphatase. One of the unknown proteins belongs to the TraB family, the other shows partial similarities (32%) to a *Sulfolobus solfataricus* (strain ATCC 33909) putative UDP-N-acetylglucosamine-dolichyl-phosphate N-acetylglucosamine-phosphotransferase.

Regarding the complete hamus subunit protein sequence, NCBI NR blastp search (Altschul et al., 1990) revealed no homologous sequences. However, three different architectures to be remotely related to parts of the protein sequence were

revealed by the conserved Domain Architecture Retrieval Tool (http://www.ncbi.nlm.nih.gov/Structure/lexington/docs/cdart_about.html): Archaeal S-layer proteins, hypothetical proteins and acetyl-transferases. Patterns attributed to archaeal S-layer proteins were observed mainly at the N-terminal region of the protein [“S_layer_N”: S-layer like family, N-terminal region (pfam 05123, aa 5-81) and “S_layer_MJ”: S-layer protein, MJ0822 family (TIGR 01564, aa 5-98)]. Closest relatives were found to be N-terminal regions of S-layer proteins from *Methanotorris igneus* (WP_013799875.1; 47% identity, e-value: 6e-05; 56.6 bits), *Methanothermococcus thermolithotrophicus* (CAC83952.2; 44% identity, e-value 5e-04, 53.9 bits) and *Thermococcus eurythermalis* (AIU70131.1; 45%; e-value: 7e-04; 53.5 bits). In contrast to these and other known S-layer proteins, the hamus subunit-protein did not exhibit a typical S_layer_C-terminus pattern.

The Hamus Protein Exhibits a Prominent β -Sandwich Fold and Thus Structurally Resembles Typical Archaeal S-Layers

The secondary structure prediction program PSIPRED identified about 56 beta strands connected by loops and flanked by helical segments in some cases (Figure S7). The first 110 amino acids showed sequence homologies to S-layer proteins from methanogenic Archaea (*Methanococcus voltae*, *Methanococcus maripaludis*, *Methanothermococcus okinawensis*, and *Methanocaldococcus vulcanius*). None of these homologous S-layer proteins have yet been investigated structurally and they do not show any sequence identity to solved structures of S-layer proteins, from bacterial *Clostridium difficile* [Protein Data Bank (pdb) entry 3cvz, (Fagan et al., 2009)], *Clostridium thermocellum* (pdb entry 4qvs) or archaeal S-layer proteins from *Methanoscincina acetivorans* (pdb entry 3u2h, 1l0q; Arbing et al., 2012). Therefore, it was not possible to find a suitable template for reasonable homology modeling just by using a BLAST algorithm.

Putative structural conservation patterns compared to the archaeal S-layer proteins were investigated by applying the

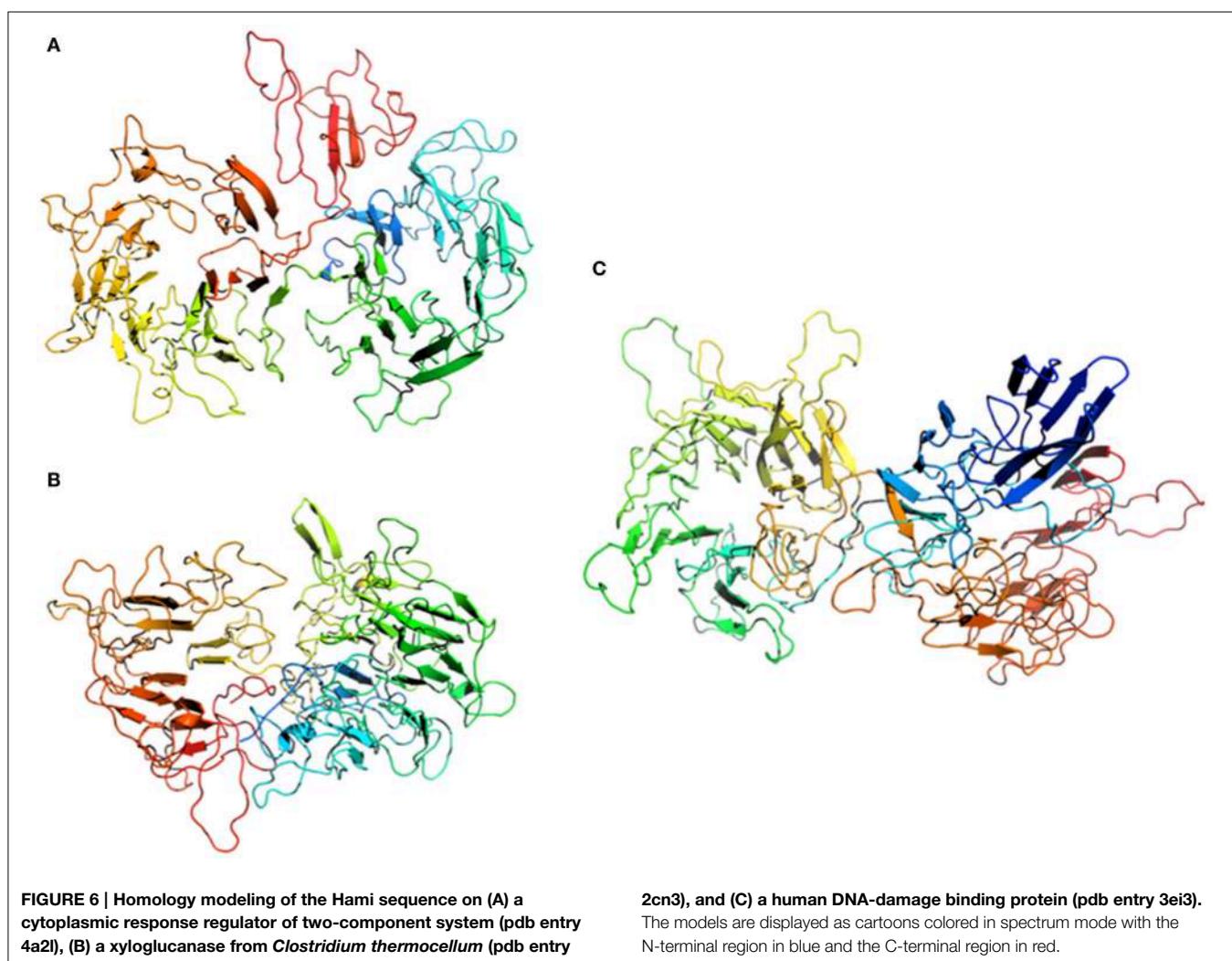
multi-sequence alignment program MAFFT in combination with the topology prediction provided by PSIPRED (pdb entry 3u2h). A parallel overlap between sequence and topology conservation to the DUF1608 domain in the S-layer protein from *M. acetivorans* was identified, when only parts of the hamus protein sequence (Ala334-Asp665) were searched. The *M. acetivorans* S-layer protein MA0829 comprises 671 aa and, similar to the hamus protein, has an N-terminal signal peptide. In addition, it contains the tandem-duplicated DUF1608 domains exclusively found in methanogenic Euryarchaeota.

In a second step, a PSIPRED prediction combined with a fold recognition search (pGenThreader) was performed, which can be applied to individual protein sequences. Three hits were indicated with a p-score of 10^{-8} and an overall coverage of more than 50%. Interestingly, none of the proposed proteins were associated with S-layer proteins: (1) a cytoplasmic response regulator of two-component system, which controls heparin and heparan sulfate acquisition and degradation in the human gut symbiont *Bacteroides thetaiotaomicron* (pdb entry 4a2l; Lowe et al., 2012), (2) a xyloglucanase from *C. thermocellum* (pdb

entry 2cn3, Martinez-Fleites et al., 2006), and (3) a human DNA-damage binding protein (pdb entry 3ei3; Scrima et al., 2008).

We have performed homology modeling based on the first three hits of pGenThreader (Figure 6). In all three models the first 110 aa were removed. Two of the models (Figures 6A,C) show a propeller-like assembly reminiscent of a WD40 repeat. Although not being suitable for homology modeling, also the lower ranking hits proposed by pGenTHREADER revealed a consistent picture with an always returning motif of the WD40 propeller domain, which is often seen in S-layer proteins (Veith et al., 2009; Klingl et al., 2011) and represents one of the most conserved domains for protein-protein interaction. In the model of the hamus protein based on the template of the DNA-damage binding protein (pdb entry 3ei3) this motif would cover the region from aa 230-600, while in the model based on the template of the xyloglucanase from *C. thermocellum* (pdb entry 2cn3) the WD40 like array would correspond to aa 440-800.

Phylogenetic analyses of the hamus protein N-terminus region (aa 1-98) revealed a distinct position within the Euryarchaeota as displayed in Figure S8. Maximum Likelihood



and Neighbor Joining analyses revealed similar results, supporting the hamus subunit protein N-terminus localization as a separate branch (Figure S8).

Discussion

The filamentous cell surface appendages of the cold-loving, uncultivated SM1 Euryarchaeon, *Ca. A. hamiconexum* (Probst et al., 2014) are composed of one major protein species. The sequence of this hamus subunit protein did not show any homologies to currently known proteins involved in microbial fiber-, pili-, flagella-, or archaella formation, but showed similarities to known archaeal S-layer proteins: Besides a typical S-layer N-terminus pattern, the hamus subunit protein was found to be slightly acidic and most likely highly glycosylated, similar to S-layer proteins from e.g., *Acidianus ambivalens* and *Metallosphaera sedula* (Veith et al., 2009). In addition, the hamus subunit protein revealed a prominent beta sheet topology and thus might be primed for self-assembly (Makabe et al., 2006). Previous modeling of beta-rich structures has shown that conformational diversity over a large number of repeats can lead to significantly different self-assemblies therein (such as the formation of fibrils, films, and ribbons) and that their final structure is determined by the way inherent flexibility is maintained via beta-sheet twists and bends (Makabe et al., 2006).

Although there is a remarkable lack of sequence similarity between archaeal S-layer proteins, which also limited the modeling possibilities, β -sandwich structures are obviously typical features of euryarchaeal S-layers (Arbing et al., 2012) and of S-layers from more distantly related archaea, such as Sulfolobales (Veith et al., 2009). Moreover, it is likely that the β -sandwich domains are structurally related to other proteins associated with enveloping functions not only in archaea but also in bacteria, fungi, and viruses, emphasizing the general principle and self-assembly nature of beta-sheet rich proteins (Arbing et al., 2012). In general, S-layer proteins (with a size range between 40 and 210 kDa; Sleytr and Sára, 1997; Sleytr et al., 1997; König et al., 2007) are able to self-assemble into different lattices of oblique, tetragonal or hexagonal architecture—and can even exhibit complex, unusual structures, such as the tetrabrachion of *Staphylothermus marinus* (Engelhardt and Peters, 1998). This S-layer associated protein complex possesses umbrella-like thread morphology and distal branched quadrupled arms (Peters et al., 1995). The S-layer itself depicts p4-symmetry and the long stalks, which anchor the protein in the cytoplasmic membrane, form a 70 nm wide pseudoperiplasm. Furthermore, this membrane anchor is associated with a protease (STABLE; Peters et al., 1995), which might have a metabolic function for this species. Although the resemblance of tetrabrachion and hamis appears striking, no similarity between both proteins could be shown on structural or sequence level.

The assembly and secretion process of bacterial and archaeal S-layers appears multifarious and obviously has evolved independently in some, even closely related microorganisms, such as *Aeromonas* species (Pugsley, 1993; Noonan and Trust, 1995; Thomas and Trust, 1995; Wattiau et al., 1996). By secreting

the premature protein into the periplasm, multimerization of the mature proteins in the cytoplasm is prevented.

This mechanism of translocation strongly resembles the formation of type IV pili (Boot and Pouwels, 1996), where pilin precursors are inserted into the cytoplasmic membrane using the Sec translocation pathway (Arts et al., 2007; Francetic et al., 2007). After cleavage of the positively charged signal peptide the highly hydrophobic N-terminus of the mature pilin is exposed and provides a scaffolding interface for the assembly of the entire pilus structure (Bardy et al., 2003; Pohlchröder et al., 2005; Ng et al., 2007; Albers and Pohlchröder, 2009). A similar process was proposed for the formation of the hamus filaments (Probst et al., 2014).

Although the hamus seems to be formed by one major protein, the presence of other proteins involved in its assembly cannot be excluded. Even supposedly simple systems, such as bacterial type IV pili, are usually composed by several pilins and require a certain set of membrane-associated proteins at the basis of the pilus structure (Mattick, 2002; Craig et al., 2004; Nudleman and Kaiser, 2004). Additional hamus-associated proteins could possibly be identified via future co-immunoprecipitation assays, which could then help to understand assembly procedure of the hamis and their potential function.

Possible functional traits of the hamus subunit protein were revealed by our combined PSIPRED prediction with a fold recognition search, which revealed three hits, a cytoplasmic response regulator, a xyloglucanase or a human DNA-damage binding protein. Although a functional relation to the DNA damage surveillance proteins serving in the initial detection of UV lesions *in vivo* is difficult to draw for the hamus protein, the first two fold-homologs can be prudently associated with a functional relationship. Xyloglucanases, for instance, hydrolyze polysaccharides from the cellulose microfibrils in plant cell walls (Hayashi and Kaida, 2010). The enzymatic reaction is central to the plant carbon cycle and might also indicate a role of the hamis for cell wall degradation and/or carbon metabolism of *Ca. A. hamiconexum*.

Interestingly, the structure of the cytoplasmic response regulator revealed a substantial conformational change on ligand binding and signal transduction, which results in a scissor-like closing. This conformation resembles the barb-like assembly at the hook of the hamus fibril structure, indicative of a signaling role in cell-cell interactions—a possible function of the hamis, which had been discussed earlier (Perras et al., 2014). To date it is unclear, whether the hamis are involved in other processes apart from cell attachment and biofilm formation. In particular, the anchorage and organization of the hamis within the cell wall could not further be resolved using electron cryo-tomography, although it remains without doubt that the hamis span both membranes.

Due to the high similarity of the N-terminal amino acid sequence of the hamus subunit to known archaeal S-layer N-termini, one could even hypothesize about a divergent evolution of the hamus subunit protein from ancestral cell surface proteins and thus a conversion of a layered structure toward a filamentous arrangement—concomitant with a loss of the original surface layer and the development of a second membrane. However, it

remains elusive if the two membranes of *Ca. A. hamiconexum* are different in their organization and whether a structural and compositional adaptation of the outer membrane has occurred due to the lack of an external S-layer. TOF-SIMS has confirmed that the membranes are mainly composed of diglycosidic diethers (C₂₀-C₂₀ archaeol and C₂₀-C₂₅ extended archaeol; Probst et al., 2014). No clear indications of tetraether lipids are observed in the SIMS mass spectra, although trialkyl lipids may be present. Similar to the sodiated series of mono-, di- and tri-glycosylated diether lipid peaks revealed in this study are those of halophilic archaea, *Haloarcula marismortui*, by LC-atmospheric pressure ionization mass spectrometry (de Souza et al., 2009).

In recent years, the field of nanobiotechnology advanced tremendously and is providing an increasing number of strategies to apply natural biomolecules in nanotechnology. For instance, spider silk proved to exhibit extraordinary properties such as strength, elasticity, biocompatibility, and biodegradability and is thus of major interest in nanobiotechnology (Gerritsen, 2002). The spider silk protein could be used in biomedical applications such as coating of implants and drug delivery or scaffolds for tissue engineering (for a review see: Schacht and Scheibel, 2014). However, due to limited availability, experiments with natural spider silk proteins proved complicated (Fox, 1975) and thus the recombinant production of engineered spider silk proteins was pushed forward (Heidebrecht and Scheibel, 2012), including testing of several protein expression systems, with different results in protein yield and property (Chung et al., 2012). Overall, the recombinant expression of the spider silk proteins took researchers decades until a satisfactory result was obtained. Since the hamus subunit protein and the spider silk protein share common features, such as elasticity, robustness and high molecular weight (Moissl et al., 2005; hami: 97 kDa; spider silk protein 250–320 kDa; Sponner et al., 2005; Ayoub et al., 2007), it is not surprising that preliminary overexpression attempts in *Escherichia coli* host strains failed so far. Recombinant expression of the spider silk protein resulted in insufficient

yield, as conventional expression strains lacked the capacity of expressing proteins with high molecular weight (Chung et al., 2012). This obstacle was finally overcome by a metabolically engineered *E. coli* expression host, which overexpressed and assembled the silk protein into a strong fiber (Xia et al., 2010). We suggest that a similar approach may be applicable to successful recombinant expression of the hami, which would pave the way for its exploitation in various fields of nanobiotechnology.

The results presented in this communication emphasize the uniqueness of the altarchaeal hami: their major protein revealed no similarities in sequence and structure to known microbial filament-forming proteins, but showed relationship to archaeal S-layers (in sequence) and beta-sheet protein complexes (in structure), that are widely found in classical macromolecular self-assembled structures. Thus, the hami and the altarchaeal cell wall (with two membranes and without S-layer) could represent a divergent form of cell organization or even a missing link between euryarchaeal ancestors and the current forms of euryarchaeal life.

Acknowledgments

Research on SM1-MSI was supported by the German Research Foundation (Deutsche Forschungsgemeinschaft), grant no. MO 1977/3-1 given to CM-E. AJP was supported by the German National Academic Foundation (Studienstiftung des deutschen Volkes). We thank Uwe-G. Maier for allocation of the EM facility, Marion Debus and Silvia Dobler for technical assistance and Reinhard Wirth and Robert Huber for support and discussions.

Supplementary Material

The Supplementary Material for this article can be found online at: <http://journal.frontiersin.org/article/10.3389/fmib.2015.00543/abstract>

References

- Albers, S.-V., and Pohlschröder, M. (2009). Diversity of archaeal type IV pilin-like structures. *Extremophiles* 13, 403–410. doi: 10.1007/s00792-009-0241-7
- Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990). Basic local alignment search tool. *J. Mol. Biol.* 215, 403–410. doi: 10.1016/S0022-2836(05)80360-2
- Arbing, M. A., Chan, S., Shin, A., Phan, T., Ahn, C. J., Rohlin, L., et al. (2012). Structure of the surface layer of the methanogenic archaeon *Methanosarcina acetivorans*. *Proc. Natl. Acad. Sci. U.S.A.* 109, 11812–11817. doi: 10.1073/pnas.1120595109
- Arts, J., van Boxtel, R., Filloux, A., Tommassen, J., and Koster, M. (2007). Export of the pseudopilin XcpT of the *Pseudomonas aeruginosa* type II secretion system via the signal recognition particle-Sec pathway. *J. Bacteriol.* 189, 2069–2076. doi: 10.1128/JB.01236-06
- Ayoub, N. A., Garb, J. E., Tinghitella, R. M., Collin, M. A., and Hayashi, C. Y. (2007). Blueprint for a high-performance biomaterial: full-length spider dragline silk genes. *PLoS ONE* 2:e514. doi: 10.1371/journal.pone.0000514
- Baker, B. J., Comolli, L. R., Dick, G. J., Hauser, L. J., Hyatt, D., Dill, B. D., et al. (2010). Enigmatic, ultrasmall, uncultivated Archaea. *Proc. Natl. Acad. Sci. U.S.A.* 107, 8806–8811. doi: 10.1073/pnas.0914470107
- Baker, B. J., Tyson, G. W., Webb, R. I., Flanagan, J., Hugenholtz, P., Allen, E. E., et al. (2006). Lineages of acidophilic archaea revealed by community genomic analysis. *Science* 314, 1933–1935. doi: 10.1126/science.1132690
- Bardy, S. L., Eichler, J., and Jarrell, K. F. (2003). Archaeal signal peptides—A comparative survey at the genome level. *Protein Sci.* 12, 1833–1843. doi: 10.1110/ps.03148703
- Beg, Q., Kapoor, M., Mahajan, L., and Hoondal, G. (2001). Microbial xylanases and their industrial applications: a review. *Appl. Microbiol. Biot.* 56, 326–338. doi: 10.1007/s002530100704
- Boot, H. J., and Pouwels, P. H. (1996). Expression, secretion and antigenic variation of bacterial S-layer proteins. *Mol. Microbiol.* 21, 1117–1123. doi: 10.1046/j.1365-2958.1996.711442.x
- Borrel, G., Parisot, N., Harris, H. M., Peyretaille, E., Gaci, N., Tottey, W., et al. (2014). Comparative genomics highlights the unique biology of *Methanomassiliicoccales*, a Thermoplasmatales-related seventh order of methanogenic archaea that encodes pyrrolysine. *BMC Genomics* 15:679. doi: 10.1186/1471-2164-15-679
- Chung, H., Kim, T. Y., and Lee, S. Y. (2012). Recent advances in production of recombinant spider silk proteins. *Curr. Opin. Biotechnol.* 23, 957–964. doi: 10.1016/j.copbio.2012.03.013

- Comolli, L. R., Baker, B. J., Downing, K. H., Siegerist, C. E., and Banfield, J. F. (2009). Three-dimensional analysis of the structure and ecology of a novel, ultra-small archaeon. *ISME J.* 3, 159–167. doi: 10.1038/ismej.2008.99
- Comolli, L. R., and Banfield, J. F. (2014). Inter-species interconnections in acid mine drainage microbial communities. *Front. Microbiol.* 5:367. doi: 10.3389/fmicb.2014.00367
- Craig, L., Pique, M. E., and Tainer, J. A. (2004). Type IV pilus structure and bacterial pathogenicity. *Nat. Rev. Microbiol.* 2, 363–378. doi: 10.1038/nrmicro885
- de Souza, L. M., Müller-Santos, M., Iacomini, M., Gorin, P. A. J., and Sasaki, G. L. (2009). Positive and negative tandem mass spectrometric fingerprints of lipids from the halophilic Archaea *Haloarcula marismortui*. *J. Lipid Res.* 50, 1363–1373. doi: 10.1194/jlr.M800478-JLR200
- Dridi, B., Fardeau, M.-L., Ollivier, B., Raoult, D., and Drancourt, M. (2012). *Methanomassiliicoccus luminyensis* gen. nov., sp. nov., a methanogenic archaeon isolated from human faeces. *Int. J. Syst. Evol. Microbiol.* 62, 1902–1907. doi: 10.1099/ijs.0.033712-0
- Dutta, K., and Kundu, P. P. (2014). A review on aromatic conducting polymers-based catalyst supporting matrices for application in microbial fuel cells. *Polym. Rev.* 54, 401–435. doi: 10.1080/15583724.2014.881372
- Eichler, J. (2003). Facing extremes: archaeal surface-layer (glyco)proteins. *Microbiology* 149, 3347–3351. doi: 10.1099/mic.0.26591-0
- Engelhardt, H., and Peters, J. (1998). Structural research on surface layers: a focus on stability, surface layer homology domains, and surface layer-cell wall interactions. *J. Struct. Biol.* 124, 276–302. doi: 10.1006/jsb1.1998.4070
- Fagan, R. P., Albesa-Jové, D., Qazi, O., Svergun, D. I., Brown, K. A., and Fairweather, N. F. (2009). Structural insights into the molecular organization of the S-layer from *Clostridium difficile*. *Mol. Microbiol.* 71, 1308–1322. doi: 10.1111/j.1365-2958.2009.06603.x
- Fox, L. R. (1975). Cannibalism in natural populations. *Annu. Rev. Ecol. Syst.* 6, 87–106. doi: 10.1146/annurev.es.06.110175.000511
- Francetic, O., Buddelmeijer, N., Lewenza, S., Kumamoto, C. A., and Pugsley, A. P. (2007). Signal recognition particle-dependent inner membrane targeting of the PilG pseudopilin component of a type II secretion system. *J. Bacteriol.* 189, 1783–1793. doi: 10.1128/JB.01230-06
- Gerritsen, V. B. (2002). The Tiptoe of an airbus. *Protein Spotlight Swiss. Prot.* 24, 1–2.
- Hasan, F., Shah, A. A., and Hameed, A. (2006). Industrial applications of microbial lipases. *Enzyme Microb. Technol.* 39, 235–251. doi: 10.1016/j.enzmictec.2005.10.016
- Hayashi, T., and Kaida, R. (2010). Functions of xyloglucan in plant cells. *Mol. Plant* 4, 17–24. doi: 10.1093/mp/ssq063
- Heidebrecht, A., and Scheibel, T. (2012). Recombinant production of spider silk proteins. *Adv. Appl. Microbiol.* 82, 115–153. doi: 10.1016/B978-0-12-407679-2.00004-1
- Henneberger, R., Moissl, C., Amann, T., Rudolph, C., and Huber, R. (2006). New insights into the lifestyle of the cold-loving SM1 euryarchaeon: natural growth as a monospecies biofilm in the subsurface. *Appl. Environ. Microbiol.* 72, 192–199. doi: 10.1128/AEM.72.1.192-199.2006
- Iino, T., Tamaki, H., Tamazawa, S., Ueno, Y., Ohkuma, M., Suzuki, K.-I., et al. (2013). *Candidatus Methanogranum caenicola*: a novel methanogen from the anaerobic digested sludge, and proposal of *Methanomassiliicoccaceae* fam. nov. and *Methanomassiliicoccales* ord. nov., for a aethanogenic lineage of the class *Thermoplasmata*. *Microbes Environ.* 28, 244–250. doi: 10.1264/jsme2.ME12189
- Junglas, B., Briegel, A., Burghardt, T., Walther, P., Wirth, R., Huber, H., et al. (2008). *Ignicoccus hospitalis* and *Nanoarchaeum equitans*: ultrastructure, cell-cell interaction, and 3D reconstruction from serial sections of freeze-substituted cells and by electron cryotomography. *Arch. Microbiol.* 190, 395–408. doi: 10.1007/s00203-008-0402-6
- Kantor, R. S., Wrighton, K. C., Handley, K. M., Sharon, I., Hug, L. A., Castelle, C. J., et al. (2013). Small genomes and sparse metabolisms of sediment-associated bacteria from four candidate phyla. *MBio* 4, e00708–e00713. doi: 10.1128/mBio.00708-13
- Klingl, A., Moissl-Eichinger, C., Wanner, G., Zweck, J., Huber, H., Thomm, M., et al. (2011). Analysis of the surface proteins of *Acidithiobacillus ferrooxidans* strain SP5/1 and the new, pyrite-oxidizing *Acidithiobacillus* isolate HV2/2, and their possible involvement in pyrite oxidation. *Arch. Microbiol.* 193, 867–882. doi: 10.1007/s00203-011-0720-y
- Klingl, A. (2014). S-layer and cytoplasmic membrane—exceptions from the typical archaeal cell wall with a focus on double membranes. *Front. Microbiol.* 5:624. doi: 10.3389/fmicb.2014.00624
- König, H., Rachel, R., and Claus, H. (2007). “Proteinaceous surface layers of archaea: ultrastructure and biochemistry,” in *Archaea: Molecular and Cell Biology*, ed R. Cavicchioli (Washington, DC: American Society of Microbiology Press), 315–340.
- Kremer, J. R., Mastronarde, D. N., and McIntosh, J. R. (1996). Computer visualization of three-dimensional image data using IMOD. *J. Struct. Biol.* 116, 71–76. doi: 10.1006/jbsb.1996.0013
- Laemmli, U. K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* 227, 680–685. doi: 10.1038/227680a0
- Lobley, A., Sadowski, M. I., and Jones, D. T. (2009). pGenTHREADER and pDomTHREADER: new methods for improved protein fold recognition and superfamily discrimination. *Bioinformatics* 25, 1761–1767. doi: 10.1093/bioinformatics/btp302
- Lowe, E. C., Baslé, A., Czjzek, M., Firbank, S. J., and Bolam, D. N. (2012). A scissor blade-like closing mechanism implicated in transmembrane signaling in a *Bacteroides* hybrid two-component system. *Proc. Natl. Acad. Sci. U.S.A.* 109, 7298–7303. doi: 10.1073/pnas.1200479109
- Makabe, K., McElheny, D., Tereshko, V., Hilyard, A., Gawlik, G., Yan, S., et al. (2006). Atomic structures of peptide self-assembly mimics. *Proc. Natl. Acad. Sci. U.S.A.* 103, 17753–17758. doi: 10.1073/pnas.0606690103
- Martinez-Fleites, C., Guerreiro, C. I., Baumann, M. J., Taylor, E. J., Prates, J. A., Ferreira, L. M., et al. (2006). Crystal structures of *Clostridium thermocellum* xyloglucanase, XGH74A, reveal the structural basis for xyloglucan recognition and degradation. *J. Biol. Chem.* 281, 24922–24933. doi: 10.1074/jbc.M603583200
- Mattick, J. S. (2002). Type IV Pili and twitching motility. *Annu. Rev. Microbiol.* 56, 289–314. doi: 10.1146/annurev.micro.56.012302.160938
- Milić, M., Leitinger, G., Pavičić, I., Žebić Avdičević, M., Dobrović, S., Goessler, W., et al. (2015). Cellular uptake and toxicity effects of silver nanoparticles in mammalian kidney cells. *J. Appl. Toxicol.* 35, 581–592. doi: 10.1002/jat.3081
- Moissl-Eichinger, C., Henneberger, R., and Huber, R. (2012). “SM1: a cold-loving archaeon with powerful nano-grappling hooks,” in *Extremophiles: Microbiology and Biotechnology*, ed R. P. Anitori (Norfolk, UK: Horizon Scientific Press), 77.
- Moissl, C., Rachel, R., Briegel, A., Engelhardt, H., and Huber, R. (2005). The unique structure of archaeal ‘hami’, highly complex cell appendages with nano-grappling hooks. *Mol. Microbiol.* 56, 361–370. doi: 10.1111/j.1365-2958.2005.04294.x
- Ng, S. Y., Chaban, B., Vandyke, D. J., and Jarrell, K. F. (2007). Archaeal signal peptidases. *Microbiology* 153, 305–314. doi: 10.1099/mic.0.2006/003087-0
- Noonan, B., and Trust, T. J. (1995). Molecular analysis of an A-protein secretion mutant of *Aeromonas salmonicida* reveals a surface layer-specific protein secretion pathway. *J. Mol. Biol.* 248, 316–327. doi: 10.1016/S0022-2836(95)80053-0
- Nudelman, E., and Kaiser, D. (2004). Pulling together with type IV pili. *J. Mol. Microbiol. Biotechnol.* 7, 52–62. doi: 10.1159/000077869
- Perras, A. K., Wanner, G., Klingl, A., Mora, M., Auerbach, A. K., Heinz, V., et al. (2014). Grappling archaea: ultrastructural analyses of an uncultivated, cold-loving archaeon, and its biofilm. *Front. Microbiol.* 5:397. doi: 10.3389/fmicb.2014.00397
- Peters, J., Nitsch, M., Kühlmorgen, B., Golbik, R., Lupas, A., Kellermann, J., et al. (1995). Tetraetherion: a Filamentous archaeabacterial surface protein assembly of unusual structure and extreme stability. *J. Mol. Biol.* 245, 385–401. doi: 10.1006/jmbi.1994.0032
- Pohlschröder, M., Giménez, M. I., and Jarrell, K. F. (2005). Protein transport in Archaea: Sec and twin arginine translocation pathways. *Curr. Opin. Microbiol.* 8, 713–719. doi: 10.1016/j.mib.2005.10.006
- Probst, A. J., Holman, H.-Y. N., Desantis, T. Z., Andersen, G. L., Birarda, G., Bechtel, H. A., et al. (2013). Tackling the minority: sulfate-reducing bacteria in an archaea-dominated subsurface biofilm. *ISME J.* 7, 635–651. doi: 10.1038/ismej.2012.133
- Probst, A. J., Weinmaier, T., Raymann, K., Perras, A., Emerson, J. B., Rattei, T., et al. (2014). Biology of a widespread uncultivated archaeon that contributes

- to carbon fixation in the subsurface. *Nat. Commun.* 5:5497. doi: 10.1038/ncomms6497
- Pugsley, A. P. (1993). The complete general secretory pathway in gram-negative bacteria. *Microbiol. Rev.* 57, 50.
- Rachel, R., Wyschkony, I., Riehl, S., and Huber, H. (2002). The ultrastructure of Ignicoccus: evidence for a novel outer membrane and for intracellular vesicle budding in an archaeon. *Archaea* 1, 9–18. doi: 10.1155/2002/307480
- Rudolph, C., Moissl, C., Henneberger, R., and Huber, R. (2004). Ecology and microbial structures of archaeal/bacterial strings-of-pears communities and archaeal relatives thriving in cold sulfidic springs. *FEMS Microbiol. Ecol.* 50, 1–11. doi: 10.1016/j.femsec.2004.05.006
- Rudolph, C., Wanner, G., and Huber, R. (2001). Natural communities of novel archaea and bacteria growing in cold sulfurous springs with a string-of-pears-like morphology. *Appl. Environ. Microbiol.* 67, 2336–2344. doi: 10.1128/AEM.67.5.2336-2344.2001
- Schacht, K., and Scheibel, T. (2014). Processing of recombinant spider silk proteins into tailor-made materials for biomaterials applications. *Curr. Opin. Biotechnol.* 29, 62–69. doi: 10.1016/j.copbio.2014.02.015
- Scrima, A., Koničková, R., Czyzewski, B. K., Kawasaki, Y., Jeffrey, P. D., Groisman, R., et al. (2008). Structural Basis of UV DNA-Damage Recognition by the DDB1–DDB2 Complex. *Cell* 135, 1213–1223. doi: 10.1016/j.cell.2008.10.045
- Sharon, I., Morowitz, M. J., Thomas, B. C., Costello, E. K., Relman, D. A., and Banfield, J. F. (2013). Time series community genomics analysis reveals rapid shifts in bacterial species, strains, and phage during infant gut colonization. *Genome Res.* 23, 111–120. doi: 10.1101/gr.142315.112
- Sharon, I., and Banfield, J. F. (2013). Microbiology. Genomes from metagenomics. *Science* 342, 1057–1058. doi: 10.1126/science.1247023
- Sleytr, U. B., Messner, P., Pum, D., and Sára, M. (1988). “Crystalline bacterial cell surface layers,” in *Proceedings of EMBO Workshop on Crystalline Bacterial Surface Layers* (Berlin: Springer Verlag). doi: 10.1007/978-3-642-73537-0
- Sleytr, U. B., Bayley, H., Sára, M., Breitwieser, A., Küpcü, S., Mader, C., et al. (1997). VI. Applications of S-layers1. *FEMS Microbiol. Rev.* 20, 151–175. doi: 10.1016/S0168-6445(97)00044-2
- Sleytr, U. B., and Sára, M. (1997). Bacterial and archaeal S-layer proteins: structure-function relationships and their biotechnological applications. *Trends Biotechnol.* 15, 20–26. doi: 10.1016/S0167-7799(96)10063-9
- Sponner, A., Schlott, B., Vollrath, F., Unger, E., Grosse, F., and Weisshart, K. (2005). Characterization of the protein components of *Nephila clavipes* dragline silk. *Biochemistry* 44, 4727–4736. doi: 10.1021/bi047671k
- Tamura, K., Stecher, G., Peterson, D., Filipski, A., and Kumar, S. (2013). MEGA6: molecular evolutionary genetics analysis version 6.0. *Mol. Biol. Evol.* 30, 2725–2729. doi: 10.1093/molbev/mst197
- Teubl, B. J., Leitinger, G., Schneider, M., Lehr, C. M., Fröhlich, E., Zimmer, A., et al. (2014). The buccal mucosa as a route for TiO₂ nanoparticle uptake. *Nanotoxicology* 29, 1–9. doi: 10.3109/17435390.2014.921343
- Thomas, S. R., and Trust, T. J. (1995). A specific PulD homolog is required for the secretion of paracrystalline surface array subunits in *Aeromonas hydrophila*. *J. Bacteriol.* 177, 3932–3939.
- Tyson, G. W., Chapman, J., Hugenholtz, P., Allen, E. E., Ram, R. J., Richardson, P. M., et al. (2004). Community structure and metabolism through reconstruction of microbial genomes from the environment. *Nature* 428, 37–43. doi: 10.1038/nature02340
- Weith, A., Klingl, A., Zolghadr, B., Lauber, K., Mentele, R., Lottspeich, F., et al. (2009). *Acidianus*, *Sulfolobus* and *Metallosphaera* surface layers: structure, composition and gene expression. *Mol. Microbiol.* 73, 58–72. doi: 10.1111/j.1365-2958.2009.06746.x
- Wattiau, P., Woestyn, S., and Cornelis, G. R. (1996). Customized secretion chaperones in pathogenic bacteria. *Mol. Microbiol.* 20, 255–262. doi: 10.1111/j.1365-2958.1996.tb02614.x
- Wilkins, M. R., Lindskog, I., Gasteiger, E., Bairoch, A., Sanchez, J. C., Hochstrasser, D. F., et al. (1997). Detailed peptide characterization using PEPTIDEMASS—a World-Wide-Web-accessible tool. *Electrophoresis* 18, 403–408. doi: 10.1002/elps.1150180314
- Xia, X.-X., Qian, Z.-G., Ki, C. S., Park, Y. H., Kaplan, D. L., and Lee, S. Y. (2010). Native-sized recombinant spider silk protein produced in metabolically engineered *Escherichia coli* results in a strong fiber. *Proc. Natl. Acad. Sci. U.S.A.* 107, 14059–14063. doi: 10.1073/pnas.1003366107
- Yelton, A. P., Comolli, L. R., Justice, N. B., Castelle, C., Denef, V. J., Thomas, B. C., et al. (2013). Comparative genomics in acid mine drainage biofilm communities reveals metabolic and structural differentiation of co-occurring archaea. *BMC Genomics* 14:485. doi: 10.1186/1471-2164-14-485

Conflict of Interest Statement: Patent is pending on the “microbial nano-tool” (Pending European Patent Application No. EP15166985.0). This patent was filed jointly by the AKP, AJP and CME. The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2015 Perras, Daum, Ziegler, Takahashi, Ahmed, Wanner, Klingl, Leitinger, Kolb-Lenz, Gribaldo, Auerbach, Mora, Probst, Bellack and Moissl-Eichinger. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Archaeal type IV pili and their involvement in biofilm formation

Mechthild Pohlschroder * and Rianne N. Esquivel

Department of Biology, University of Pennsylvania, Philadelphia, PA, USA

Type IV pili are ancient proteinaceous structures present on the cell surface of species in nearly all bacterial and archaeal phyla. These filaments, which are required for a diverse array of important cellular processes, are assembled employing a conserved set of core components. While type IV pilins, the structural subunits of pili, share little sequence homology, their signal peptides are structurally conserved allowing for *in silico* prediction. Recently, *in vivo* studies in model archaea representing the euryarchaeal and crenarchaeal kingdoms confirmed that several of these pilins are incorporated into type IV adhesion pili. In addition to facilitating surface adhesion, these *in vivo* studies also showed that several predicted pilins are required for additional functions that are critical to biofilm formation. Examples include the subunits of *Sulfolobus acidocaldarius* Ups pili, which are induced by exposure to UV light and promote cell aggregation and conjugation, and a subset of the *Haloflexax volcanii* adhesion pilins, which play a critical role in microcolony formation while other pilins inhibit this process. The recent discovery of novel pilin functions such as the ability of haloarchaeal adhesion pilins to regulate swimming motility may point to novel regulatory pathways conserved across prokaryotic domains. In this review, we will discuss recent advances in our understanding of the functional roles played by archaeal type IV adhesion pili and their subunits, with particular emphasis on their involvement in biofilm formation.

OPEN ACCESS

Edited by:

Sonja-Verena Albers,
University of Freiburg, Germany

Reviewed by:

Vladimir Pelicic,
Imperial College London, UK
Ken Jarrell,
Queen's University, Canada
Olivera Francetic,
Institut Pasteur, France

*Correspondence:

Mechthild Pohlschroder,
Department of Biology, University of
Pennsylvania, Philadelphia,
PA 19104, USA
pohlschr@sas.upenn.edu

Specialty section:

This article was submitted to Microbial Physiology and Metabolism, a section of the journal *Frontiers in Microbiology*

Received: 05 January 2015

Accepted: 20 February 2015

Published: 24 March 2015

Citation:

Pohlschroder M and Esquivel RN (2015) Archaeal type IV pili and their involvement in biofilm formation. *Front. Microbiol.* 6:190.
doi: 10.3389/fmicb.2015.00190

Introduction

Archaea and bacteria alike cope with stress by forming biofilms, multicellular communities encased in a structure consisting of polysaccharide layers (Monds and O'Toole, 2009; Haussler and Fuqua, 2013; Orell et al., 2013a). The initial steps in this process involve adherence to surfaces and interactions between cells (Wozniak and Parsek, 2014). A diverse set of surface filaments facilitate these interactions with biotic and abiotic surfaces. Within the bacterial domain, these filaments include the chaperone-usher pathway dependent pili, which are associated with the outer membranes of some gram-negative bacteria (Busch and Waksman, 2012) and the sortase-dependent cell wall-associated pili found in many gram-positive bacteria (Schneewind and Missiakas, 2013). Additionally, structurally conserved amyloid fibers consisting of autoaggregating polymeric fibrils that are composed of folded β -sheets have been identified in many bacterial phyla and may also play roles in archaeal biofilm formation (Blanco et al., 2012; Chimileski et al., 2014). The archaea also express a wide variety of additional, structurally diverse adhesion filaments. These filaments include the Mth60 fimbriae of *Methanothermobacter thermautotrophicus* (Thoma et al., 2008), 5 nm diameter filaments that adhere to organic surfaces such as chitin, the cannulae, 25 nm diameter hollow tubes that are found on the hyperthermophilic *Pyrodictium*, which mediate interactions between

cells (Horn et al., 1999), and the hami of SM1 euryarchaea, which form grappling hooks that allow attachment to surfaces in cold sulfurous marshes (Perras et al., 2014).

Based on currently available genomic analyses and experimental data, only one known type of adhesion filament, the type IV pilus, appears to be present in nearly all phyla across both prokaryotic domains (Pohlschroder et al., 2011; Giltner et al., 2012; Szabo and Pohlschroder, 2012). These structures, which were first identified, and have been best studied in gram-negative bacteria, have been defined as filamentous protein complexes composed of subunits known as type IV pilins (Burrows, 2012a). The precursors of type IV pilins have N-terminal signal peptides that target them to the Sec pathway, which transports them across the plasma membrane where they are processed by a prepilin peptidase – either PilD, in bacteria, or PibD, in archaea – prior to incorporation into the pilus (Strom et al., 1993; Albers et al., 2003; Bardy and Jarrell, 2003). Unlike other Sec signal peptides, where the processing site follows the hydrophobic (H)-domain, the PilD/PibD processing site precedes the H-domain. The hydrophobic portions of these N-terminal domains, upon processing, serve as a scaffold at the core of the pilus to facilitate assembly. In addition to the conserved signal peptidases, the only components required for type IV pilus biosynthesis in all organisms appear to be PilB, an ATPase, and PilC, a transmembrane protein that has been proposed to anchor the pilus to the membrane (Nunn et al., 1990; Pelicic, 2008; Lassak et al., 2012; Takhar et al., 2013). Although additional components involved in pilus biosynthesis have been identified, none are conserved across the prokaryotic phyla.

In addition to surface adhesion, various type IV pili have evolved a variety of functions including nutrient scavenging (Zolghadr et al., 2011), mediating electron transport (Lovley, 2012), and facilitating protein transport across the periplasmic space of some gram-negative bacteria (Nivaskumar and Francetic, 2014). In fact, some type IV pilus-like structures have evolved in such a way that their functions no longer include surface adhesion. For example, the competence system of many gram-positive bacteria, which is critical for DNA uptake, has no effect on cell-cell interactions or surface adhesion, although assembly of these structures does require the type IV pilus biosynthesis machinery (Chen and Dubnau, 2004; Chen et al., 2006).

The development of several model archaeal systems (Leigh et al., 2011) as well as *in silico* tools that allow rapid, accurate identification of type IV pilins (Szabó et al., 2007), has allowed significant advances in the understanding of the roles played by type IV pili in the cellular processes of archaea. In this review, we will highlight recent advances that have been made in characterizing a diverse set of archaeal type IV pili, with particular attention paid to how these structures facilitate archaeal biofilm formation. These studies, which have focused primarily on a subset of type IV pili in the model crenarchaea, *Sulfolobus solfataricus* and *Sulfolobus acidocaldarius*, as well as the euryarchaeal model systems, *Methanococcus maripaludis* and *Haloferax volcanii*, have illuminated the strategies that have allowed organisms to thrive in the extreme environments that archaea often inhabit. The study of these evolutionarily

conserved surface filaments in a diverse array of archaea has highlighted the divergent as well as the conserved aspects of type IV pili across a highly diverse range of prokaryotes, both within and across domains.

Evolutionary Relationship between Type IV Pili and Type IV Pilus-Like Structures

Type IV Pili

Biofilm formation is likely an evolutionarily ancient strategy adapted to protect microorganisms against highly stressful environmental conditions (Lopez et al., 2010). The presence of type IV pili in nearly all the phyla of bacteria and archaea (Szabó et al., 2007; Imam et al., 2011) indicates that these surface filaments are also ancient, perhaps being a prerequisite for the adaptations that allowed the establishment of these multicellular structures early on during the evolutionary history of the prokaryotes. The fact that most prokaryotic organisms still express type IV pili underscores the key roles they have played in the biological processes that allowed prokaryotes to flourish. Membrane-associated type IV pilins, pilus subunits, if not part of a pilus can promote adhesion in some organisms (Esquivel and Pohlschroder, 2014), suggesting that the earliest form of surface attachment was facilitated through membrane-anchored adhesion proteins that evolved to be incorporated into surface filaments, perhaps providing a mechanism that allowed cells to adhere more efficiently to surfaces.

Possibly reflecting the different types of surfaces to which type IV pili can adhere, the amino acid sequences of the pilins are highly diverse (Szabó et al., 2007; Imam et al., 2011). In fact, the only segment of the protein sequence that is largely conserved among all type IV pilins is limited to the unique prepilin-peptidase processing site in the N-terminal Sec signal peptide of pilin precursors (Giltner et al., 2012). The diversity of pilin sequences may also be due in part to the fact that many type IV pili have evolved to carry out additional functions. For example, *Geobacter sulfurreducens* expresses a type IV pilus that facilitates the transfer of electrons to extracellular electron acceptors such as insoluble metals (Reguera et al., 2005). Some additional pilus functions take advantage of the fact that many type IV pili have acquired the ability to retract, thus providing cells with a molecular ratchet that allows them to move along surfaces in a process that has come to be known as twitching motility (Bradley, 1972; Burrows, 2012b). The ability of pili to retract may also play a key role in maintaining close contact between cells while they exchange DNA (Ajon et al., 2011), and might also facilitate the uptake of viruses that attach to these surface structures (Kim et al., 2012). In some gram-negative bacteria, type IV pili, which are anchored to the inner membrane and cross the outer membrane through a secretin pore, can facilitate protein transport to the extracellular environment. For example, in *Vibrio cholera*, toxin co-regulated type IV pili are not only essential for surface adhesion, but are also required for secretion of the soluble colonization factor, TcpF (Kirn et al., 2003).

Type IV Pilus-Like Structures

Some cell surface structures are assembled using homologs of the same components required for the biosynthesis of type IV pili, and share structural similarities with type IV pili, but have functions that do not include adhesion. For example, the piston-like structure of the type II secretion machinery of gram-negative bacteria, such as *Klebsiella oxytoca*, which is thought to facilitate the transport of proteins across the periplasm, is composed of type IV pilin-like proteins, and its substrates are secreted via pores composed of secretin (Giltner et al., 2012; Campos et al., 2013; Nivaskumar and Francetic, 2014). In fact, although this structure is not found on the cell surface of wild-type cells, the overexpression of its major piston subunit can result in the production of surface filaments (Sauvonnet et al., 2000; Nivaskumar and Francetic, 2014). In gram-positive bacteria, the assembly of a competence system that facilitates DNA uptake requires the presence of the PilD homolog, ComC, to process the subunits of a high molecular weight DNA-binding surface structure, as well as the PilB and PilC homologs, ComGA and ComGB, respectively. DNA binding pili have been visualized in *Streptococcus pneumoniae* (Laurenceau et al., 2013; Balaban et al., 2014). However, in the well-studied Com system of *Bacillus subtilis* these structures do not promote adhesion nor do they form surface filaments under the range of conditions that have been tested (Chen and Dubnau, 2004).

The key components required for the biosynthesis of archaeal flagella – rotating surface structures that drive swimming motility – are homologous to the core components of the type IV pilus-biosynthesis machinery (Jarrell and McBride, 2008). These structures are also composed of subunits that are processed by a prepilin peptidase, and the precursors of these flagellin subunits have signal peptides that are structurally similar to those of pilin precursors. It is likely that the flagella, which require several additional components to function properly (see review on flagella in this special issue), evolved from simpler type IV pili, a subset of which appears to promote surface adhesion in archaea, perhaps by overcoming surface tension barriers (Davey and O'Toole, 2000; Lassak et al., 2012). The lack of similarity between archaeal flagella and the flagella of bacteria has lead to a proposal to change the name of these structures to archaella (Jarrell and Albers, 2012). However, considering that flagella have not been defined based on their composition or the composition of their biosynthesis machineries, but rather by their function (Diniz et al., 2012), it seems fitting that these rotating surface structures remain known as flagella, just as these structures with analogous functions in bacteria and eukaryotes are. Consistent with this argument, the nomenclature used for pili seems to follow a similar logic. Pili are generally regarded as surface filaments that facilitate adhesion to biotic or abiotic surfaces. However, various types of pili, including the type IV pili, as well as others, such as the sortase-dependent pili, are named based on the composition of their biosynthesis machinery and structural subunits. Similarly, the name “archaeal flagella” clearly distinguishes these archaeal motility structures from the “bacterial flagella” which in turn are easily distinguished from the “eukaryotic flagella,” while these names still indicate

that these are all surface structures that propel swimming motility.

Finally, a fascinating surface structure that is assembled using the same core pilus biosynthesis components described above is the *S. solfataricus* bindosome, which plays an important role in nutrient uptake (Zolghadr et al., 2007). The subunits of this structure, while containing processing sites that are typical of a type IV pilin, are at least four times the size of an average pilin and exhibit significant homology to substrate-binding proteins. Hence, these proteins are unlikely to have evolved from type IV pilins but rather were originally substrate-binding proteins that seem to have hijacked the type IV pilus biosynthesis machinery to assemble a surface structure that allows more efficient substrate scavenging.

In silico analyses have predicted a large number of genes encoding proteins that contain putative type IV pilin signal peptides in the genomes of species representing all archaeal phyla. These predicted proteins vary in size from less than 150 amino acids to well over 1000 (Szabó et al., 2007). Potential functions for most of these predicted proteins have not been identified, as this vast array of potential type IV pilus subunits has only recently begun to be characterized. As we learn more about the functions of these structures, we may also begin to tease out the evolutionary history of the type IV pili and their related structures.

Pilus Biosynthesis

Pilin Transport and Signal Peptide Processing

Type IV pilin precursors are transported across the cytoplasmic membrane via the Sec pathway in an SRP-dependent manner (Arts et al., 2007; Francetic et al., 2007). Like other Sec signal peptides, all type IV pilin signal peptides contain a charged N-terminus followed by a hydrophobic (H) domain. However, rather than following the H-domain, as in the signal peptides of other Sec substrates, the peptidase-processing site in the signal peptides of prepilins precedes it (Strom et al., 1993; Albers et al., 2003; Ng et al., 2009). Processing of the pilin precursors occurs either simultaneously with, or following, the lateral insertion of the H-domain into the cytoplasmic membrane, which, upon transport of the rest of the pilin through the Sec pore, may anchor the pilin to the membrane before it is incorporated into a pilus (Giltner et al., 2012).

In bacteria, the prepilin peptidase PilD is a bifunctional enzyme that cleaves and *N*-methylates type IV pilin precursors (Strom et al., 1993). While PilD exhibits only a low degree of homology with the archaeal prepilin peptidase, PibD, which lacks a region homologous to the bacterial protein domain required for pilin methylation, the prepilin peptidases of both domains are integral membrane aspartic acid proteases. Both proteases have similar catalytic sites, each containing two aspartic acids, of which the second one is part of a conserved GxHyD motif identified in many aspartic acid proteases (Szabo et al., 2006; Henche et al., 2014). Site-directed mutagenesis of the *Sulfolobales* *pibD* codons encoding the conserved aspartic acids resulted in

mutant peptidases that are unable to process pilins as well as pilin-like subunits. Similarly, these conserved aspartic acids are required for processing of flagellins by the PibD homolog, FlaK, in *M. maripaludis* (Bardy and Jarrell, 2003; Szabó et al., 2007; Esquivel et al., 2013).

Although all type IV pilins and type IV pilin-like proteins share specific structural similarities, including a prepilin peptidase-processing site that precedes the H-domain, the importance of some commonly occurring characteristics within the tripartite structure of the signal peptide varies depending upon the particular prepilin peptidase involved in processing the pilin precursor (**Table 1**). For example, while proper processing of all type IV pilin precursors appears to require a G/A/S at the -1 position, a charged amino acid at position -2 is only critical in the recognition of archaeal pilin precursors by the prepilin peptidase (Szabó et al., 2007; Imam et al., 2011). Conversely, while the fifth amino acid of the mature bacterial pilin is nearly always a glutamate or aspartate, only a subset of archaeal pilins contain a negatively charged amino acid at the +5 position. Using these criteria, as well as a few additional conserved parameters, including the length and hydrophobicity of the H-domain of the signal peptide and the position of the processing motif, rule-based algorithms for genome-wide identification of genes that encode prepilins and prepilin-like proteins in bacteria (PilFind) and archaea (FlaFind) were successfully developed (Szabó et al., 2007; Imam et al., 2011; Esquivel et al., 2013).

Most archaea express one prepilin peptidase that recognizes all prepilins as well as prepilin-like proteins such as the flagellins and bindosome subunits (Albers et al., 2003). However, a subset of euryarchaea expresses an additional, much larger, prepilin peptidase (EppA). Although substrates recognized by this second prepilin peptidase are predicted by FlaFind, they contain a highly conserved +1 glutamine, confirmed by top-down mass spectrometry (Ng et al., 2010), and the +5 glutamic acid as is commonly, and primarily, found in bacterial type IV pilins (Giltner et al., 2012). In fact, the *M. maripaludis* FlaK paralog, EppA, requires the presence of the +1 glutamine to process a pilin precursor, and it cannot process the *M. maripaludis* flagellin subunit (Szabó et al., 2007; Ng et al., 2009). Interestingly, while the PibD prepilin peptidases of *S. solfataricus* and *H. volcanii* appear to exhibit little stringency with regard to amino acid residues at positions +1 and +3 of the pilin precursors, the *M. maripaludis* FlaK does not process precursors having a +1 glutamine, preventing it from processing EppA substrates (Ng et al., 2009). It is likely that in other euryarchaea, which encode predicted EppA processed pilins, as well as PibD/FlaK processed pilins, the prepilin peptidases are similarly stringent. Expressing a single prepilin peptidase that processes both adhesion pilins and motility-promoting flagellins, as opposed to distinct prepilin peptidases for each, requires that distinct mechanisms regulate flagellin and pilin function. Details of some of these mechanisms have recently begun to emerge and are discussed below.

N-Glycosylation

In bacteria, type IV pilins are often O-glycosylated and the components of the pathways involved in this protein modification

have been well-characterized (Nothaft and Szymanski, 2010). This post-translational modification has been implicated in a variety of functions, including colony morphology, surface motility, and regulating pilus composition, with the specific roles played differing from species to species in bacteria (Smedley et al., 2005; Takahashi et al., 2012; Vik et al., 2012). O-glycosylation also affects how the immune systems of higher eukaryotes, such as humans, respond to invasive pathogenic species (Lizcano et al., 2012).

While O-glycosylation has not yet been reported for any archaeal pilins, some subsets of archaeal type IV pilins are predicted to be N-glycosylated (**Table 1**; Jarrell et al., 2014). A detailed description of the archaeal glycosylation (Agl) pathways is presented by Eichler in this special issue (Kandiba and Eichler, 2014). Briefly, sugars are initially assembled on a dolichol phosphate lipid carrier, are then “flipped” across the cytoplasmic membrane, and are finally transferred to the target protein by an oligosaccharyltransferase at a conserved Asn-X-Ser/Thr motif (Jarrell et al., 2014). Although the archaeal glycosylation pathways involved in glycosylation of pilin-like proteins contain a conserved AglB oligosaccharyltransferase (Abu-Qarn and Eichler, 2006; Chaban et al., 2006; VanDyke et al., 2009), the composition of the polysaccharide moiety added to the modified protein varies between species and even between the subunits of different surface filaments. For example, *M. maripaludis* flagellins are decorated with a tetrasaccharide that is similar to the pentasaccharide found on the type IV pilin EpdE, which contains an additional hexose attached as a branch to the linking GalNAc subunit (Ng et al., 2010). Interestingly, while *M. maripaludis* flagella processing by FlaK is not required for N-glycosylation it was recently shown that EppA-dependent signal peptide processing of EpdD and EpdE is required for the N-glycosylation of these pilins (Nair and Jarrell, 2015). Although glycosylation is required for the synthesis of the *M. maripaludis* flagella, its type IV pili appear to be stable in a Δ $aglB$ strain (VanDyke et al., 2009). However, it remains to be determined whether the pili in this strain are functional. Thus far, with regards to glycosylation, only the deletion of the *M. maripaludis* acetyltransferase, which is required for the synthesis of the second sugar of the polysaccharide, affects proper cell surface attachment, as only few cell-associated pili could be identified in this strain, while culture supernatants contained pili (VanDyke et al., 2008).

The *H. volcanii* flagellins, FlgA1 and FlgA2, are also decorated with a pentasaccharide in an AglB-dependent manner; however, in this case, the pentasaccharide contains a hexose, mannose, two hexuronic acids, and a methyl ester of a hexuronic acid (Tripepi et al., 2012), the same that was previously identified on the *H. volcanii* s-layer (Abu-Qarn et al., 2007; **Table 1**). While the composition of polysaccharides attached to *H. volcanii* type IV pilins has not yet been determined, PilA1–PilA4 are also predicted to be N-glycosylated (**Table 1**). Interestingly, recent studies identified an alternative Agl pathway that differentially glycosylates the *H. volcanii* S-layer under low salt conditions (Guan et al., 2011). Furthermore, AglB-dependent glycosylation is diminished under these conditions (Kaminski et al., 2013). Hence, the *H.*

TABLE 1 |*In vivo* characterized archaeal type IV pilins.

Organism (# FlaFind pos.)	Pilus	Pilin	PilB/PilC	Prepilin peptidase		N-glycosylation: Sequons/predicted ⁴	Assembly	Role in biofilm formation	
				Paralog	Processing site ³			Adhesion	Microcolony formation
<i>M. maripaludis</i> (16)	Epd	EpdA	EpdL/EpdU & K	EppA	R AQI	4/3	+/-	ND	ND
		EpdB		EppA	K GQV	7/4	+	ND	ND
		EpdC		EppA	K GQV	4/3	+	ND	ND
		EpdD		EppA	K GQL	1/0	+	ND	ND
		EpdE ¹		EppA	K GQI	4/3	+	ND	ND
<i>S. acidocaldarius</i> (20)	Ups	UpsA ²	UpsE/UpsF	PibD	K AlS	4/3	+	+/-	+ ⁵
		UpsB ²		PibD	K GlS	4/4	+	+/-	+ ⁵
		AapA	AapE/AapF	PibD	R AlS	2/2	+	+	ND
		AapB ¹		PibD	R AlS	4/3	+	+	ND
<i>H. volcanii</i> (47)	PilA	PilA1 ²	PilB3/PilC3	PibD	D AVS	3/2	+	-	-
		PilA2 ²		PibD	D SAV	3/3	+	+	+
		PilA3 ²		PibD	D AVS	3/2	+	+	+
		PilA4 ²		PibD	R AVS	4/3	+	-	+
		PilA5 ²		PibD	R AVS	0/0	+	+	+
		PilA6 ²		PibD	R AVS	1/1	+	+	+

¹Determined as major pilins by mass spectrometry.²Promote pilus biosynthesis in absence of other known pilins.³Processing site is C-terminal to the bolded amino acid; predicted by FlaFind (Szabó et al., 2007).⁴Number of Asn-Xaa-Ser/Thr sequons/# Predicted by Net-N-glyc 1.0.⁵Aggregates formed in liquid media. See text for details.

volcanii pili might also display differential glycosylation under low salt conditions, although whether the glycosylation of these pili involves this newly identified pathway has yet to be elucidated. Preliminary data support *in silico* predictions suggesting that PilA1–PilA4 are *N*-glycosylated in an AglB-dependent manner and also show that these pilins probably inhibit microcolony formation that is promoted by PilA5 and PilA6 (Esquivel et al., 2013). Differential glycosylation of these adhesion pilins may be a regulatory mechanism that results in the inhibition of microcolony formation under stress conditions, such as low salt (see below). Since loss of AglB-dependent glycosylation of the flagellins also inhibits *H. volcanii* flagella biosynthesis (Triepel et al., 2012), low salt conditions might promote biofilm formation by inhibiting flagella-dependent motility as well as alleviating the inhibition of microcolony formation that is dependent upon the glycosylation of PilA1–PilA4.

Thus far, only a small subset of pilins have been investigated for post-translational modifications; future molecular and cellular biological analyses combined with improving mass spectrometry methods, will undoubtedly reveal additional modifications and the roles they play in the biosynthesis, function, and regulation of archaeal type IV pili. Substantial progress has already been made in *Sulfolobales* in regards to flagellin glycosylation (Meyer and Albers, 2013), and it will be interesting to determine whether the crenarchaeal pilins are also glycosylated as predicted by *in silico* data (**Table 1**). A more thorough characterization of these modifications will lead to a better understanding of the mechanisms that regulate the biosynthesis and functions of these evolutionarily ancient surface filaments and may result in insights into the regulation of biofilm formation and dispersal under extreme conditions.

Major and Minor Pilins

In any given archaeon the number of genes that encode predicted PibD-processed substrates can vary greatly (**Table 1**). While many of these genes likely encode the subunits of specific pilus-like structures, depending on growth conditions, an archaeon might produce differing sets of type IV pili. As in bacteria (Kuchma et al., 2012), archaeal pili can be composed of major and minor pilins. For example, while EpdE, as determined by Mass spectrometry, is the major pilin of *M. maripaludis*, two additional pilins, EpdB, and EpdC, are also required for piliation, and cells lacking EpdA have reduced piliation under the conditions tested. The exact functions of these minor pilins, whose genes are co-regulated with pilus biosynthesis genes, are largely unknown (**Figure 1**; Ng et al., 2010). Six additional genes encoding putative minor pilins *mmp0528*, *mmp0600*, *mmp0601*, *mmp0709*, *mmp0903*, and *mmp1283* (*epdD*) were recently investigated to determine whether the proteins they encode are involved in pilus assembly. Investigations of specific deletion mutants for each gene determined that only the deletion of *epdD* results in the loss of piliation (Nair et al., 2014). While pili containing EpdE as the major pilin appear to be the only type IV pili that are synthesized in this methanogen under the standard laboratory conditions tested, it is unclear whether the other predicted pilins might be involved in related functions, such as regulating pili-assembly or perhaps in forming

distinct pili under different conditions or when attached to different surfaces rather than expressed in planktonic cells (see below).

The differential expression of pilins has indeed been demonstrated in *S. solfataricus*, which encodes 28 FlaFind positives (Szabó et al., 2007). However, the only known type IV pilus produced by *S. solfataricus* is induced through exposure to ultraviolet (UV) light. This UV-inducible pilus, known as Ups, which plays important roles in both aggregation and surface attachment, is believed to be composed of UpsA and UpsB, two pilins encoded by genes adjacent to the pilus-biosynthesis genes *upsE* and *upsF* (**Figure 1** and see below). When overexpressing these pilins, *S. solfataricus* produces long, irregular pili (Fröls et al., 2008). Consistent with each of them being major pilins, individual deletions of these homologous genes in *S. acidocaldarius* still results in piliation, although fewer pili are observed (van Wolferen et al., 2013). Whether these pilins form mixed pili in wild-type cells is not known. Perhaps additional pilins, encoded by genes that are not co-regulated with the biosynthesis genes, are required for pilus formation, as demonstrated in *M. maripaludis*. Of the additional 24 FlaFind positive genes in *S. solfataricus*, one encodes the flagellin and three encode known substrate-binding proteins, the functions of the remaining 20 predicted PibD substrates remain elusive.

In addition to the Ups pili, *S. acidocaldarius* also produces the archaeal adhesive or Aap pilus, which plays a major role in surface adhesion (Henché et al., 2012a). The function and assembly of the Aap pilus requires the presence of at least two pilins, AapA and AapB, both of which are encoded by genes located adjacent to Aap pilus biosynthesis genes, *aapE* and *aapF* (**Figure 1** and see below). Unlike the Ups pili, the deletion of either gene encoding these Aap pilin subunits results in the absence of Aap pili. As only AapB was identified by mass spectrometry of purified Aap filaments, AapB appears to be the major pilin under the tested conditions (Henché et al., 2012a). Interestingly, while *aapB* expression is downregulated during stationary phase, *aapA* expression is increased, suggesting that the composition of the pili produced by cells varies depending upon growth conditions. As in *S. solfataricus*, aside from the pilins *aapA* and *aapB*, the eight substrate-binding proteins and one flagellin, the functional roles of most of its 20 FlaFind positives are not known.

Interestingly, in *H. volcanii*, the deletion of six adhesion pilin genes *pilA1-pilA6*, none of which is co-regulated with pilus-biosynthesis genes, is required to inhibit pilus-formation (**Table 1** and **Figures 1 and 2**; Esquivel et al., 2013). Among the 47 predicted PibD substrates of *H. volcanii*, only these six pilins contain an almost completely conserved H-domain, PilA2 containing one additional N-terminal serine. Each of these six pilins can form a functional type IV pilus when expressed individually in a Δ *pilA[1-6]* strain, suggesting that each can serve as the major pilin. However, it should be noted that, while pili in a wild-type *H. volcanii* strain are 8–12 nm in width and can be up to 4 μ m long, individually expressed PilA1–PilA6 in Δ *pilA[1-6]* make very short pili (Esquivel et al., 2013). Preliminary data strongly indicate that only a subset of these six pilins is expressed during planktonic growth while another set is expressed under

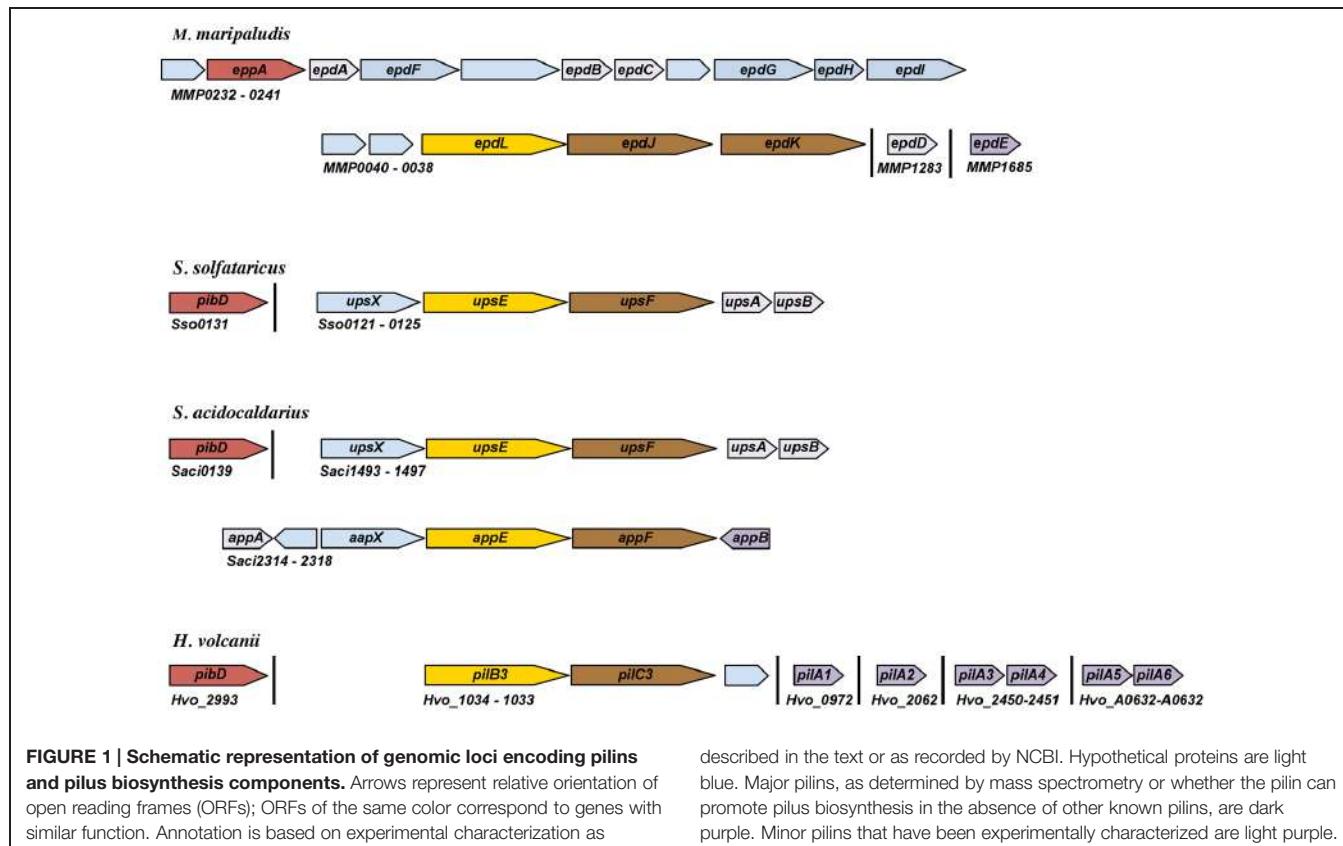


FIGURE 1 | Schematic representation of genomic loci encoding pilins and pilus biosynthesis components. Arrows represent relative orientation of open reading frames (ORFs); ORFs of the same color correspond to genes with similar function. Annotation is based on experimental characterization as

described in the text or as recorded by NCBI. Hypothetical proteins are light blue. Major pilins, as determined by mass spectrometry or whether the pilin can promote pilus biosynthesis in the absence of other known pilins, are dark purple. Minor pilins that have been experimentally characterized are light purple.

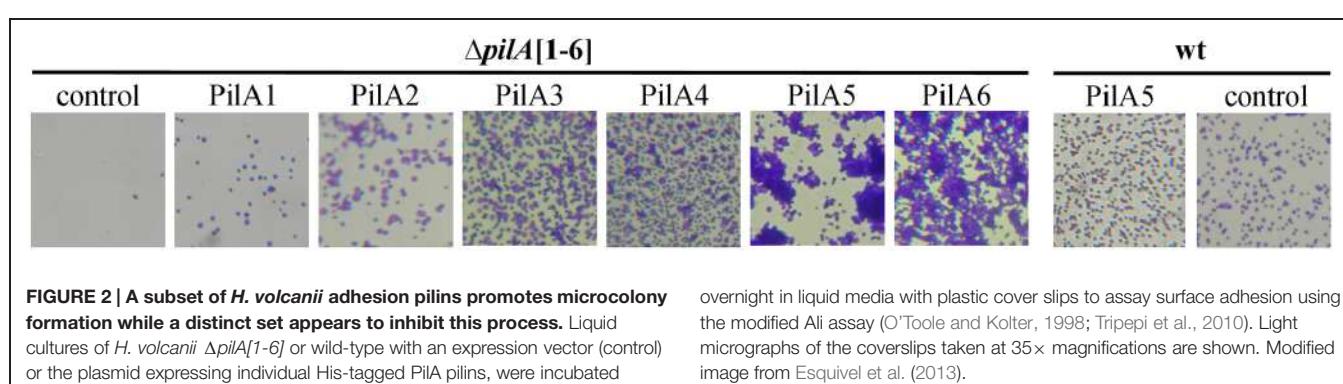


FIGURE 2 | A subset of *H. volcanii* adhesion pilins promotes microcolony formation while a distinct set appears to inhibit this process. Liquid cultures of *H. volcanii* $\Delta pilA[1-6]$ or wild-type with an expression vector (control) or the plasmid expressing individual His-tagged PilA pilins, were incubated

overnight in liquid media with plastic cover slips to assay surface adhesion using the modified All assay (O'Toole and Kolter, 1998; Tripepi et al., 2010). Light micrographs of the coverslips taken at 35 \times magnifications are shown. Modified image from Esquivel et al. (2013).

sessile conditions. The critical importance of the pilin H-domain for pilus-biosynthesis was demonstrated by the fact that a fusion protein PilA1Hybrid, in which the conserved pilin H-domain is replaced with an unrelated H-domain, cannot form pili in the $\Delta pilA[1-6]$ strain (Esquivel et al., 2013; see below).

Pilus-Assembly

Throughout prokaryotes, assembly of type IV pili occurs using energy obtained by the evolutionarily conserved ATPase, PilB. While it is known that the hydrolysis of ATP by this VirB11 ATPase provides the energy, it is not known how this energy is transferred to the pilin and how the pilin is moved from the membrane into the pilus. While the transmembrane protein, PilC, is proposed to be the membrane anchor for the pilus, to date there is

no definitive evidence for this function. However, genes encoding either of these evolutionarily conserved biosynthesis components are essential for piliation (Table 1; Pelicic, 2008).

M. maripaludis type IV pili, which have a diameter of ~6 nm, have a hollow lumen unlike other type IV pili with available structures. Similar to bacterial pili, they require a *pilB* homolog (*epdL*) for assembly. Moreover, *pilB* is co-regulated with two *pilC* paralogs (*epdJ* and *epdK*), which are both essential for piliation (Hendrickson et al., 2004; Nair et al., 2014). The requirement for two PilC paralogs is reminiscent of the requirement for two co-regulated *pilC* paralogs, *tadB* and *tadC* in *Aggregatibacter actinomycetemcomitans* (Kachlany et al., 2000). It has been proposed that the single ATPase of the Tad system may interact with one version of the conserved membrane proteins in pilin

addition and with the other in pilin removal (Burrows, 2012b). However, whether archaeal type IV pili can retract has not yet been determined (see below). It should be noted that, while an archaeal retraction ATPase, PilT, has not been identified, *M. maripaludis*, in addition to EpdL and FlaI, does encode one additional VirB11-like ATPase, which, like PilT, is not encoded by a gene co-localized with a *pilC* (Nair et al., 2014). While this protein is not required for pilus-biosynthesis in *M. maripaludis*, it remains to be determined whether it is involved in retraction.

Unlike the *M. maripaludis* genome, which appears to contain only a single *fla* and a single *pil* biosynthesis operon, many archaea, despite encoding only one, or, at most, two PibD paralogs, contain several operons that encode PilB and PilC homologs. The *S. solfataricus* genome, for example, contains the *bas* operon, which encodes among other proteins, the PilB and PilC homologs that are required for the incorporation of substrate-binding proteins into a bindosome, as well as the previously noted *ups* operon, which encodes UpsE and UpsF, PilB and PilC homologs, respectively (She et al., 2001; Zolghadr et al., 2007; van Wolferen et al., 2013). While wild-type cells, upon UV-induction form pili of up to 16 μm in length that have a diameter of approximately 10 nm, the deletion of *upsE* results in a non-piliated strain (Fröls et al., 2008). At least a subset of the *ups* pilins are also encoded by this operon, as noted above. *S. acidocaldarius* shares the *fla*, *bas* and *ups* operons with *S. solfataricus*, but also has an additional *pil* operon that encodes the PilB and PilC paralogs, AapE and AapF, respectively (Chen et al., 2005). Cells lacking either AapE or AapF do not form Aap pili, long filaments, 8–10 nm in diameter, that unlike the *M. maripaludis* pili, are not hollow (Henche et al., 2012a). Interestingly, a strain expressing AapE, but lacking UpsE, fails to assemble Ups pili upon UV-induction, suggesting that neither PilB paralog can complement the function of the other (Henche et al., 2012b). Similarly, in *H. volcanii*, which contains five putative *pil* operons, a *pilB3-C3* deletion strain, which lacks the PilB and PilC homologs required for assembling PilA1–PilA6 pili, does not have pili on the cell surface (Esquivel and Pohlschroder, 2014). While microarray data have indicated that three of the five *pilBpilC*-containing operons are not transcribed at detectable levels under experimental conditions tested to date, *pilB4* and *pilC4* are expressed under these conditions. The lack of PilA pilus formation in the Δ *pilB3-C3* background suggests that PilB4 and PilC4 are specifically involved in the biosynthesis of a pilus-like structure composed of two relatively large pilins encoded by FlaFind positive genes that are co-regulated with these pilus-biosynthesis components (Szabó et al., 2007).

Similar to *M. maripaludis*, the genomes of both *Sulfolobales* strains examined, as well as *H. volcanii*, encode a PilB homolog that is not co-regulated with a PilC homolog (Hartman et al., 2010). Perhaps these “orphan” PilB homologs are involved in pilus retraction, similar to PilT in bacteria. Although, as noted above, twitching motility has not yet been observed in any archaea, a recent report has shown evidence for *H. volcanii* social motility, where waves of this haloarchaeon were observed moving through a liquid medium (Chimileski et al., 2014). While this differs from surface motility, type IV pili may be required for

this process, as is true for the bacteria *Stigmatella aurantiaca* and *Myxococcus xanthus* (Tan et al., 2013).

Similarly, following initial *H. volcanii* attachment where an even distribution of cells is seen across a surface, the flagella-independent transition into microcolonies may also require type IV pili retraction (Esquivel et al., 2013). Determination of the effect that deleting the *H. volcanii* “orphan” *pilB* homolog has on these functions may lead to intriguing new lines of inquiry into the roles played by type IV pili in a variety of archaeal cellular processes.

A number of well-studied gram-negative pilus-biosynthesis components, such as PilQ, an outer membrane pore known as a secretin, as well as additional components required for the biosynthesis of this pore, are not found in monoderm archaea (Ayers et al., 2010). Moreover, as noted, several types of post-translational modifications of type IV pilins, such as methylation and O-glycosylation, have been identified in bacteria (Nothaft and Szymanski, 2010; Kim et al., 2011). Unlike PilD, as noted above, the archaeal pre pilin peptidase, PibD, is not a methyltransferase (Albers et al., 2003) and neither methylated pilins, nor a methyltransferase that specifically methylates archaeal pilins, have thus far been identified. However, the analysis of archaeal type IV pilin post-translational modifications has been limited and *in vivo* studies, rather than *in silico* analyses, are more likely to identify any novel archaeal biosynthesis components. Two examples are the identification of the AglB-dependent type IV pilin N-glycosylation, which appears to be limited to the archaea (Jarrell et al., 2014) as well as AapX, a protein required for biosynthesis of the *S. acidocaldarius* Aap pili, but whose exact role is still elusive (Henche et al., 2012a).

Roles of Type IV Pili and its Pilins in Biofilm Formation

Type IV pili play important roles in several processes required for surface-associated biofilm formation, including surface attachment and microcolony formation, as well as the aggregation of cells in liquid media (Giltner et al., 2012; Lassak et al., 2012; Esquivel and Pohlschroder, 2014). Minor pilins also play crucial roles in regulating the assembly of pili in bacteria (Cisneros et al., 2012; Nguyen et al., 2015). Archaea have been identified in biofilms established in a diverse variety of microbial ecosystems, from acidic hot spring mats to methane-rich marine sediments and hypersaline lakes (Whitaker et al., 2005; Coman et al., 2013; Aoki et al., 2014). Several model systems of archaea, including *M. maripaludis* and *H. volcanii*, as well as the *Sulfolobales* species, *S. acidocaldarius* and *S. solfataricus*, can also form biofilms (Fröls et al., 2012; Henche et al., 2012b; Fröls, 2013; Brileya et al., 2014). Recent molecular biological analyses of biofilm formation in these organisms have revealed that archaeal type IV pili, like their evolutionarily conserved bacterial counterparts, are involved in surface adhesion and cell aggregation (Jarrell et al., 2011; Henche et al., 2012a; Esquivel et al., 2013). Moreover, archaeal pilins also seem to be involved in regulating microcolony formation and flagella-dependent motility, functions that, while previously unrecognized, might be broadly

conserved across the prokaryotes (Esquivel and Pohlschroder, 2014).

Adhesion to Abiotic Surfaces

M. maripaludis can attach to a diverse set of materials, including glass, nickel, gold, silicon, and molybdenum, as has been determined by electron microscopy (**Table 1**; Jarrell et al., 2011). Attachment of cells in shaking cultures to all the abiotic surfaces tested thus far is EppA-dependent. However, it has not yet been determined which of the nine predicted methanogen pilins, or combination thereof, are required for attachment to these different surfaces. While cells grown in liquid media appear to have pili composed of the major pilin EpdE, under the conditions where cells adhere to various surfaces the composition of the pili may be distinct (Jarrell et al., 2011). While 14 *M. maripaludis* FlaFind positive pilins are predicted to be processed by EppA, only five of these have thus far been shown to be required for pilus assembly under the conditions tested (Ng et al., 2010; Nair et al., 2014).

Type IV pilus-dependent adhesion to different surfaces has also been demonstrated for the *Sulfolobales* model organisms using shaking cultures. While wild-type *S. solfataricus* can attach to mica, glass, pyrite, and carbon-coated gold grids (**Figure 3**), the *S. solfataricus* *upsE* deletion mutant will not adhere to any of these surfaces. Type IV pili are more highly expressed during growth on a surface as determined by electron microscopy of cells attached to the four surfaces (Zolghadr et al., 2009). Whether the known Ups pili subunits, UpsA and UpsB are crucial for adhesion to the various surfaces noted above is not yet known.

Interestingly, the patterns of *S. solfataricus* attachment to a surface are distinct, depending upon the specific material, perhaps indicating that the pili in each case have a distinct pilin composition.

Consistent with *S. acidocaldarius* containing a second adhesion pilus, deleting *upsE* does not lead to defective adhesion to glass in this closely related crenarchaeon, nor does deleting the gene encoding the *S. acidocaldarius* membrane protein AapF affect adhesion to this surface (Hence et al., 2012b). However, consistent with at least one of these pili being required for adhesion, a Δ *upsE* Δ *aapF* strain has a severe adhesion defect. Interestingly, a Δ *aapF* Δ *flaJ* strain also exhibits a defective adhesion phenotype, supporting previous studies that indicated a role for *Sulfolobales* flagella in adhesion (Zolghadr et al., 2009; Hence et al., 2012b). While flagella also appear to play a role in surface adhesion in *M. maripaludis* (Jarrell et al., 2011), the *H. volcanii* flagella do not. Adhesion was assessed by the accumulation of cells at the air-liquid interface of a glass coverslip incubated in static liquid culture (O'Toole and Kolter, 1998; Tripepi et al., 2010). Perhaps a flagella-driven force is not necessary to overcome the lower surface tensions that are found under high salt conditions compared to the surface tension that most prokaryotes must master in order to maintain contact with a surface during attachment (O'Toole et al., 2000). However, *H. volcanii* does require PilA1–PilA6 for attachment to a glass or plastic surface (Esquivel et al., 2013). Surprisingly, cells of the non-adhering Δ *pilA[1-6]* strain not only produce pili when expressing any of the individual pilins in *trans*, their ability

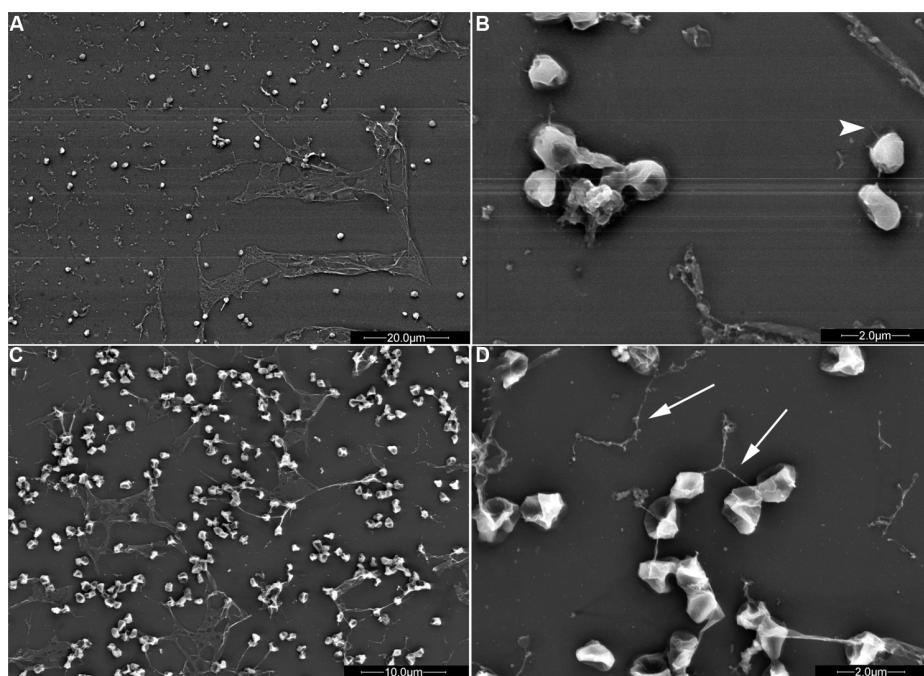


FIGURE 3 | *S. solfataricus* adhesion patterns differ depending on the surface. *S. solfataricus* cells attached to mica (**A,B**; **B** is an enlargement of **A**) produce significantly more sheath-like structures compared to cells attached to glass (**C,D**; **D** is an enlargement of **C**). Pili (arrowheads) and flagella (arrows) are

indicated. Bars: 20 μ m (**A**), 2 μ m (**B,D**), and 10 μ m (**C**; Zolghadr et al., 2009). Mica or glass surfaces were incubated in shaking cultures for 2 days in liquid medium, followed by electron microscopy to observe surface adhesion. Image courtesy of Sonja Albers, University of Freiburg.

to adhere to glass and plastic is also restored, consistent with the ability of each of these pilins to form a functional pilus. Interestingly, the ability of these cells to adhere to both surfaces varies depending on the pilin expressed. For instance, cells expressing either PilA1 or PilA2 have the lowest affinity for a glass or plastic surface under the conditions tested (**Figure 2**). PilA1 and PilA2 might be better adapted for binding to surfaces that are more frequently encountered in the natural environment by *H. volcanii*. Halophilic archaea have been isolated from brine shrimp, suggesting that chitin could be an ideal surface on which to test the adhesive capabilities of the six *H. volcanii* adhesion pilins (Riddle et al., 2013). Similar differential adhesive properties have been noted for specific pili in other species, including *V. cholerae*, which uses the MshA pilus to adhere to chitin and the TcpA pilus to attach to epithelial cells (Watnick et al., 1999; Krebs and Taylor, 2011).

Finally, Δ pilA[1-6] cells expressing a PilA1Hybrid in *trans*, in which the conserved H-domain of the pilin is replaced with an unrelated hydrophobic domain, are unable to adhere to glass, consistent with the inability of these cells to assemble pili (see above; Esquivel et al., 2013). Additionally, the Δ pilB3-C3 strain, which lacks the core biosynthesis components required for PilA pilus biosynthesis, retains some ability to adhere to glass, indicating that membrane-associated pilins can promote some surface adhesion (see above). Thus, the lack of adhesion by Δ pilA[1-6] cells expressing the PilA1Hybrid indicates that the H-domain has an additional role other than proper pilus assembly.

It will be intriguing to discover additional surfaces to which these varied archaea can adhere, and to determine whether the expression of as yet uncharacterized type IV pili or perhaps a distinct combination of pili or pilins or both is required to complete this crucial initial step in the formation of biofilms on these surfaces. Different environmental conditions may also induce the expression of a distinct set of pili. Although archaeal biofilms have been observed in natural environments and examined for composition (Wilmes et al., 2008), little work has been done to identify the archaeal type IV pili that are expressed in the cells that inhabit these natural environments or what additional functions these structures might perform in multispecies communities.

Cell Aggregation and Microcolony Formation

Biofilm formation is typically initiated by cells adhering to an abiotic surface followed by the formation of type IV pilus-dependent cell aggregates called microcolonies, which are encased in a polysaccharide matrix (Monds and O'Toole, 2009; Haussler and Fuqua, 2013; Orell et al., 2013a). While wild-type microcolony formation is not apparent even after 24 h, an *H. volcanii* Δ pilA[1-6] strain expressing either PilA5 or PilA6 in *trans* will form microcolonies after about 8 h (Esquivel et al., 2013). Since heterologous expression of either of these pilins in wild-type cells does not promote microcolony formation, at least a subset of the remaining adhesion pilins, PilA1–PilA4, must inhibit the formation of these cell aggregates. Not only is this the first indication that a specific subset of pilins can promote microcolony formation, it is also the first time that a distinct subset of pilins has been suggested to inhibit it. This regulatory

mechanism might prevent cell aggregation of planktonic cells that express a subset of pili, allowing cells to quickly attach to surfaces when necessary. Whether the cell aggregation that results in microcolony formation requires retractile type IV pili or is accomplished through another mechanism is not yet known.

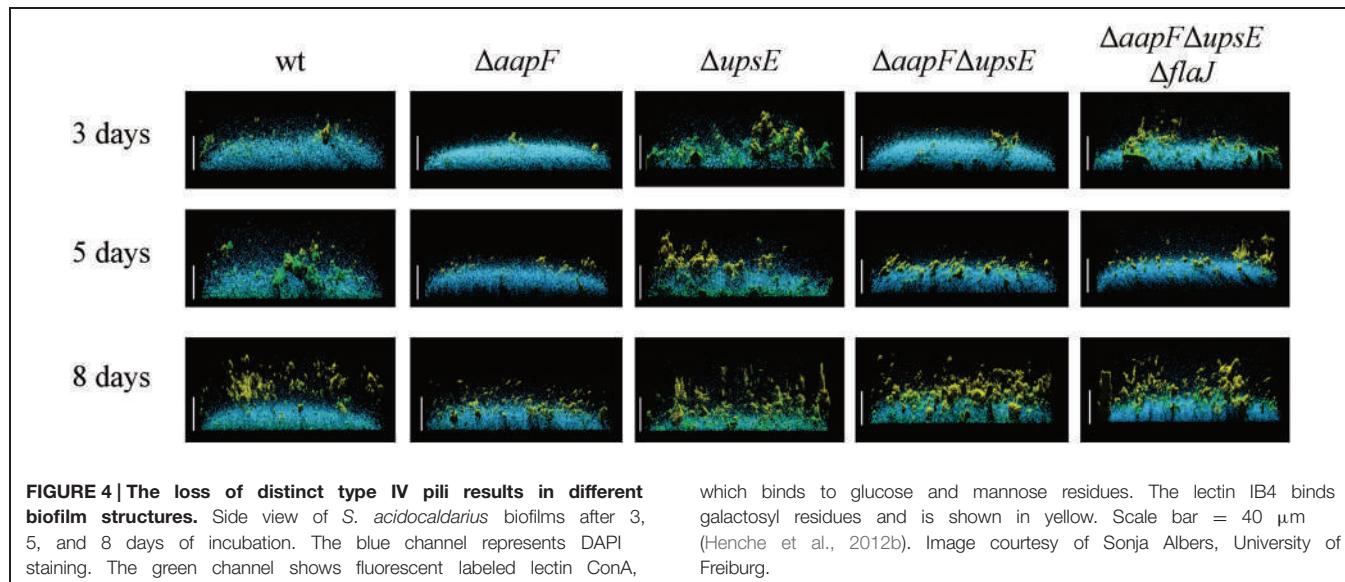
Pilus-dependent cell-to-cell interactions have been demonstrated for *S. solfataricus* and *S. acidocaldarius* in liquid media, where the Ups pili promote the formation of cell aggregates upon UV exposure (Fröls et al., 2008; Ajon et al., 2011). In this case, cellular aggregation, which is required for DNA exchange between the cells of these two species, is believed to be part of a response to DNA damage. In addition, these aggregates might be precursors of floating, pellicle biofilms that can protect cells from UV light as well as other stresses. While, as described above, unlike Δ upsE strains, which lack type IV pili, Δ upsA and Δ upsB strains, when analyzed by EM still had some pili; however, the UV-induced aggregation defect was similar in all strains, suggesting that the remaining pili could not significantly promote this cell–cell interaction (van Wolferen et al., 2013). It is not yet known whether the composition of Ups pili required for UV-induced aggregation in liquid media and those involved in adhesion to abiotic surfaces (see above) are identical.

H. volcanii can also exchange DNA through mating (Rosenshine et al., 1989) and biofilm formation promotes this DNA exchange (Chimileski et al., 2014). Interestingly, under non-biofilm conditions this process is not dependent on type IV pili, as deleting *pibD* does not affect mating despite abolishing pilus-assembly and surface adhesion (Tripepi et al., 2010).

Biofilm Maturation

The maturation of a surface biofilm, following surface adhesion and microcolony formation, results in large cell aggregates encased in exopolysaccharide (EPS) with differing morphologies between species (Fröls, 2013). The effects of type IV pili on the maturation of an archaeal biofilm have only recently been examined and are limited to studies in the *Sulfolobales*.

S. solfataricus liquid cultures grown in petri dishes form low-density carpets, covering the entire surface of the petri dish after 3 days of incubation. Thin cell-to-cell connections between aggregates within these biofilms have been observed by confocal liquid scanning microscopy. These connections can be stained by GSII, which binds N-acetylglucosamine residues, suggesting that these thin connections may be formed by glycosylated type IV pili (Koerdt et al., 2010). *S. acidocaldarius* wild-type cells, under similar growth conditions, establish thicker biofilms in which towering cell aggregates form, leaving some uncovered space between aggregates. An *S. solfataricus* *upsE* mutant strain, forms a significantly less dense biofilm, with an increase in the number of cell aggregates (Koerdt et al., 2010). Similar to *S. solfataricus*, a *upsE* deletion mutant in *S. acidocaldarius* forms a less dense biofilm that consists of loose aggregates, although these biofilms still maintain high, tower-like structures (**Figure 4**; Henche et al., 2012b). Staining of the biofilms with ConA and IB4 lectins, which bind to mannose/glucose and galactosyl sugar residues, respectively, has suggested that EPS production of these sugars is highly induced in the *upsE* deletion strain of *S. acidocaldarius*. Contrary to the phenotype



seen upon loss of the Ups pilus, the deletion of *aapF* in *S. acidocaldarius* results in a denser biofilm with a decreased thickness and lacking the towering structures (Hенче et al., 2012a,b). These results suggest that the Aap pili might be involved in maintaining a certain distance between the cell aggregates in the biofilm, perhaps to allow for optimal nutrient flow. Conversely, it is conceivable that the Aap pili facilitate surface (twitching) motility, which promotes microcolony formation following initially uniform adhesion to the surface. Hence, in the absence of these pili, cells are unable to form microcolonies and will rather expand evenly into a dense biofilm. Alternatively, the deletion of Ups pili may inhibit microcolony formation.

A subset of *H. volcanii* type IV pili promotes microcolony formation while another distinct set of pilins inhibits this type of cell aggregation, which is consistent with the observation that, while *H. volcanii* form microcolonies 2 days after inoculation into static liquid cultures (Chimileski et al., 2014), a Δ pilA[1-6] strain expressing either PilA5 or PilA6, begins forming microcolonies within 8 h of inoculation (Esquivel et al., 2013). As discussed above, it may be that biofilms mature more quickly when PilA1–PilA4 are not expressed. Cells expressing only PilA1 or PilA2 do not adhere as well-compared to the wild-type while cells expressing either PilA3 or PilA4 appear to adhere better, but they do not form microcolonies. *H. volcanii* biofilms observed after at least 7 days of incubation, form tall cell aggregate towers having a flaky appearance (Fröls et al., 2012; Chimileski et al., 2014). It will be intriguing to determine whether, similar to the *S. acidocaldarius* *aap* mutants, the mature pilA[5-6] mutant *H. volcanii* biofilm is denser. Much like the *Sulfolobales* biofilms, staining with ConA reveals EPS in this haloarchaeal biofilm. Interestingly, congo red also stains 3 day old biofilms, indicating the presence of amyloid protein. Since cells in mature biofilms behave differently than during biofilm formation, it is important to continue to examine the roles type IV pili play in maintaining, as well as forming, a biofilm.

Regulation of Flagella-Dependent Motility

To initiate biofilm formation, both bacteria and archaea must have regulatory mechanisms that facilitate the transition of cells from a planktonic to a sessile state when local conditions warrant it (McDougald et al., 2011). In the planktonic state, prokaryotic cells produce functional flagella that allow them to move through the environment, seeking nutrients and avoiding unfavorable conditions, while sessile cells use type IV pili to attach to abiotic surfaces and form microcolonies (O'Toole and Kolter, 1998; Ghosh and Albers, 2011). Thus, there exists an inverse relationship in these states between the expression of functional flagella and type IV pili and the respective sets of genes that encode components of the biosynthesis machinery for each (Fröls et al., 2008; Karatan and Watnick, 2009; Pohlschroder et al., 2011; Esquivel et al., 2013). Deleting either the gene encoding the major pilin *aapB* or pilin biosynthesis genes *aapE* and *aapF* results in hypermotility in the crenarchaeon *S. acidocaldarius* (Hенче et al., 2012a). Quantitative RT-PCR has revealed that deleting *aapF* leads to an increase in the expression of the flagellin gene, *flaB*, and the flagella gene *flaJ*, encoding a PilC homolog. This result is consistent with the hypermotility phenotype observed for the *aapF* deletion mutant (Hенче et al., 2012a,b). Deleting *upsE* also leads to an increase in the expression levels of the *fla* genes, albeit somewhat less significantly (Hенче et al., 2012b). The expression levels of these genes seem to be linked through a regulatory mechanism, which is not surprising given their roles at different stages during the formation, maintenance and dispersal of biofilms. However, in bacteria, where a similar inverse expression has also been observed, regulation of the changes in expression of the proteins involved in the assembly and function of the pili and flagella is often controlled by local concentrations of cyclic-di-GMP (Bordeleau et al., 2014; Martinez and Vadyvaloo, 2014). While c-di-GMP has not been shown to play a role in regulating archaeal pili and flagella expression, a recent study in *S. acidocaldarius* demonstrated that the deletion of

a gene encoding the Lrs14 transcription regulator, Saci0446, significantly upregulates *aapA* expression and increases biofilm formation while at the same time downregulating *flaB* and causing impaired motility (Orell et al., 2013b). Thus far, additional specific molecules that regulate biofilm formation in archaea are unknown.

Contrary to the models described above, deleting *H. volcanii pilB3* and *pilC3*, homologs of *aapE* and *aapF*, respectively, has no obvious effect on motility. More surprising, deleting all six *H. volcanii* genes that encode PilA pilins ($\Delta\text{pilA}[1-6]$) results in cells with a severe motility defect (Esquivel and Pohlschroder, 2014). Moreover, expressing any one of these pilins in *trans* in a $\Delta\text{pilA}[1-6]$ deletion mutant restores swimming motility (Table 1). However, when the PilA1Hybrid (See Adhesion section above) is expressed in the $\Delta\text{pilA}[1-6]$ strain, motility is not restored, indicating that these pilins, and more specifically, their conserved hydrophobic domain, play an important role in regulating flagella-dependent motility (Esquivel and Pohlschroder, 2014). Consistent with this hypothesis, in a $\Delta\text{pilA}[1-6]$ strain, heterologous expression of a FlgA1Hybrid in which the flagellin hydrophobic domain is replaced with the conserved PilA H-domain results in a restoration of motility, albeit to less than wild-type levels, despite the fact that this hybrid flagellin does not complement a motility defect caused by the deletion of *flgA1*. Furthermore, overexpressing any one of the *pilA* genes in a wild-type strain causes hypermotility (Esquivel and Pohlschroder,

2014). Considering that these pilins do not directly interact with either the flagellins or the flagellum, they appear to regulate motility within the membrane, perhaps by sequestering a protein that inhibits flagella biosynthesis (Figure 5). The regulation of flagella-dependent motility by proteins required for biofilm formation is reminiscent of the inhibition of swimming motility by the *B. subtilis* oligosaccharyltransferase, a bifunctional enzyme that is not only critical for EPS biosynthesis, it also acts like a clutch while interacting with the flagellum (Blair et al., 2008).

Finally, although there is not yet *in vivo* data supporting this hypothesis, as noted above, we know most archaea use a single prepilin peptidase, PibD to process both flagellins and pilins before they are incorporated into a filament (Albers et al., 2003; Szabó et al., 2007; Triepi et al., 2010). Thus, perhaps when prokaryotic cells attach to a surface, and pilin precursor expression increases, the PibD available for processing the flagellins might become more limited. Alternatively, in organisms such as *M. maripaludis* that have two PibD homologs, the availability of a second peptidase might allow the cells to shift more rapidly between high levels of flagella and high levels of type IV pili. While much work needs to be done to determine the details of the various regulatory mechanisms outlined here, the fact that the biosynthesis and functions of the flagella and type IV pili are regulated at several different levels underscores the importance of an ability to quickly transition from a planktonic to a sessile state and *vice versa*.

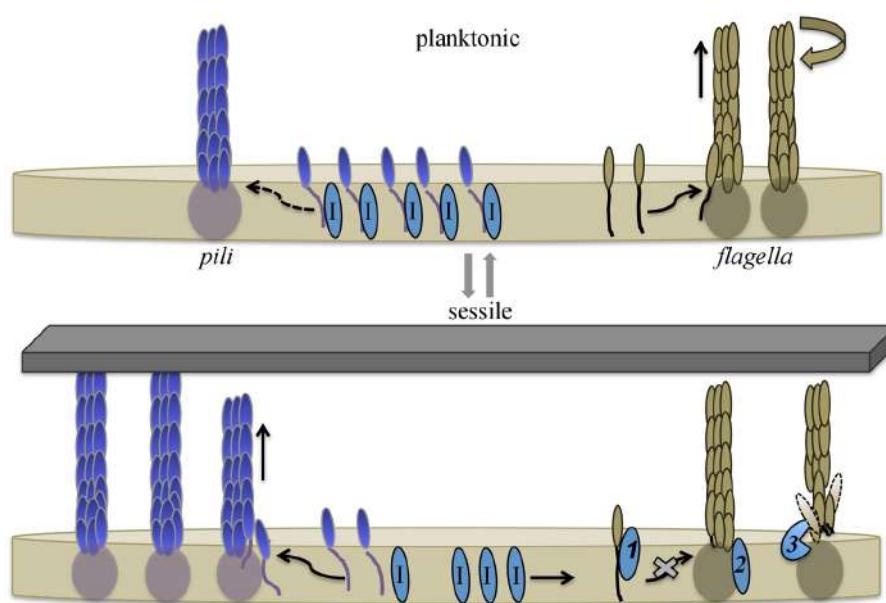


FIGURE 5 | Model for pilin-mediated inhibition of swimming motility. During planktonic growth *H. volcanii* cells synthesize flagellins that are readily incorporated into flagella, supporting swimming motility. Cells also express type IV pilins, which are incorporated into pili at a slow rate. The H-domain of membrane-associated pilins interacts with, and hence sequesters, a protein that directly or indirectly inhibits flagella motility. Upon adhesion, pilus-assembly kinetics shift, and the incorporation of pilins into pili increases, depleting the membrane of pilins and releasing the inhibitor proteins. The released inhibitors interfere with flagella

biosynthesis and/or stability. Taken together, this allows cells to rapidly respond to environmental conditions that favor biofilm formation over motility. Three possible mechanisms through which an inhibitor might hinder swimming motility are: (1) direct interaction with flagellins, preventing the incorporation of subunits into the flagellum; (2) inactivation of a flagella-biosynthesis component(s); or (3) degradation/destabilization of the flagella. Dashed arrow indicates limited incorporation into pili. Putative inhibitor proteins are labeled with I. Image modified from (Esquivel and Pohlschroder, 2014).

Concluding Remarks

During the past decade, biochemical and molecular biological studies combined with sophisticated microscopy, on a diverse set of archaeal models, have clearly demonstrated the critical importance of the evolutionarily conserved type IV pili in archaeal biofilm formation. These studies have also confirmed that the core components of the type IV pilus biosynthesis machinery are conserved across prokaryotic domains, and, conversely, revealed that, as compared to the bacterial filaments, there are novel aspects to the biosynthesis, regulation, and functions of archaeal type IV pili. A number of unique, previously unidentified, components of the archaeal type IV pilus biosynthesis machinery have been identified recently, and detailed analyses during the coming decade will be crucial in determining the roles these proteins play in pilus assembly. These analyses will include advanced approaches such as co-purifications coupled with sophisticated mass spectrometry as well as insertional mutagenesis as exemplified by the transposon screen that was recently developed for use in *H. volcanii* (Kiljunen et al., 2014). Such approaches should facilitate the identification of as yet unknown, but likely present, proteins that are critical for pilus assembly or function. Forthcoming studies of archaeal type IV pili will also focus on determining the significance of distinct sets of type IV pili that are commonly found within a single archaeon, including pili for which assembly depends on distinct sets of PilB and PilC paralogs. Among the prokaryotes that have been investigated, archaeal, as well as bacterial, species, can express up to six distinct PilB and PilC sets of paralogs. Furthermore, type IV pili composed of distinctly unrelated pilin subunits can still depend upon the same core components for assembly.

To date, the functions of the vast majority of predicted archaeal type IV pilins, including those encoded by the genomes of model archaeal systems, remain undetermined, perhaps because the conditions under which these pilins function have not been experimentally replicated. For instance, these predicted pilins may be the subunits of pili that are required for attachment to surfaces that have not yet been assayed. Crucial insights into the roles these pilins play in adhesion, microcolony formation, biofilm maturation, and dispersal can be gained by performing *in vivo* assays, along with RNAseq and mass spectrometry

on wild-type cells, as well as specific mutants, that are isolated during various stages of biofilm formation; currently, assays are performed predominantly on planktonic cells. The dynamics of pilus diversity on the cell surface, the dynamics of pilus subunit composition, as well as changes in the post-translational modifications of pilins during different stages of biofilm formation are not only poorly understood in archaea, but also in bacteria. Finally, while these studies are currently in their infancy, the tools needed to study the regulatory mechanisms that control transitions between planktonic to sessile cell states, where type IV pili appear to play key roles in archaea, are already available. Despite the discovery of previously unknown archaeal type IV pilus biosynthesis components, the molecular machinery involved in assembling type IV pilus-like structures in archaea appears to be significantly less complex than its counterpart in bacteria, simplifying detailed analysis of the molecular machinery involved in type IV pilus biosynthesis. Hence, future analyses of archaeal type IV pilus-biosynthesis may not only reveal further details about archaeal pilus-biosynthesis but about pilus biosynthesis in general. Comparisons between various archaeal systems, and between archaeal and bacterial systems, will help elucidate the types of adaptations that have allowed prokaryotes to thrive under a diverse variety of environmental conditions and may also provide useful information to aid in the development of a diverse set of industrial applications.

Decades of research in bacteria have provided a solid foundation for the study of archaeal type IV pili. Now, studies of the archaeal type IV pili will lead to important insights into the evolutionary history of these ancient cell surface structures and may lead to the identification of novel functions and regulatory mechanisms, which might also set in motion new lines of research on bacterial pili.

Acknowledgments

MP and RE were supported by National Aeronautics and Space Administration grant NNX10AR84G. RE was also supported by the National Institutes of Health Microbial Pathogenesis and Genomics Training grant 5T32AI060516. We thank the Pohlschroder lab for helpful discussions.

References

- Abu-Qarn, M., and Eichler, J. (2006). Protein N-glycosylation in Archaea: defining *Haloferax volcanii* genes involved in S-layer glycoprotein glycosylation. *Mol. Microbiol.* 61, 511–525. doi: 10.1111/j.1365-2958.2006.05252.x
- Abu-Qarn, M., Yurist-Doutsch, S., Giordano, A., Trauner, A., Morris, H. R., Hitchen, P., et al. (2007). *Haloferax volcanii* AglB and AglD are involved in N-glycosylation of the S-layer glycoprotein and proper assembly of the surface layer. *J. Mol. Biol.* 374, 1224–1236. doi: 10.1016/j.jmb.2007.10.042
- Ajon, M., Frols, S., van Wolferen, M., Stoecker, K., Teichmann, D., Driessen, A. J., et al. (2011). UV-inducible DNA exchange in hyperthermophilic archaea mediated by type IV pili. *Mol. Microbiol.* 82, 807–817. doi: 10.1111/j.1365-2958.2011.07861.x
- Albers, S. V., Szabo, Z., and Driessen, A. J. (2003). Archaeal homolog of bacterial type IV prepilin signal peptidases with broad substrate specificity. *J. Bacteriol.* 185, 3918–3925. doi: 10.1128/JB.185.13.3918-3925.2003
- Aoki, M., Ehara, M., Saito, Y., Yoshioka, H., Miyazaki, M., Miyashita, A., et al. (2014). A long-term cultivation of an anaerobic methane-oxidizing microbial community from deep-sea methane-seep sediment using a continuous-flow bioreactor. *PLoS ONE* 9:e105356. doi: 10.1371/journal.pone.0105356
- Arts, J., van Boxtel, R., Filloux, A., Tommassen, J., and Koster, M. (2007). Export of the pseudopilin XcpT of the *Pseudomonas aeruginosa* type II secretion system via the signal recognition particle-Sec pathway. *J. Bacteriol.* 189, 2069–2076. doi: 10.1128/JB.01236-06
- Ayers, M., Howell, P. L., and Burrows, L. L. (2010). Architecture of the type II secretion and type IV pilus machineries. *Future Microbiol.* 5, 1203–1218. doi: 10.2217/fmb.10.76
- Balaban, M., Battig, P., Muschiol, S., Tirier, S. M., Wartha, F., Normark, S., et al. (2014). Secretion of a pneumococcal type II secretion system pilus correlates with DNA uptake during transformation. *Proc. Natl. Acad. Sci. U.S.A.* 111, E758–E765. doi: 10.1073/pnas.1313860111

- Bardy, S. L., and Jarrell, K. F. (2003). Cleavage of preflagellins by an aspartic acid signal peptidase is essential for flagellation in the archaeon *Methanococcus voltae*. *Mol. Microbiol.* 50, 1339–1347. doi: 10.1046/j.1365-2958.2003.03758.x
- Blair, K. M., Turner, L., Winkelman, J. T., Berg, H. C., and Kearns, D. B. (2008). A molecular clutch disables flagella in the *Bacillus subtilis* biofilm. *Science* 320, 1636–1638. doi: 10.1126/science.1157877
- Blanco, L. P., Evans, M. L., Smith, D. R., Badtke, M. P., and Chapman, M. R. (2012). Diversity, biogenesis and function of microbial amyloids. *Trends Microbiol.* 20, 66–73. doi: 10.1016/j.tim.2011.11.005
- Bordeleau, E., Purcell, E. B., Lafontaine, D. A., Fortier, L. C., Tamayo, R., and Burrus, V. (2014). C-di-GMP riboswitch-regulated type IV pili contribute to aggregation of *Clostridium difficile*. *J. Bacteriol.* 197, 819–832. doi: 10.1128/jb.02340-14
- Bradley, D. E. (1972). Evidence for the retraction of *Pseudomonas aeruginosa* RNA phage pili. *Biochem. Biophys. Res. Commun.* 47, 142–149. doi: 10.1016/S0006-291X(72)80021-4
- Brileya, K. A., Camilleri, L. B., Zane, G. M., Wall, J. D., and Fields, M. W. (2014). Biofilm growth mode promotes maximum carrying capacity and community stability during product inhibition syntrophy. *Front. Microbiol.* 5:693. doi: 10.3389/fmicb.2014.00693
- Burrows, L. L. (2012a). Prime time for minor subunits of the type II secretion and type IV pilus systems. *Mol. Microbiol.* 86, 765–769. doi: 10.1111/mmi.12034
- Burrows, L. L. (2012b). *Pseudomonas aeruginosa* twitching motility: type IV pili in action. *Annu. Rev. Microbiol.* 66, 493–520. doi: 10.1146/annurev-micro-092611-150055
- Busch, A., and Waksman, G. (2012). Chaperone-usher pathways: diversity and pilus assembly mechanism. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 367, 1112–1122. doi: 10.1098/rstb.2011.0206
- Campos, M., Cisneros, D. A., Nivaskumar, M., and Francetic, O. (2013). The type II secretion system – a dynamic fiber assembly nanomachine. *Res. Microbiol.* 164, 545–555. doi: 10.1016/j.resmic.2013.03.013
- Chabani, B., Voisin, S., Kelly, J., Logan, S. M., and Jarrell, K. F. (2006). Identification of genes involved in the biosynthesis and attachment of *Methanococcus voltae* N-linked glycans: insight into N-linked glycosylation pathways in archaea. *Mol. Microbiol.* 61, 259–268. doi: 10.1111/j.1365-2958.2006.05226.x
- Chen, I., and Dubnau, D. (2004). DNA uptake during bacterial transformation. *Nat. Rev. Microbiol.* 2, 241–249. doi: 10.1038/nrmicro844
- Chen, I., Provvedi, R., and Dubnau, D. (2006). A macromolecular complex formed by a pilin-like protein in competent *Bacillus subtilis*. *J. Biol. Chem.* 281, 21720–21727. doi: 10.1074/jbc.M604071200
- Chen, L., Brugger, K., Skovgaard, M., Redder, P., She, Q., Torarinsson, E., et al. (2005). The genome of *Sulfobolbus acidocaldarius*, a model organism of the Crenarchaeota. *J. Bacteriol.* 187, 4992–4999. doi: 10.1128/jb.187.14.4992-4999.2005
- Chimileski, S., Franklin, M. J., and Papke, R. T. (2014). Biofilms formed by the archaeon *Haloferax volcanii* exhibit cellular differentiation and social motility, and facilitate horizontal gene transfer. *BMC Biol.* 12:65. doi: 10.1186/s12915-014-0065-5
- Cisneros, D. A., Pehau-Arnaudet, G., and Francetic, O. (2012). Heterologous assembly of type IV pili by a type II secretion system reveals the role of minor pilins in assembly initiation. *Mol. Microbiol.* 30, 598–603. doi: 10.1111/mmi.12033
- Coman, C., Druga, B., Hegedus, A., Sicora, C., and Dragos, N. (2013). Archaeal and bacterial diversity in two hot spring microbial mats from a geothermal region in Romania. *Extremophiles* 17, 523–534. doi: 10.1007/s00792-013-0537-5
- Davey, M. E., and O’Toole, G. A. (2000). Microbial biofilms: from ecology to molecular genetics. *Microbiol. Mol. Biol. Rev.* 64, 847–867. doi: 10.1128/MMBR.64.4.847-867.2000
- Diniz, M. C., Pacheco, A. C. L., Farias, K. M., and de Oliveira, D. M. (2012). The eukaryotic flagellum makes the day: novel and unforeseen roles uncovered after post-genomics and proteomics data. *Curr. Protein Pept. Sci.* 13, 524–546. doi: 10.2174/138920312803582951
- Esquivel, R. N., and Pohlschroder, M. (2014). A conserved type IV pilin signal peptide H-domain is critical for the post-translational regulation of flagella-dependent motility. *Mol. Microbiol.* 93, 494–504. doi: 10.1111/mmi.12673
- Esquivel, R. N., Xu, R., and Pohlschroder, M. (2013). Novel archaeal adhesion pilins with a conserved N terminus. *J. Bacteriol.* 195, 3808–3818. doi: 10.1128/JB.00572-13
- Francetic, O., Buddelmeijer, N., Lewenza, S., Kumamoto, C. A., and Pugsley, A. P. (2007). Signal recognition particle-dependent inner membrane targeting of the PilG Pseudopilin component of a type II secretion system. *J. Bacteriol.* 189, 1783–1793. doi: 10.1128/JB.01230-06
- Fröls, S. (2013). Archaeal biofilms: widespread and complex. *Biochem. Soc. Trans.* 41, 393–398. doi: 10.1042/bst20120304
- Fröls, S., Ajon, M., Wagner, M., Teichmann, D., Zolghadr, B., Folea, M., et al. (2008). UV-inducible cellular aggregation of the hyperthermophilic archaeon *Sulfobolbus solfataricus* is mediated by pili formation. *Mol. Microbiol.* 70, 938–952. doi: 10.1111/j.1365-2958.2008.06459.x
- Fröls, S., Dyall-Smith, M., and Pfeifer, F. (2012). Biofilm formation by haloarchaea. *Environ. Microbiol.* 14, 3159–3174. doi: 10.1111/j.1462-2920.2012.02895.x
- Ghosh, A., and Albers, S. V. (2011). Assembly and function of the archaeal flagellum. *Biochem. Soc. Trans.* 39, 64–69. doi: 10.1042/BST0390064
- Giltner, C. L., Nguyen, Y., and Burrows, L. L. (2012). Type IV pilin proteins: versatile molecular modules. *Microbiol. Mol. Biol. Rev.* 76, 740–772. doi: 10.1128/MMBR.00035-12
- Guan, Z., Naparstek, S., Calo, D., and Eichler, J. (2011). Protein glycosylation as an adaptive response in Archaea: growth at different salt concentrations leads to alterations in *Halofexax volcanii* S-layer glycoprotein N-glycosylation. *Environ. Microbiol.* 14, 743–753. doi: 10.1111/j.1462-2920.2011.02625.x
- Hartman, A. L., Norais, C., Badger, J. H., Delmas, S., Haldenby, S., Madupu, R., and et al. (2010). The complete genome sequence of *Halofexax volcanii* DS2, a model archaeon. *PLoS ONE* 5:e9605. doi: 10.1371/journal.pone.0009605
- Haussler, S., and Fuqua, C. (2013). Biofilms 2012: new discoveries and significant wrinkles in a dynamic field. *J. Bacteriol.* 195, 2947–2958. doi: 10.1128/JB.00239-13
- Henche, A. L., Ghosh, A., Yu, X., Jeske, T., Egelman, E., and Albers, S. V. (2012a). Structure and function of the adhesive type IV pilus of *Sulfobolbus acidocaldarius*. *Environ. Microbiol.* 14, 3188–3202. doi: 10.1111/j.1462-2920.2012.02898.x
- Henche, A. L., Koerdt, A., Ghosh, A., and Albers, S. V. (2012b). Influence of cell surface structures on crenarchaeal biofilm formation using a thermostable green fluorescent protein. *Environ. Microbiol.* 14, 779–793. doi: 10.1111/j.1462-2920.2011.02638.x
- Henche, A. L., van Wolferen, M., Ghosh, A., and Albers, S. V. (2014). Dissection of key determinants of cleavage activity in signal peptidase III (SPaseIII) PibD. *Extremophiles* 18, 905–913. doi: 10.1007/s00792-014-0675-4
- Hendrickson, E. L., Kaul, R., Zhou, Y., Bovee, D., Chapman, P., Chung, J., et al. (2004). Complete genome sequence of the genetically tractable hydrogenotrophic methanogen *Methanococcus maripaludis*. *J. Bacteriol.* 186, 6956–6969. doi: 10.1128/jb.186.20.6956-6969.2004
- Horn, C., Paulmann, B., Kerlen, G., Junker, N., and Huber, H. (1999). In vivo observation of cell division of anaerobic hyperthermophiles by using a high-intensity dark-field microscope. *J. Bacteriol.* 181, 5114–5118.
- Imam, S., Chen, Z., Roos, D. S., and Pohlschröder, M. (2011). Identification of surprisingly diverse type IV pili, across a broad range of gram-positive bacteria. *PLoS ONE* 6:e28919. doi: 10.1371/journal.pone.0028919
- Jarrell, K. F., and Albers, S. V. (2012). The archaellum: an old motility structure with a new name. *Trends Microbiol.* 20, 307–312. doi: 10.1016/j.tim.2012.04.007
- Jarrell, K. F., Ding, Y., Meyer, B. H., Albers, S. V., Kaminski, L., and Eichler, J. (2014). N-linked glycosylation in Archaea: a structural, functional, and genetic analysis. *Microbiol. Mol. Biol. Rev.* 78, 304–341. doi: 10.1128/mmbr.0052-13
- Jarrell, K. F., and McBride, M. J. (2008). The surprisingly diverse ways that prokaryotes move. *Nat. Rev. Microbiol.* 6, 466–476. doi: 10.1038/nrmicr.01900
- Jarrell, K. F., Stark, M., Nair, D. B., and Chong, J. P. (2011). Flagella and pili are both necessary for efficient attachment of *Methanococcus maripaludis* to surfaces. *FEMS Microbiol. Lett.* 319, 44–50. doi: 10.1111/j.1574-6968.2011.02264.x
- Kachlany, S. C., Planet, P. J., Bhattacharjee, M. K., Kollia, E., DeSalle, R., Fine, D. H., et al. (2000). Nonspecific adherence by *Actinobacillus actinomycetemcomitans*

- requires genes widespread in bacteria and archaea. *J. Bacteriol.* 182, 6169–6176. doi: 10.1128/JB.182.21.6169-6176.2000
- Kaminski, L., Guan, Z., Yurist-Doutsch, S., and Eichler, J. (2013). Two distinct N-glycosylation pathways process the *Haloflexax volcanii* S-Layer glycoprotein upon changes in environmental salinity. *mBio* 4, e00716-13. doi: 10.1128/mBio.00716-13
- Kandiba, L., and Eichler, J. (2014). Archaeal S-layer glycoproteins: post-translational modification in the face of extremes. *Front. Microbiol.* 5:661. doi: 10.3389/fmicb.2014.00661
- Karatan, E., and Watnick, P. (2009). Signals, regulatory networks, and materials that build and break bacterial biofilms. *Microbiol. Mol. Biol. Rev.* 73, 310–347. doi: 10.1128/mmbr.00041-08
- Kiljunen, S., Pajunen, M. I., Dilks, K., Storf, S., Pohlschroder, M., and Savilahti, H. (2014). Generation of comprehensive transposon insertion mutant library for the model archaeon, *Haloflexax volcanii*, and its use for gene discovery. *BMC Biol.* 12:103. doi: 10.1186/s12915-014-0103-3
- Kim, S., Rahman, M., Seol, S. Y., Yoon, S. S., and Kim, J. (2012). *Pseudomonas aeruginosa* bacteriophage PA1O requires type IV pili for infection and shows broad bactericidal and biofilm removal activities. *Appl. Environ. Microbiol.* 78, 6380–6385. doi: 10.1128/aem.00648-12
- Kim, Y. H., Park, K. H., Kim, S. Y., Ji, E. S., Kim, J. Y., Lee, S. K., et al. (2011). Identification of trimethylation at C-terminal lysine of pilin in the cyanobacterium *Synechocystis* PCC 6803. *Biochem. Biophys. Res. Commun.* 404, 587–592. doi: 10.1016/j.bbrc.2010.11.133
- Kirn, T. J., Bose, N., and Taylor, R. K. (2003). Secretion of a soluble colonization factor by the TCP type 4 pilus biogenesis pathway in *Vibrio cholerae*. *Mol. Microbiol.* 49, 81–92. doi: 10.1046/j.1365-2958.2003.03546.x
- Koerdt, A., Godeke, J., Berger, J., Thormann, K. M., and Albers, S. V. (2010). Crenarchaeal biofilm formation under extreme conditions. *PLoS ONE* 5:e14104. doi: 10.1371/journal.pone.0014104
- Krebs, S. J., and Taylor, R. K. (2011). Protection and attachment of *Vibrio cholerae* mediated by the toxin-coregulated pilus in the infant mouse model. *J. Bacteriol.* 193, 5260–5270. doi: 10.1128/JB.00378-11
- Kuchma, S. L., Griffin, E. F., and O'Toole, G. A. (2012). Minor pilins of the type IV pilus system participate in the negative regulation of swarming motility. *J. Bacteriol.* 194, 5388–5403. doi: 10.1128/JB.00899-12
- Lassak, K., Ghosh, A., and Albers, S. V. (2012). Diversity, assembly and regulation of archaeal type IV pilus-like and non-type-IV pilus-like surface structures. *Res. Microbiol.* 163, 630–644. doi: 10.1016/j.resmic.2012.10.024
- Laurenceau, R., Pehau-Arnaudet, G., Baconnais, S., Gault, J., Malosse, C., Dujeancourt, A., et al. (2013). A type IV pilus mediates DNA binding during natural transformation in *Streptococcus pneumoniae*. *PLoS Pathog.* 9:e1003473. doi: 10.1371/journal.ppat.1003473
- Leigh, J. A., Albers, S. V., Atomi, H., and Allers, T. (2011). Model organisms for genetics in the domain Archaea: methanogens, halophiles, Thermococcales and Sulfolobales. *FEMS Microbiol. Rev.* 35, 577–608. doi: 10.1111/j.1574-6976.2011.00265.x
- Lizcano, A., Sanchez, C. J., and Orihuela, C. J. (2012). A role for glycosylated serine-rich repeat proteins in gram-positive bacterial pathogenesis. *Mol. Oral Microbiol.* 27, 257–269. doi: 10.1111/j.2041-1014.2012.00653.x
- Lopez, D., Vlamakis, H., and Kolter, R. (2010). Biofilms. *Cold Spring Harb. Perspect. Biol.* 2:a000398. doi: 10.1101/cshperspect.a000398
- Lovley, D. R. (2012). Long-range electron transport to Fe(III) oxide via pili with metallic-like conductivity. *Biochem. Soc. Trans.* 40, 1186–1190. doi: 10.1042/bst20120131
- Martinez, L. C., and Vadyvaloo, V. (2014). Mechanisms of post-transcriptional gene regulation in bacterial biofilms. *Front. Cell. Infect. Microbiol.* 4:38. doi: 10.3389/fcimb.2014.00038
- McDougald, D., Rice, S. A., Barraud, N., Steinberg, P. D., and Kjelleberg, S. (2011). Should we stay or should we go: mechanisms and ecological consequences for biofilm dispersal. *Nat. Rev. Microbiol.* 10, 39–50. doi: 10.1038/nrmicro2695
- Meyer, B. H., and Albers, S. V. (2013). Hot and sweet: protein glycosylation in Crenarchaeota. *Biochem. Soc. Trans.* 41, 384–392. doi: 10.1042/bst20120296
- Monds, R. D., and O'Toole, G. A. (2009). The developmental model of microbial biofilms: ten years of a paradigm up for review. *Trends Microbiol.* 17, 73–87. doi: 10.1016/j.tim.2008.11.001
- Nair, D. B., and Jarrell, K. F. (2015). Pilin processing follows a different temporal route than that of Archaeolins in *Methanococcus maripaludis*. *Life* 5, 85–101. doi: 10.3390/life5010085
- Nair, D. B., Uchida, K., Aizawa, S., and Jarrell, K. F. (2014). Genetic analysis of a type IV pilus-like locus in the archaeon *Methanococcus maripaludis*. *Arch. Microbiol.* 196, 179–191. doi: 10.1007/s00203-014-0956-4
- Ng, S. Y., VanDyke, D. J., Chaban, B., Wu, J., Nosaka, Y., Aizawa, S., et al. (2009). Different minimal signal peptide lengths recognized by the archaeal prepilin-like peptidases FlaK and PibD. *J. Bacteriol.* 191, 6732–6740. doi: 10.1128/JB.00673-09
- Ng, S. Y., Wu, J., Nair, D. B., Logan, S. M., Robotham, A., Tessier, L., et al. (2010). Genetic and mass spectrometry analyses of the unusual type IV-like pili of the archaeon *Methanococcus maripaludis*. *J. Bacteriol.* 193, 804–814. doi: 10.1128/jb.00822-10
- Nguyen, Y., Sugiman-Marangos, S., Harvey, H., Bell, S. D., Charlton, C. L., Junop, M. S., et al. (2015). *Pseudomonas aeruginosa* minor pilins prime type IVa pilus assembly and promote surface display of the PilY1 adhesin. *J. Biol. Chem.* 290, 601–611. doi: 10.1074/jbc.M114.616904
- Nivaskumar, M., and Francetic, O. (2014). Type II secretion system: a magic beanstalk or a protein escalator. *Biochim. Biophys. Acta* 1843, 1568–1577. doi: 10.1016/j.bbamcr.2013.12.020
- Nothaft, H., and Szymanski, C. M. (2010). Protein glycosylation in bacteria: sweeter than ever. *Nat. Rev. Microbiol.* 8, 765–778. doi: 10.1038/nrmicro2383
- Nunn, D., Bergman, S., and Lory, S. (1990). Products of three accessory genes, pilB, pilC, and pilD, are required for biogenesis of *Pseudomonas aeruginosa* pili. *J. Bacteriol.* 172, 2911–2919.
- Orell, A., Frots, S., and Albers, S. V. (2013a). Archaeal biofilms: the great unexplored. *Annu. Rev. Microbiol.* 67, 337–354. doi: 10.1146/annurev-micro-092412-155616
- Orell, A., Peeters, E., Vassen, V., Jachlewski, S., Schalles, S., Siebers, B., et al. (2013b). Lrs14 transcriptional regulators influence biofilm formation and cell motility of Crenarchaea. *ISME J.* 7, 1886–1898. doi: 10.1038/ismej.2013.68
- O'Toole, G., Kaplan, H. B., and Kolter, R. (2000). Biofilm formation as microbial development. *Annu. Rev. Microbiol.* 51, 49–79. doi: 10.1146/annurev.micro.54.1.49
- O'Toole, G. A., and Kolter, R. (1998). Flagellar and twitching motility are necessary for *Pseudomonas aeruginosa* biofilm development. *Mol. Microbiol.* 30, 295–304. doi: 10.1046/j.1365-2958.1998.01062.x
- Pelicic, V. (2008). Type IV pili: e pluribus unum? *Mol. Microbiol.* 68, 827–837. doi: 10.1111/j.1365-2958.2008.06197.x
- Perras, A. K., Wanner, G., Klingl, A., Mora, M., Auerbach, A. K., Heinz, V., et al. (2014). Grappling archaea: ultrastructural analyses of an uncultivated, cold-loving archaeon, and its biofilm. *Front. Microbiol.* 5:397. doi: 10.3389/fmicb.2014.00397
- Pohlschroder, M., Ghosh, A., Tripepi, M., and Albers, S. V. (2011). Archaeal type IV pilus-like structures—evolutionarily conserved prokaryotic surface organelles. *Curr. Opin. Microbiol.* 14, 357–363. doi: 10.1016/j.mib.2011.03.002
- Reguera, G., McCarthy, K. D., Mehta, T., Nicoll, J. S., Tuominen, M. T., and Lovley, D. R. (2005). Extracellular electron transfer via microbial nanowires. *Nature* 435, 1098–1101. doi: 10.1038/nature03661
- Riddle, M. R., Baxter, B. K., and Avery, B. J. (2013). Molecular identification of microorganisms associated with the brine shrimp *Artemia franciscana*. *Aquat. Biosyst.* 9, 7. doi: 10.1186/2046-9063-9-7
- Rosenshine, I., Tchelet, R., and Mevarech, M. (1989). The mechanism of DNA transfer in the mating system of an archaeabacterium. *Science* 245, 1387–1389. doi: 10.1126/science.2818746
- Sauvonnet, N., Vignon, G., Pugsley, A. P., and Gounon, P. (2000). Pilus formation and protein secretion by the same machinery in *Escherichia coli*. *EMBO J.* 19, 2221–2228. doi: 10.1093/emboj/19.10.2221
- Schneewind, O., and Missiakas, D. (2013). Sec-secretion and sortase-mediated anchoring of proteins in Gram-positive bacteria. *Biochim. Biophys. Acta* 1843, 1687–1697. doi: 10.1016/j.bbamcr.2013.11.009
- She, Q., Singh, R. K., Confalonieri, F., Zivanovic, Y., Allard, G., Awayez, M. J., et al. (2001). The complete genome of the crenarchaeon *Sulfolobus solfataricus* P2. *Proc. Natl. Acad. Sci. U.S.A.* 98, 7835–7840. doi: 10.1073/pnas.141222098
- Smedley, J. G., III, Jewell, E., Roguskie, J., Horzempa, J., Syboldt, A., Stolz, D. B., et al. (2005). Influence of pilin glycosylation on *Pseudomonas aeruginosa*

- 1244 pilus function. *Infect. Immun.* 73, 7922–7931. doi: 10.1128/iai.73.12.7922-7931.2005
- Strom, M. S., Nunn, D. N., and Lory, S. (1993). A single bifunctional enzyme, PilD, catalyzes cleavage and N-methylation of proteins belonging to the type IV pilin family. *Proc. Natl. Acad. Sci. U.S.A.* 90, 2404–2408. doi: 10.1073/pnas.90.6.2404
- Szabo, Z., Albers, S. V., and Driessens, A. J. (2006). Active-site residues in the type IV prepilin peptidase homologue PilD from the archaeon *Sulfolobus solfataricus*. *J. Bacteriol.* 188, 1437–1443. doi: 10.1128/jb.188.4.1437-1443.2006
- Szabo, Z., and Pohlschroder, M. (2012). Diversity and subcellular distribution of archaeal secreted proteins. *Front. Microbiol.* 3:207. doi: 10.3389/fmicb.2012.00207
- Szabó, Z., Stahl, A. O., Albers, S. V., Kissinger, J. C., Driessens, A. J., and Pohlschröder, M. (2007). Identification of diverse archaeal proteins with class III signal peptides cleaved by distinct archaeal prepilin peptidases. *J. Bacteriol.* 189, 772–778. doi: 10.1128/JB.01547-06
- Takahashi, H., Yanagisawa, T., Kim, K. S., Yokoyama, S., and Ohnishi, M. (2012). Meningococcal PilV potentiates *Neisseria meningitidis* type IV pilus-mediated internalization into human endothelial and epithelial cells. *Infect. Immun.* 80, 4154–4166. doi: 10.1128/iai.00423-12
- Takhar, H. K., Kemp, K., Kim, M., Howell, P. L., and Burrows, L. L. (2013). The platform protein is essential for type IV pilus biogenesis. *J. Biol. Chem.* 288, 9721–9728. doi: 10.1074/jbc.M113.453506
- Tan, Z., Li, H., Pan, H., Zhou, X., Liu, X., Luo, N., et al. (2013). Characterization of four type IV pilin homologues in *Stigmatella aurantiaca* DSM17044 by heterologous expression in *Myxococcus xanthus*. *PLoS ONE* 8:e75105. doi: 10.1371/journal.pone.0075105
- Thoma, C., Frank, M., Rachel, R., Schmid, S., Nather, D., Wanner, G., et al. (2008). The Mth60 fimbriae of *Methanothermobacter thermoautotrophicus* are functional adhesins. *Environ. Microbiol.* 10, 2785–2795. doi: 10.1111/j.1462-2920.2008.01698.x
- Tripepi, M., Imam, S., and Pohlschroder, M. (2010). *Haloflexax volcanii* flagella are required for motility but are not involved in PilD-dependent surface adhesion. *J. Bacteriol.* 192, 3093–3102. doi: 10.1128/jb.00133-10
- Tripepi, M., You, J., Temel, S., Onder, O., Brisson, D., and Pohlschroder, M. (2012). N-glycosylation of *Haloflexax volcanii* flagellins requires known AgI proteins and is essential for biosynthesis of stable flagella. *J. Bacteriol.* 194, 4876–4887. doi: 10.1128/jb.00731-12
- VanDyke, D. J., Wu, J., Logan, S. M., Kelly, J. F., Mizuno, S., Aizawa, S., et al. (2009). Identification of genes involved in the assembly and attachment of a novel flagellin N-linked tetrasaccharide important for motility in the archaeon *Methanococcus maripaludis*. *Mol. Microbiol.* 72, 633–644. doi: 10.1111/j.1365-2958.2009.06671.x
- VanDyke, D. J., Wu, J., Ng, S. Y., Kanbe, M., Chaban, B., Aizawa, S., et al. (2008). Identification of a putative acetyltransferase gene, MMP0350, which affects proper assembly of both flagella and pili in the archaeon *Methanococcus maripaludis*. *J. Bacteriol.* 190, 5300–5307. doi: 10.1128/jb.00474-08
- van Wolferen, M., Ajon, M., Driessens, A. J., and Albers, S. V. (2013). Molecular analysis of the UV-inducible pilin operon from *Sulfolobus acidocaldarius*. *MicrobiologyOpen* 2, 928–937. doi: 10.1002/mbo3.128
- Vik, A., Aspholm, M., Anonsen, J. H., Borud, B., Roos, N., and Koomey, M. (2012). Insights into type IV pilus biogenesis and dynamics from genetic analysis of a C-terminally tagged pilin: a role for O-linked glycosylation. *Mol. Microbiol.* 85, 1166–1178. doi: 10.1111/j.1365-2958.2012.08166.x
- Watnick, P. I., Fullner, K. J., and Kolter, R. (1999). A role for the mannose-sensitive hemagglutinin in biofilm formation by *Vibrio cholerae* El Tor. *J. Bacteriol.* 181, 3606–3609.
- Whitaker, R. J., Grogan, D. W., and Taylor, J. W. (2005). Recombination shapes the natural population structure of the hyperthermophilic archaeon *Sulfolobus islandicus*. *Mol. Biol. Evol.* 22, 2354–2361. doi: 10.1093/molbev/msi233
- Wilmes, P., Remis, J. P., Hwang, M., Auer, M., Thelen, M. P., and Banfield, J. F. (2008). Natural acidophilic biofilm communities reflect distinct organismal and functional organization. *ISME J.* 3, 266–270. doi: 10.1038/ismej.2008.90
- Wozniak, D. J., and Parsek, M. R. (2014). Surface-associated microbes continue to surprise us in their sophisticated strategies for assembling biofilm communities. *F1000prime Rep.* 6, 26. doi: 10.12703/p6-26
- Zolghadr, B., Klingl, A., Koerdt, A., Driessens, A. J., Rachel, R., and Albers, S. V. (2009). Appendage-mediated surface adherence of *Sulfolobus solfataricus*. *J. Bacteriol.* 192, 104–110. doi: 10.1128/jb.01061-09
- Zolghadr, B., Klingl, A., Rachel, R., Driessens, A. J., and Albers, S. V. (2011). The bindosome is a structural component of the *Sulfolobus solfataricus* cell envelope. *Extremophiles* 15, 235–244. doi: 10.1007/s00792-010-0353-0
- Zolghadr, B., Weber, S., Szabo, Z., Driessens, A. J., and Albers, S. V. (2007). Identification of a system required for the functional surface localization of sugar binding proteins with class III signal peptides in *Sulfolobus solfataricus*. *Mol. Microbiol.* 64, 795–806. doi: 10.1111/j.1365-2958.2007.05697.x

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

Copyright © 2015 Pohlschroder and Esquivel. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Archaeal viruses at the cell envelope: entry and egress

Emmanuelle R. J. Quemin¹ and Tessa E. F. Quax^{2*}

¹ Department of Microbiology, Institut Pasteur, Paris, France, ² Molecular Biology of Archaea, Institute for Biology II - Microbiology, University of Freiburg, Freiburg, Germany

The cell envelope represents the main line of host defense that viruses encounter on their way from one cell to another. The cytoplasmic membrane in general is a physical barrier that needs to be crossed both upon viral entry and exit. Therefore, viruses from the three domains of life employ a wide range of strategies for perforation of the cell membrane, each adapted to the cell surface environment of their host. Here, we review recent insights on entry and egress mechanisms of viruses infecting archaea. Due to the unique nature of the archaeal cell envelope, these particular viruses exhibit novel and unexpected mechanisms to traverse the cellular membrane.

OPEN ACCESS

Edited by:

Mechthild Pohlschroder,
University of Pennsylvania, USA

Reviewed by:

Jason W. Cooley,
University of Missouri, USA

Jerry Eichler,

Ben Gurion University of the Negev,
Israel

*Correspondence:

Tessa E. F. Quax,

Molecular Biology of Archaea,
Institute for Biology II - Microbiology,

University of Freiburg,

Schänzlestrasse 1, 79104 Freiburg,
Germany

tessa.quax@biologie.uni-freiburg.de

Specialty section:

This article was submitted to
Microbial Physiology and Metabolism,
a section of the journal
Frontiers in Microbiology

Received: 07 January 2015

Accepted: 19 May 2015

Published: 05 June 2015

Citation:

Quemin ERJ and Quax TEF (2015)
Archaeal viruses at the cell envelope:
entry and egress.
Front. Microbiol. 6:552.
doi: 10.3389/fmicb.2015.00552

Introduction

Members of the three domains of life, Archaea, Bacteria and Eukarya, are all subject to viral infections. Viruses have been isolated from various environments, where they are often abundant, outnumbering prokaryotic cells by a factor of 10 (Bergh et al., 1989; Borsheim et al., 1990; Suttle, 2007). Viruses infecting archaea tend to display high morphological and genetic diversity compared to viruses of bacteria and eukaryotes (Pina et al., 2011). Several archaeal viral families have members, which display unique shapes that are not found amongst other viruses, such as a bottle, droplet or spiral (Prangishvili, 2013).

The cell envelope represents a major barrier for all viruses. In fact, the cell membrane has to be traversed twice by viruses to establish successful infection, first upon entry and secondly during exit. In order to cross the cell envelope, viruses have developed various strategies, each adapted to the membrane environment of their host.

The combination of high-throughput approaches with more classical techniques has shed light on the process of viral entry and release in some archaeal virus-host model systems. However, the detailed molecular mechanisms underlying the various stages of the viral life cycle remain poorly understood in archaea in general (Quemin et al., 2014). Recently, a few studies have focused on the adsorption at the surface of the archaeal host cell before viral entry and release of viral particles at the end of the infection cycle (Bize et al., 2009; Brumfield et al., 2009; Ceballos et al., 2012; Quemin et al., 2013; Deng et al., 2014). This has delivered the very first insights into the fashion in which viruses interact with the archaeal membrane.

The cell surface of archaea is fundamentally different from bacteria (Albers and Meyer, 2011). Archaeal membranes have an alternative lipid composition and generally lack a cell wall of peptidoglycan. In addition, the motility structures present at the surface of archaea are constructed from different building blocks than their bacterial counterparts (Pohlschroder et al., 2011). Gram positive bacteria contain a lipid bilayer covered by a thick peptidoglycan cell wall and gram negative cells are surrounded by two membranes with a thinner peptidoglycan in the periplasmic space in between. While bacteria typically contain a cell wall polymer of peptidoglycan (Tytas et al., 2012),

peptidoglycan cell walls are absent from archaea. Instead, most archaea are surrounded by a thin proteinaceous surface layer (S-layer) that consists of glycosylated proteins, which are anchored in the cell membrane. In contrast to the peptidoglycan, which has a molecular composition that can be very similar from one species to another, S-layer proteins show a great diversity (Fagan and Fairweather, 2014). Hence, archaea exhibit specific features, in particular at the cell surface, which are not shared with bacteria and influence the mechanisms at play in the course of infection.

The first studies on archaeal viral entry and egress have shown that some archaeal viruses employ entry strategies that superficially resemble those of bacterial viruses (Quemin et al., 2013; Deng et al., 2014), while others utilize surprisingly novel exit mechanisms (Brumfield et al., 2009; Quax et al., 2011). Here we will give an overview of the first studies reporting viral interaction with the archaeal cell envelope, focusing on hyperthermophilic crenarchaeal viruses. Furthermore, current research permits comparison with corresponding mechanisms taking place during the viral cycle of bacterial viruses. We will discuss how features of cell surfaces compel viruses to employ specific strategies for entry and egress.

Viral Entry

A virus is able to infect only a few strains or species. Such specificity in interaction of viruses with their host is determined by the characteristics of entry, which in turn rely on the nature and structural peculiarities of the cell envelope. Adsorption as the first key step of the viral cycle is one of the most restrictive in terms of host range, depending on the accessibility and number of receptors present at the cell surface (Poranen et al., 2002). Structural proteins are found within the viral particle in metastable conformation and it is the interaction with the host cell, which leads to a more stable, lower-energy conformation of these proteins (Dimitrov, 2004). Indeed, virus entry and genome uncoating are energy-dependent processes and irreversible conformational change of the capsid proteins (CP) during adsorption triggers the release of the genome from the extracellular virions (Molineux and Panja, 2013). As a general rule, entry can be subdivided in two steps. For the well-studied viruses infecting bacteria, the first contact with the host is reversible and then, viruses attach irreversibly to a specific, saturable cell envelope receptor. Primary and secondary adsorptions can take place with the same receptor or, more frequently involve different players. Common cellular determinants in bacteria are peptidoglycan, lipopolysaccharide S (LPS), or cellular appendages (Poranen et al., 2002). Subsequently, delivery of the viral genome into the cellular cytoplasm happens through the cell wall and bacterial membrane. Indeed, the nature of the host cell wall has a great influence on the viral entry mechanism and different cell types expose diverse external envelope structures. Three main entry strategies have been reported for viral entry in bacteria: genome release through an icosahedral vertex; dissociation of virion at the cell envelope; and virion penetration via membrane fusion (Poranen et al., 2002). Thus far insights into

the mechanisms of entry by archaeal viruses have been based on coincidental observations. However, more recently a few detailed analyses have provided a better understanding of the molecular mechanisms at play in archaeal virus-host systems from geothermal environments.

Interaction with Cellular Appendages

Filamentous, flexible viruses of the *Lipothrixviridae* family have been classified into four different genera partly based on the virion core and terminal structures. Indeed, the exposed filaments can vary in number from one (AFV9, *Acidianus* filamentous virus 9) to six (SIFV, *Sulfolobus islandicus* filamentous virus) or even form complex structures like claws (AFV1) or brushes (AFV2; Arnold et al., 2000; Bettstetter et al., 2003; Haring et al., 2005b; Bize et al., 2008). The high diversity of terminal structures observed in this particular family strongly suggests their involvement in cellular adsorption processes. Indeed, AFV1 particles terminate with claws that mediate attachment to cellular pili (Bettstetter et al., 2003). In the case of AFV2, the “bottle brush,” a complex collar termini with two sets of filaments, should be able to interact with the surface of host cells directly since its specific host doesn't show any extracellular appendages (Haring et al., 2005b). In addition, SIFV virions display mop-like structures found in open or closed conformations (Arnold et al., 2000). Hence, lipothrixviruses are decorated with diverse and unique terminal structures that play a major role in recognition and interaction with the host cell.

In a similar manner, the stiff, filamentous ruddivirus SIRV2 (*Sulfolobus islandicus* rod-shaped virus 2) was also shown to bind host pili by the three terminal fibers of virions. SIRV2 is one of the more appealing models to study virus-host interactions in archaea (Prangishvili et al., 2013). Recently published analyses concluded that adsorption occurs within the first minute of infection, much more efficiently than in halophilic archaeal systems for which binding requires several hours (Kukkaro and Bamford, 2009). The particles of SIRV2 specifically attach to the tip of host pili-like structures leading to a strong and irreversible interaction between the viral and cellular determinants (Figure 1A). Subsequently, viruses are found on the side of the appendages indicating a progression toward the cell surface where DNA entry is concomitant with virion disassembly (Quemin et al., 2013; Figures 1C,D). Thus, the three fibers located at the virion termini represent the viral anti-receptors involved in recognition of host cells and are responsible for the primary adsorption (Figure 1B). It is noteworthy that both ends of the virions have an equal binding capacity as previously noticed for the lipothrixvirus AFV1 (Bettstetter et al., 2003). The families *Lipothrixviridae* and *Rudiviridae* belong to the order *Ligamenvirales* and are known to attach to extracellular filaments (Prangishvili and Krupovic, 2012). Although AFV1 is capable of binding the side of host pili, a feature shared with bacterial leviviruses, cystoviruses and some tailed bacteriophages (Poranen et al., 2002), the interaction of SIRV2 with *Sulfolobus* filaments occurs initially via the tip. This resembles more closely the primary adsorption observed in the inoviruses (Rakonjac et al., 2011). All these data suggest that linear archaeal viruses employ a common strategy for the initiation of infection although

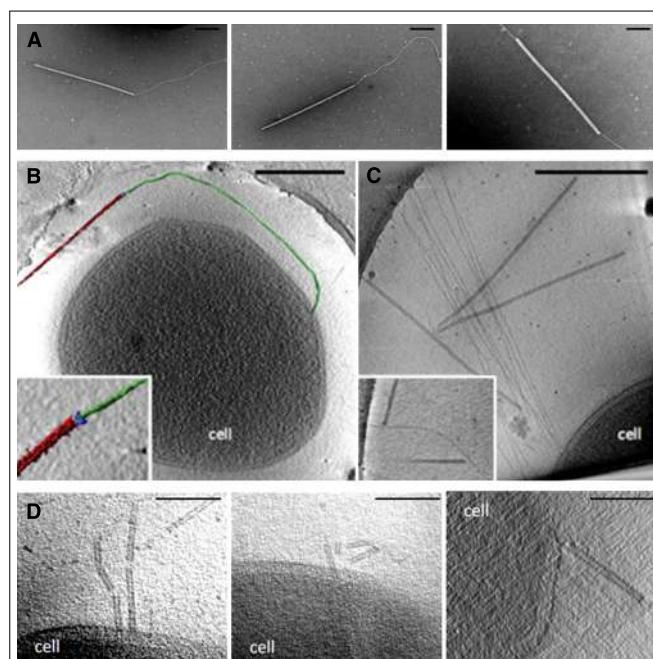


FIGURE 1 | Entry of SIRV2 in *S. islandicus* LAL14/1 cells.

(A) Transmission electron micrographs showing that SIRV2 virions interact with purified cellular filaments. Stained with 2% uranyl acetate for 2 min. Scale bar, 200 nm. Electron micrographs of SIRV2 interaction with *S. islandicus* LAL14/1 cells. Samples were collected 1 min post-infection and flash-frozen for electron cryotomography (cryo-ET). The virions interact both at the filament tips (B) and along the length of the filaments (C). The lower left panel (B) also shows a segmented tomographic volume of the SIRV2 virion (red) attached to the tip of an *S. islandicus* filament (green). The three terminal virion fibers that appear to mediate the interaction are shown in blue (the inset depicts a magnified view of the interaction between the virion fibers and the tip of the filament). The inset in the lower right panel (C) depicts two virions bound to the sides of a single filament. Scale bars, 500 nm. (D) Tomographic slices through *S. islandicus* LAL14/1 cells at 1 min after infection with SIRV2 reveals partially disassembled SIRV2 virions at the cell surface. Adapted from (Quemin et al., 2013). Scale bar, 100 nm.

the molecular mechanisms involved are most likely to be distinct.

Interaction with Cell Surface

As a general rule, viral entry implies direct or indirect binding to the cell surface depending on whether a primary adsorption step is required. In the case of SIRV2, analysis of virus-resistant strains provided interesting candidates for the receptors of SIRV2 virions at the cell surface. In fact, two operons were identified: sso2386-2387 and sso3139-3141 (Deng et al., 2014). The former encodes proteins homologous to components of type IV pili and the latter presumably a membrane-associated cell surface complex. In *S. acidocaldarius*, the assembly ATPase, AapE, and the central membrane protein, AapF, homologous to Sso2386 and Sso2387, respectively, are both essential for the assembly of the type IV adhesive pilus (Hence et al., 2012). The sso3139-3141 operon is thought to encode a membrane bound complex, which could function as a secondary receptor for SIRV2 (Deng et al., 2014).

While entry of ravidiviruses, and filamentous archaeal viruses in general, relies on two coordinated adsorption steps, other

systems interact spontaneously with the cell surface. As far back as 1984, SSV1 (*Sulfolobus* spindle-shaped virus 1) was reported to exist in different states: isolated particles, incorporated in typical rosette-like aggregates or even bound to cell-derived membrane (Martin et al., 1984). The best known member of the *Fuselloviridae* family displays a lemon-shaped morphotype with terminal fibers at one of the two pointed ends (Stedman et al., 2015). The set of short, thin filaments of the α -fuselloviruses are involved in viral attachment and association with host-derived structures in general. However, the β -fuselloviruses, SSV6 and ASV1 (*Acidianus* spindle-shaped virus 1), exhibit more pleomorphic virions with three or four thick, slightly curved fibers (Krupovic et al., 2014). Although these appendages do not interact with each other as observed for SSV1, some genomic features strongly suggest that the fibers are composed of host-attachment proteins (Redder et al., 2009). Notably, one gene common to all family members (SSV1_C792) and two genes in β -fuselloviruses (SSV6_C213 and SSV6_B1232) encode for the protein responsible for terminal fibers. This protein shares a similar fold with the adsorption protein P2 of bacteriophage PRD1 (Grahn et al., 2002; Redder et al., 2009). In addition, the pointed end of the enveloped virus ABV (*Acidianus* bottle-shaped virus), from the *Ampullaviridae* family, is involved in attachment to membrane vesicles and formation of virion aggregates (Haring et al., 2005a). Therefore, even if data are still scarce, interaction with cellular membranes appears to be a common feature of hyperthermophilic archaeal viruses that contain a lipidic envelope. This particularly interesting feature merits further investigation.

Release of Viral Genome

Receptor recognition and binding typically induce a cascade of events that start with structural reorganization of the virions and lead to viral genome penetration through the cell envelope (Dimitrov, 2004). Non-enveloped viruses either inject the genome into the cell interior while leaving the empty capsid associated with the cell envelope or deliver the nucleic acids concomitantly with disassembly of the virion at the cell surface. Superficially, the entry of SIRV2 is similar to that of Ff inoviruses or flagellotrophic phages, which bind F-pili and flagella respectively (Guerrero-Ferreira et al., 2011; Rakonjac et al., 2011). First, the interaction with host pili-like structures has been shown and secondly, partially broken particles have been observed at the cellular membrane (Quemin et al., 2013; Figure 1). Notably, no archaeal retraction pili has been identified so far and flagella (called archaella in archaea) of *Sulfolobus* are considerably thicker than the filaments to which SIRV2 binds (Lassak et al., 2012). Additional experiments are needed in order to determine whether the mechanisms of SIRV2 translocation and genome delivery are related to those employed by Ff inoviruses and flagellotrophic bacteriophages, or are completely novel.

Lipid-containing viruses display unusual virion architecture and appear to make direct contact with the plasma membrane. It is reasonable to assume that enveloped viruses rely on a fundamentally different entry mechanism to that employed by non-enveloped filamentous viruses, such as ravidiviruses. They might deliver their genetic material into the cell interior by fusion between the cytoplasmic membrane and the viral

envelope in a similar fashion to the eukaryotic enveloped viruses (Vaney and Rey, 2011). ATV (*Acidianus* two-tailed virus) resembles fuselloviruses with virions extruded from host cells as lemon-shaped. However, ATV has been classified within the *Bicaudaviridae* partly due to its peculiar life cycle (Haring et al., 2005c). Surprisingly, at temperatures close to that of its natural habitat (85°C), the released tail-less particles show the formation of two long tails protruding from the pointed ends. These extracellular developed tubes contain a thin filament inside and terminate in an anchor-like structure, not observed in the tail-less progeny. The two virion forms, tail-less and two-tailed, were reported to be infectious, thereby indicating that the termini are not involved in the initial stages of infection (Prangishvili et al., 2006b). However, genomic analysis as well as molecular studies highlighted some viral encoded proteins that could be important during infection. For example, the three largest open reading frames (ORFs) and one of the CPs have putative coiled-coil domains, which are usually associated with specific protein–protein interactions and protein complex formation. Moreover, two other proteins carry proline-rich regions (ORF567 and ORF1940) similar to the protein TPX and are abundant during infection by lipothrixvirus TTV1 (*Thermoproteus tenax* virus 1; Neumann and Zillig, 1990). Notably, in particular the motif TPTP has been implicated in host protein recognition for the African swine fever virus (Kay-Jackson, 2004). Finally, pull-down experiments provided evidence for a strong interaction between the ATV protein P529 and OppAss as well as cellular Sso1273, encoding a viral AAA ATPase. The cellular OppAss, an N-linked glycoprotein, is most likely part of the binding components of the ABC transporter system. It is encoded within the same operon and could serve as a receptor. It has also been proposed that the AAA ATPase would trigger ATV host cell receptor recognition. This is based on the hypothetical requirement of its endonuclease activity for the cleavage of the circular viral DNA prior to entry in the cell (Erdmann et al., 2011).

The case of the bottle-shaped virus ABV is also particularly intriguing. The enveloped particles display an elaborate organization with a funnel-shaped body composed of the “stopper,” the nucleoprotein core and the inner core. Presumably, the so-called “stopper” takes part in binding to the cellular receptor and is the only component to which the viral genome is directly attached. Therefore, it has been suggested that the “stopper” could play the role of an “injection needle” in a manner similar to that found in bacterial viruses. Actually, it is well known that head-tail bacteriophages belonging to the *Caudovirales* order use this transmembrane pathway for channeling and delivery of nucleic acids (Poranen et al., 2002). The inner core of ABV virions is the most labile part and could undergo structural changes that would facilitate the release of viral DNA (Haring et al., 2005a). Whether the energy accumulated in the structure after packaging of the supercoiled nucleoprotein is sufficient to transport the whole genetic material into the cytoplasm is unclear. However, relaxation of the nucleoprotein filament, wound up as an inverse cone, concomitantly with its funneling into the cell could be an efficient way of utilizing the energy stored during packaging for DNA injection as previously observed in bacteria (Poranen et al., 2002).

How archaeal viruses interact with the cell surface and deliver the viral genome into the host cytoplasm is still puzzling. Some systems, ruddiviruses and lipothrixviruses, show similarities to their bacterial counterparts while others, fuselloviruses, bicaudavirus and ampullavirus, could be related to eukaryotic viruses. Identification of the pathways utilized by both filamentous and unique lipid-containing viruses represents a great challenge and one of the main issues that should be tackled in the near future. It is noteworthy that the S-layer is generally composed of heavily glycosylated proteins and many archaeal viruses exhibit glycosylated capsid proteins. The fact that several glycosyltransferases are encoded in viral genomes (Krupovic et al., 2012) is particularly intriguing. Indeed, protein glycosylation is an important process, which could be involved in virion stability and/or interaction with the host cell (Markine-Goriaynoff et al., 2004; Meyer and Albers, 2013).

Strategies for Viral Escape from the Host Cell

The last and essential step of the viral infection cycle is escape of viral particles from the host cell. So far, the egress mechanism has been analyzed for only a small subset of archaeal viruses (Torsvik and Dundas, 1974; Bize et al., 2009; Brumfield et al., 2009; Snyder et al., 2013a). Some viruses are completely lytic, while others are apparently stably produced without causing evident cell lysis (Bettstetter et al., 2003). In addition, there are temperate archaeal viruses with a lysogenic life cycle for which induction of virion production in some cases leads to cell disruption (Janekovic et al., 1983; Schleper et al., 1992; Prangishvili et al., 2006b).

The release mechanisms utilized by archaeal viruses can be divided in two categories: those for which the cell membrane is disrupted and those where the membrane integrity remains intact. The strategy for egress is linked with the assembly mechanism of new virions. Some archaeal viruses are known to mature inside the cell cytoplasm and provoke lysis, such as STIV1 (*Sulfolobus* turreted icosahedral virus) and SIRV2 (Bize et al., 2009; Brumfield et al., 2009; Fu et al., 2010). However, most non-lytic viruses undergo final maturation concomitantly with passage through the cell membrane (Roine and Bamford, 2012) or even in the extracellular environment, as observed for ATV (Haring et al., 2005c).

Cell Membrane Disruption

Lysis by Complete Membrane Disruption

Disruption of cell membranes can be caused by lytic or temperate viruses. In case of temperate viruses the cell lysis occurs typically after induction of virus replication and virion formation. Virion production of lysogenic viruses can be induced by various stimuli such as; UV radiation, addition of mitomycin C, starvation or shift from aerobic to anaerobic growth (Janekovic et al., 1983; Schleper et al., 1992; Prangishvili et al., 2006b; Mochizuki et al., 2011).

The first archaeal viruses were isolated from hypersaline environments long before archaea were recognized as a separate domain of life (Torsvik and Dundas, 1974; Wais et al., 1975). These viruses infect halophiles, which belong to the phylum

Euryarchaeota. The viral particles exhibit a head-and-tail morphology classical for bacterial viruses. Infection with these viruses resulted in complete lysis of the cells, suggested by a decrease in culture turbidity. Later on, more euryarchaeal viruses were isolated from hypersaline or anaerobic environments, and several of these viruses displayed non-head-tail morphologies such as icosahedral or spindle shapes. Again, in some cases, optical density diminishes with time after viral infection, indicating that a part of these viruses initiate cell lysis (Bath and Dyall-Smith, 1998; Porter et al., 2005; Jaakkola et al., 2012). However, several euryarchaeal viruses apparently do not cause cell lysis.

Amongst hyperthermophilic crenarchaeal viruses there has only been a single report of a decrease in the turbidity of infected cultures (Prangishvili et al., 2006a). In this case, induction of virion production of the lysogenic viruses TTV1-3 led to cell lysis, which was measured by decreasing turbidity (Janekovic et al., 1983). Lysis induced by archaeal viruses can either be coupled with virion production (Jaakkola et al., 2012), or take place after the largest virion burst, therefore raising the possibility of an additional release mechanism in such systems (Bath and Dyall-Smith, 1998; Porter et al., 2005, 2013). Although measurement of optical density is a classical method for the characterization of viral cycles and decrease in turbidity has been observed for several archaeal viruses, no molecular mechanism to achieve complete membrane disruption in archaea has been proposed as yet.

Bacterial virus-host systems are widely studied and as a result the mechanism of lysis used by bacterial viruses is better understood. Bacterial viruses typically induce cell lysis by degradation of the cell wall, which is achieved by muralytic endolysins (Young, 2013). In addition, most bacterial viruses encode small proteins named holins (Bernhardt et al., 2001a,b; Catalao et al., 2013). Holins usually accumulate harmlessly in the bacterial cell membrane until a critical concentration is reached and nucleation occurs. Nucleation results in formation of two dimensional aggregates, “holin rafts,” that rapidly expand and create pores in lipid layers through which the endolysins can reach the cell wall (Young, 2013). In gram negative bacteria the presence of an outer membrane requires additional virus-encoded proteins, spanins, which are suggested to induce fusion of the inner and outer membrane (Berry et al., 2012). After an initial degradation of the peptidoglycan cell wall, the cells burst due to osmotic pressure, explaining total loss of turbidity observed for infected bacterial cultures (Berry et al., 2012). Accurate timing of lysis is essential for successful virus reproduction and is achieved by regulation of holin expression (Young, 2013). Since archaea lack a peptidoglycan cell wall, endolysin-holin egress systems are not effective in archaea. Only a few archaeal species contain a peptidoglycan-like cell wall consisting of pseudomurein polymers (Albers and Meyer, 2011). The oligosaccharide backbone and amino acid interbridges of murein and pseudomurein are different, rendering bacterial endolysins ineffective to pseudomurein (Visweswaran et al., 2011). However, pseudomurein degrading enzymes are encoded by a few archaeal viruses infecting methanogens; the integrated provirus ψ M100 from *Methanothermobacter wolfeii* and the virus ψ M1 infecting *M. marburgensis* (Luo et al., 2001). How these intracellularly

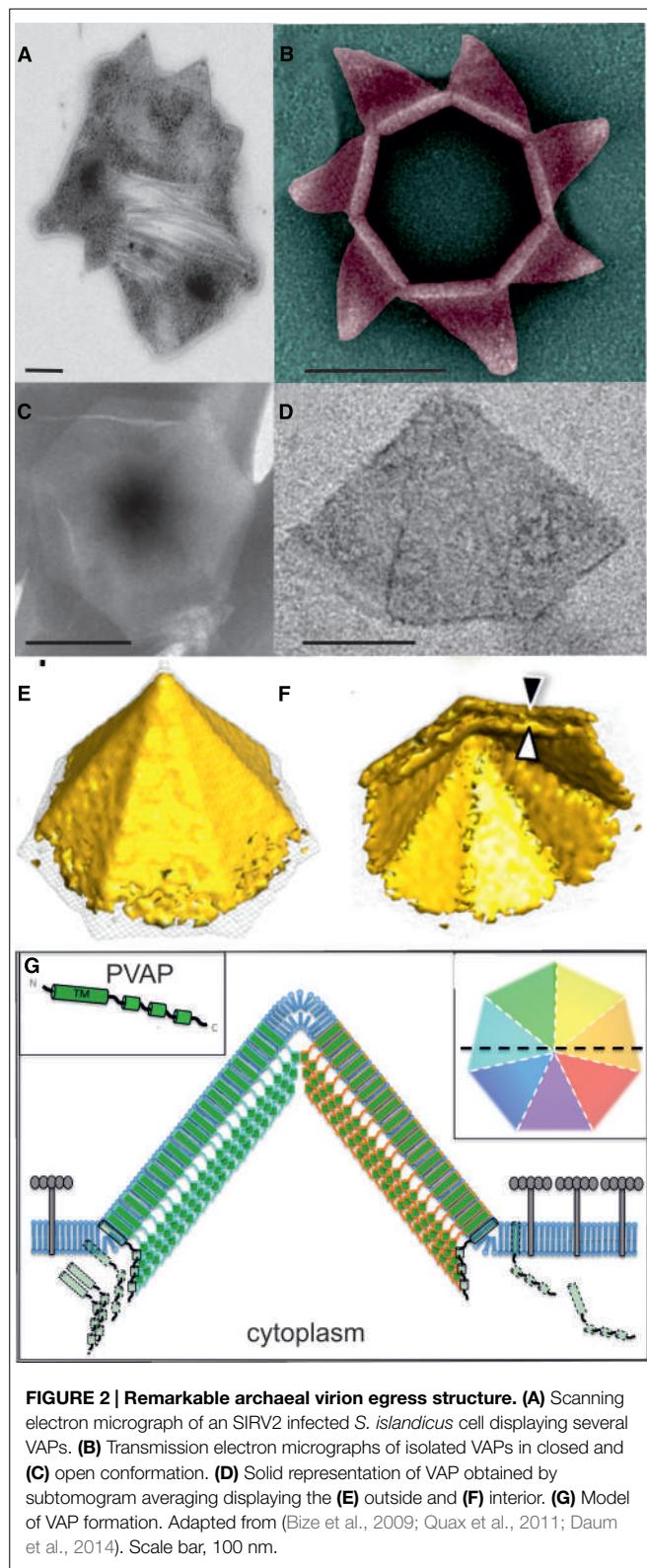
produced viral endolysis traverse the archaeal cell membrane in order to degrade the pseudomurein cell wall is not clear, since the mandatory pore forming holins have not been identified in the genomes of these viruses. The possible presence of archaeal holins could be currently overlooked, as genes encoding holins share generally very little sequence similarity, making it difficult to predict their presence in genomes (Saier and Reddy, 2015).

The large majority of archaea lack a pseudomurein cell wall. Therefore instead of an endolysin-holin system, a fundamentally different lysis mechanism would be required for release of virions from these cell wall lacking archaea. One hypothesis is that archaeal viruses employ holins to disrupt the cell membrane, possibly combined with proteolytic enzymes in order to degrade the S-layer. To date there are about a dozen holin homologs identified in archaeal genomes based on sequence similarity (Reddy and Saier, 2013), but none of the predicted proteins have been tested *in vivo*. Moreover, not a single holin-encoding gene has been identified in the genomes of currently isolated archaeal viruses (Reddy and Saier, 2013; Saier and Reddy, 2015). In addition, specific enzymes capable of S-layer degradation are currently unknown and S-layer proteins and sugars display a large diversity in different species (Albers and Meyer, 2011). Thus in contrast to bacterial endolysins that degrade peptidoglycan cell walls of virtually all bacteria, specific tailor made proteases would be required to degrade archaeal S-layers of different species.

Lysis by Formation of Defined Apertures

The egress mechanism of only two archaeal viruses (STIV1 and SIRV2) has been studied in high molecular detail. Both employ a release mechanism that relies on the formation of pyramidal shaped egress structures, which are unique to archaeal systems (Bize et al., 2009; Brumfield et al., 2009; Quax et al., 2011; Snyder et al., 2011). At first glance, both viruses were regarded as non-lytic viruses, since a decrease in cell culture turbidity was never observed (Prangishvili et al., 1999; Rice et al., 2004). However, the use of several electron microscopy techniques clearly showed that the two viruses induced cell lysis (Bize et al., 2009; Brumfield et al., 2009). Their particular lysis mechanism yields empty cell ghosts explaining the maintenance of culture turbidity.

Infection by SIRV2 and STIV1 leads to formation of several pyramidal shaped structures on the cell membrane of *S. islandicus* and *S. solfataricus* respectively (Bize et al., 2009; Brumfield et al., 2009; Prangishvili and Quax, 2011; **Figure 2A**). These virus-associated pyramids (VAPs) exhibit sevenfold rotational symmetry and protrude through the S-layer (Quax et al., 2011; Snyder et al., 2011; **Figures 2B–D**). At the end of the infection cycle, the seven facets of the VAPs open outward, generating large apertures through which assembled virions exit from the cell (Fu et al., 2010; Quax et al., 2011; Daum et al., 2014; **Figure 2B**). The baseless VAP consist of multiple copies of a 10 kDa viral encoded protein, PVAP (STIV1_C92/SIRV2_P98) (Quax et al., 2010; Snyder et al., 2013a). This protein contains a transmembrane domain, but lacks a signal sequence and seems to be inserting in membranes based on hydrophobicity of its transmembrane domain (Quax et al., 2010; Daum et al., 2014). PVAP has the remarkable property to form pyramidal structures in virtually all biological membranes, as was demonstrated by heterologous



expression of PVAP in archaea, bacteria and eukaryotes (Quax et al., 2011; Snyder et al., 2013a; Daum et al., 2014).

Nucleation of the PVAP-induced structure starts on the cell membrane, most likely with the formation of a heptamer of

PVAP subunits (Daum et al., 2014). The structures develop by the outward expansion of their seven triangular facets. They reach sizes of up to 200 nm in diameter, both in natural and heterologous systems (Quax et al., 2011; Daum et al., 2014). In contrast to bacterial holin rafts, the formation of VAPs is not a sudden process depending on a critical protein concentration. PVAP transcripts steadily increase throughout the infection cycle and PVAP integrates in the membrane until late stages of infection (Quax et al., 2010, 2013; Maaty et al., 2012). Although VAPs are slowly formed, their actual opening is quite rapid (Bize et al., 2009; Brumfield et al., 2009; Snyder et al., 2011; Daum et al., 2014). The nature of the signal triggering this opening has not been identified yet. VAPs, formed after heterologous PVAP expression, in bacteria and eukaryotes were never observed in open conformation, suggesting that an archaeal specific factor is required (Daum et al., 2014). It has been proposed that the archaeal ESCRT (Endosomal Sorting Complex Required for Transport) machinery could be involved in the STIV1 VAP-based exit (Snyder et al., 2013b). Considering that genes encoding ESCRT machinery are specifically down regulated during SIRV2 infection (Quax et al., 2013), and that STIV1 contains in contrast to SIRV2 an inner lipid layer (Veesler et al., 2013), STIV1 requirement of the ESCRT system might be independent from VAP-induced lysis.

The ultrastructure of VAPs of SIRV2 was studied by whole cell cryo-tomography and subtomogram averaging. This revealed the presence of two layers, of which the outer one is continuous with the cell membrane and presumably formed by the N-terminal transmembrane domain (Daum et al., 2014; Figures 2E,F). The inner layer represents a protein sheet formed by tight protein–protein interactions of the C-terminal domain of the protein (Daum et al., 2014). The strong interactions between PVAP monomers are suggested to exclude most lipids and membrane proteins from the VAP assembly site, in a similar fashion as holin raft formation (White et al., 2011; Figure 2G). S-layer proteins are anchored in the membrane, and consequently will be excluded from the VAP assembly site, providing a strategy for VAP protrusion through the S-layer.

The described VAP-based egress mechanism is archaeal specific. Homologues of PVAP are only found amongst some archaeal viruses (Quax et al., 2010). However, the majority of archaeal viruses lack PVAP, suggesting that they rely on a different and as yet unknown mechanism for egress.

Viral Extrusion without Membrane Disruption

While the first isolated archaeal viruses were lytic, subsequent characterization of more viruses revealed that the large majority do not cause lysis of the host cell. To date, lytic viruses make up half of the viruses infecting euryarchaeota, and only three in crenarchaea (Torsvik and Dundas, 1974; Wais et al., 1975; Janekovic et al., 1983; Bize et al., 2009; Brumfield et al., 2009; Pina et al., 2011). In addition, some studies indicate that free virions can be observed before disruption of archaeal cells, suggesting that another egress mechanism exists, which preserves cell membrane integrity. It might be possible that some lytic archaeal viruses have been currently overlooked due to special characteristics of their lysis mechanism, as was the case for STIV1 and SIRV2 (Prangishvili et al., 1999; Rice et al., 2004). Nevertheless, the low

number of lytic archaeal viruses contrasts with the situation in bacteria, for which lytic viruses are very common. The majority of archaeal viruses are thought to be continuously produced without integrating into the host genome or killing their hosts (Pina et al., 2014). This equilibrium between viruses and cells is referred to as a “stable carrier state” (Bettstetter et al., 2003; Prangishvili and Garrett, 2005; Prangishvili et al., 2006a). The nature of this stable carrier state and the mechanisms by which virions are extruded from archaea without causing cell lysis, remain poorly understood.

In contrast to the situation in archaea, the majority of bacterial viruses are lytic. Almost all bacterial viruses exit via the holin based mechanism described above. However, an exception to the rule are the bacterial filamentous viruses belonging to the *Inoviridae* that egress without causing cell lysis (Rakonjac et al., 2011). The majority of the inoviruses infect gram negative bacteria. Assembly of inoviruses is finalized during particle extrusion. The interaction between the packaging signal of the viral genome and the cellular membrane initiates the exit step (Russel and Model, 1989). Virally encoded proteins are thought to form pores in the inner membrane through which the DNA is extruded. Multiple copies of the major CP accumulate in the inner membrane and associate with the ssDNA viral genome while it is passing through the virus-induced pores (Rakonjac et al., 1999). A barrel-like structure in the outer membrane permits the release of progeny and is composed of multiple copies of a virus-encoded protein with homology to proteins of type II secretion systems and type IV pili (Marciano et al., 2001). Alternatively, other inoviruses use the host secretion machinery to traverse the outer membrane (Davis et al., 2000; Bille et al., 2005). Even though replication of the viral genome and constituents might burden the cell, the infection of inoviruses does not lead to cell death and is a continuous process. There are several archaeal filamentous viruses known. However, filamentous archaeal viruses are not related to the bacterial inoviruses, nor encode homologs of the secretion-like proteins involved in egress of inoviruses (Janekovic et al., 1983; Bize et al., 2009; Quax et al., 2010; Pina et al., 2014). Therefore the filamentous archaeal viruses must rely on an alternative mechanism for viral extrusion from the cell.

Interestingly, lipid-containing archaeal viruses are quite common (Roine and Bamford, 2012). There are some archaeal icosahedral viruses that possess an inner membrane, such as STIV and SH1 (Bamford et al., 2005; Khayat et al., 2005; Porter et al., 2005). In addition, the filamentous lipothrixviruses (Janekovic et al., 1983; Arnold et al., 2000; Bettstetter et al., 2003), the spherical virus PSV (*Pyrobaculum* spherical virus; Haring et al., 2004) and the pleiomorphic euryarchaeal viruses (Pietila et al., 2009, 2013) all contain an external lipid envelope. The lipids are typically derived from the host cell. Several eukaryotic viruses contain a membrane that is usually obtained during “budding,” a process by which particles egress without disturbing

the membrane integrity. Eukaryotic enveloped viruses either encode their own scission proteins, or hijack vesicle formation machinery of their host (Rossmann and Lamb, 2013). Archaea are also reported to produce vesicles (Soler et al., 2008; Ellen et al., 2011), and the machinery responsible for vesicle production might be utilized by lipid envelope containing viruses in archaea as well. In particular, the pleiomorphic viruses infecting euryarchaea are likely to be released through budding as their envelope has the same lipid composition as the host they infect (Pietila et al., 2009; Roine et al., 2010).

The most common scission machinery employed by eukaryotic viruses is the ESCRT system (Votteler and Sundquist, 2013). In eukaryotes these proteins are responsible for endosomal sorting in the multi vesicular body. Well-characterized viruses such as Ebola and human immunodeficiency virus (HIV) use the ESCRT proteins during egress (Harty et al., 2000; Weissenhorn et al., 2013). Interestingly, proteins homologous to ESCRT components have been identified in several archaea, where they are involved in cell division (Lindas et al., 2008; Samson et al., 2008; Makarova et al., 2010; Pelve et al., 2011). These proteins represent potential players in budding-like extrusion processes in archaea. The mechanism underlying the release of temperate archaeal viruses remains largely unexplored and represents an appealing area of research that should shed light on original and unconventional strategies.

Concluding Remarks

The last few years have shown a steady increase in an understanding of archaeal virus-host interactions, therefore revealing the first insights into viral interactions with the archaeal membrane. Viruses have developed various strategies to cross the membrane. These strategies are adapted to the nature of the cell envelope of their host. Some archaeal viruses employ fascinating novel mechanisms, while others appear to rely on processes that at first sight are analogous to their bacterial counterparts. Additional research will help to determine to which extent bacterial, eukaryotic and archaeal virospheres are evolutionary related. The uniqueness of the archaeal cell surface, and the diversity of the currently described archaeal entry and egress mechanisms, argue in favor of future discovery of more innovative and surprising molecular mechanisms.

Acknowledgments

We thank Dr. D. Prangishvili and Dr. S. Gill for critical reading of the manuscript, useful suggestions and comments. This work was supported by l’Agence Nationale de la Recherche, project “EXAVIR,” by a FWO Pegasus Marie Curie fellowship to TQ. and by a grant from the French government and the Université Pierre et Marie Curie, Paris VI to EQ.

References

- Albers, S. V., and Meyer, B. H. (2011). The archaeal cell envelope. *Nat. Rev. Microbiol.* 9, 414–426. doi: 10.1038/nrmicro2576

- Arnold, H. P., Zillig, W., Ziese, U., Holz, I., Crosby, M., Utterback, T., et al. (2000). A novel lipothrixvirus, SIFV, of the extremely thermophilic crenarchaeon *Sulfolobus*. *Virology* 267, 252–266. doi: 10.1006/viro.1999.0105

- Bamford, D. H., Ravantti, J. J., Ronnholm, G., Laurinavicius, S., Kukkaro, P., Dyall-Smith, M., et al. (2005). Constituents of SH1, a novel lipid-containing virus infecting the halophilic euryarchaeon *Haloarcula hispanica*. *J. Virol.* 79, 9097–9107. doi: 10.1128/jvi.79.14.9097-9107.2005
- Bath, C., and Dyall-Smith, M. L. (1998). His1, an archaeal virus of the Fuselloviridae family that infects *Haloarcula hispanica*. *J. Virol.* 72, 9392–9395.
- Bergh, O., Borsheim, K. Y., Bratbak, G., and Heldal, M. (1989). High abundance of viruses found in aquatic environments. *Nature* 340, 467–468. doi: 10.1038/340467a0
- Bernhardt, T. G., Struck, D. K., and Young, R. (2001a). The lysis protein E of phi X174 is a specific inhibitor of the MraY-catalyzed step in peptidoglycan synthesis. *J. Biol. Chem.* 276, 6093–6097. doi: 10.1074/jbc.M007638200
- Bernhardt, T. G., Wang, I. N., Struck, D. K., and Young, R. (2001b). A protein antibiotic in the phage Qbeta virion: diversity in lysis targets. *Science* 292, 2326–2329. doi: 10.1126/science.1058289
- Berry, J., Rajaire, M., Pang, T., and Young, R. (2012). The spanin complex is essential for lambda lysis. *J. Bacteriol.* 194, 5667–5674. doi: 10.1128/JB.01245-12
- Bettstetter, M., Peng, X., Garrett, R. A., and Prangishvili, D. (2003). AFV1, a novel virus infecting hyperthermophilic archaea of the genus *Acidianus*. *Virology* 315, 68–79. doi: 10.1016/S0042-6822(03)00481-1
- Bille, E., Zahar, J. R., Perrin, A., Morelle, S., Kriz, P., Jolley, K. A., et al. (2005). A chromosomally integrated bacteriophage in invasive meningococci. *J. Exp. Med.* 201, 1905–1913. doi: 10.1084/jem.20050112
- Bize, A., Karlsson, E. A., Ekefjard, K., Quax, T. E., Pina, M., Prevost, M. C., et al. (2009). A unique virus release mechanism in the Archaea. *Proc. Natl. Acad. Sci. U.S.A.* 106, 11306–11311. doi: 10.1073/pnas.0901238106
- Bize, A., Peng, X., Prokofeva, M., Maclellan, K., Lucas, S., Forterre, P., et al. (2008). Viruses in acidic geothermal environments of the Kamchatka Peninsula. *Res. Microbiol.* 159, 358–366. doi: 10.1016/j.resmic.2008.04.009
- Borsheim, K. Y., Bratbak, G., and Heldal, M. (1990). Enumeration and biomass estimation of planktonic bacteria and viruses by transmission electron microscopy. *Appl. Environ. Microbiol.* 56, 352–356.
- Brumfield, S. K., Ortmann, A. C., Ruigrok, V., Suci, P., Douglas, T., and Young, M. J. (2009). Particle assembly and ultrastructural features associated with replication of the lytic archaeal virus *Sulfolobus* turreted icosahedral virus. *J. Virol.* 83, 5964–5970. doi: 10.1128/jvi.02668-08
- Catalao, M. J., Gil, F., Moniz-Pereira, J., Sao-Jose, C., and Pimentel, M. (2013). Diversity in bacterial lysis systems: bacteriophages show the way. *FEMS Microbiol. Rev.* 37, 554–571. doi: 10.1111/1574-6976.12006
- Ceballos, R. M., Marceau, C. D., Marceau, J. O., Morris, S., Clore, A. J., and Stedman, K. M. (2012). Differential virus host-ranges of the Fuselloviridae of hyperthermophilic Archaea: implications for evolution in extreme environments. *Front. Microbiol.* 3:295. doi: 10.3389/fmicb.2012.00295
- Daum, B., Quax, T. E., Sachse, M., Mills, D. J., Reimann, J., Yildiz, O., et al. (2014). Self-assembly of the general membrane-remodeling protein PVAP into sevenfold virus-associated pyramids. *Proc. Natl. Acad. Sci. U.S.A.* 111, 3829–3834. doi: 10.1073/pnas.1319245111
- Davis, B. M., Lawson, E. H., Sandkvist, M., Ali, A., Sozhamannan, S., and Waldor, M. K. (2000). Convergence of the secretory pathways for cholera toxin and the filamentous phage, CTXphi. *Science* 288, 333–335. doi: 10.1126/science.288.5464.333
- Deng, L., He, F., Bhoobalan-Chitty, Y., Martinez-Alvarez, L., Guo, Y., and Peng, X. (2014). Unveiling cell surface and type IV secretion proteins responsible for archaeal *Rudivirüs* entry. *J. Virol.* 88, 10264–10268. doi: 10.1128/jvi.01495-14
- Dimitrov, D. S. (2004). Virus entry: molecular mechanisms and biomedical applications. *Nat. Rev. Microbiol.* 2, 109–122. doi: 10.1038/nrmicro817
- Ellen, A. F., Rohulya, O. V., Fusetti, F., Wagner, M., Albers, S. V., and Driessens, A. J. (2011). The sulfolobicin genes of *Sulfolobus acidocaldarius* encode novel antimicrobial proteins. *J. Bacteriol.* 193, 4380–4387. doi: 10.1128/jb.05028-11
- Erdmann, S., Scheele, U., and Garrett, R. A. (2011). AAA ATPase p529 of *Acidianus* two-tailed virus ATV and host receptor recognition. *Virology* 421, 61–66. doi: 10.1016/j.virol.2011.08.029
- Fagan, R. P., and Fairweather, N. F. (2014). Biogenesis and functions of bacterial S-layers. *Nat. Rev. Microbiol.* 12, 211–222. doi: 10.1038/nrmicro3213
- Fu, C. Y., Wang, K., Gan, L., Lanman, J., Khayat, R., Young, M. J., et al. (2010). In vivo assembly of an archaeal virus studied with whole-cell electron cryotomography. *Structure* 18, 1579–1586. doi: 10.1016/j.str.2010.10.005
- Grahn, A. M., Daugelavicius, R., and Bamford, D. H. (2002). The small viral membrane associated protein P32 is involved in bacteriophage PRD1 DNA entry. *J. Virol.* 76, 4866–4872. doi: 10.1128/jvi.76.10.4866-4872.2002
- Guerrero-Ferreira, R. C., Viollier, P. H., Ely, B., Poindexter, J. S., Georgieva, M., Jensen, G. J., et al. (2011). Alternative mechanism for bacteriophage adsorption to the motile bacterium *Caulobacter crescentus*. *Proc. Natl. Acad. Sci. U.S.A.* 108, 9963–9968. doi: 10.1073/pnas.1012388108
- Haring, M., Peng, X., Brugger, K., Rachel, R., Stetter, K. O., Garrett, R. A., et al. (2004). Morphology and genome organization of the virus PSV of the hyperthermophilic archaeal genera *Pyrobaculum* and *Thermoproteus*: a novel virus family, the Globuloviridae. *Virology* 323, 233–242. doi: 10.1016/j.virol.2004.03.002
- Haring, M., Rachel, R., Peng, X., Garrett, R. A., and Prangishvili, D. (2005a). Viral diversity in hot springs of Pozzuoli, Italy, and characterization of a unique archaeal virus, *Acidianus* bottle-shaped virus, from a new family, the Ampullaviridae. *J. Virol.* 79, 9904–9911. doi: 10.1128/JVI.79.15.9904-9911.2005
- Haring, M., Vestergaard, G., Brugger, K., Rachel, R., Garrett, R. A., and Prangishvili, D. (2005b). Structure and genome organization of AFV2, a novel archaeal lipothrixvirus with unusual terminal and core structures. *J. Bacteriol.* 187, 3855–3858. doi: 10.1128/JB.187.11.3855-3858.2005
- Haring, M., Vestergaard, G., Rachel, R., Chen, L., Garrett, R. A., and Prangishvili, D. (2005c). Virology: independent virus development outside a host. *Nature* 436, 1101–1102. doi: 10.1038/4361101a
- Harty, R. N., Brown, M. E., Wang, G., Huibregtse, J., and Hayes, F. P. (2000). A PPxY motif within the VP40 protein of Ebola virus interacts physically and functionally with a ubiquitin ligase: implications for filovirus budding. *Proc. Natl. Acad. Sci. U.S.A.* 97, 13871–13876. doi: 10.1073/pnas.250277297
- Henché, A. L., Ghosh, A., Yu, X., Jeske, T., Egelman, E., and Albers, S. V. (2012). Structure and function of the adhesive type IV pilus of *Sulfolobus acidocaldarius*. *Environ. Microbiol.* 14, 3188–3202. doi: 10.1111/j.1462-2920.2012.02898.x
- Jaakkola, S. T., Penttilä, R. K., Vilén, S. T., Jalasvuori, M., Ronnholm, G., Bamford, J. K., et al. (2012). Closely related archaeal *Haloarcula hispanica* icosahedral viruses HHIV-2 and SH1 have nonhomologous genes encoding host recognition functions. *J. Virol.* 86, 4734–4742. doi: 10.1128/jvi.06666-11
- Janekovic, D., Wunderl, S., Holz, I., Zillig, W., Gierl, A., and Neumann, H. (1983). TTV1, TTV2 and TTV3, a family of viruses of the extremely thermophilic, anaerobic, sulfur reducing archaeabacterium *Thermoproteus tenax*. *Mol. Gen. Genet.* 192, 39–45. doi: 10.1007/BF00327644
- Kay-Jackson, P. C. (2004). The CD2v protein of African swine fever virus interacts with the actin-binding adaptor protein SH3P7. *J. Gen. Virol.* 85, 119–130. doi: 10.1099/vir.0.19435-0
- Khayat, R., Tang, L., Larson, E. T., Lawrence, C. M., Young, M., and Johnson, J. E. (2005). Structure of an archaeal virus capsid protein reveals a common ancestry to eukaryotic and bacterial viruses. *Proc. Natl. Acad. Sci. U.S.A.* 102, 18944–18949. doi: 10.1073/pnas.0506383102
- Krupovic, M., Quemin, E. R., Bamford, D. H., Forterre, P., and Prangishvili, D. (2014). Unification of the globally distributed spindle-shaped viruses of the Archaea. *J. Virol.* 88, 2354–2358. doi: 10.1128/JVI.02941-13
- Krupovic, M., White, M. F., Forterre, P., and Prangishvili, D. (2012). Postcards from the edge: structural genomics of archaeal viruses. *Adv. Virus Res.* 82, 33–62. doi: 10.1016/B978-0-12-394621-8.00012-1
- Kukkaro, P., and Bamford, D. H. (2009). Virus-host interactions in environments with a wide range of ionic strengths. *Environ. Microbiol. Rep.* 1, 71–77. doi: 10.1111/j.1758-2229.2008.00007.x
- Lassak, K., Ghosh, A., and Albers, S. V. (2012). Diversity, assembly and regulation of archaeal type IV pili-like and non-type-IV pili-like surface structures. *Res. Microbiol.* 163, 630–644. doi: 10.1016/j.resmic.2012.10.024
- Lindas, A. C., Karlsson, E. A., Lindgren, M. T., Ettema, T. J., and Bernander, R. (2008). A unique cell division machinery in the Archaea. *Proc. Natl. Acad. Sci. U.S.A.* 105, 18942–18946. doi: 10.1073/pnas.0809467105
- Luo, Y., Pfister, P., Leisinger, T., and Wasserfallen, A. (2001). The genome of archaeal prophage PsiM100 encodes the lytic enzyme responsible for autolysis of Methanothermobacter wolfeii. *J. Bacteriol.* 183, 5788–5792. doi: 10.1128/jb.183.19.5788-5792.2001
- Maaty, W. S., Steffens, J. D., Heinemann, J., Ortmann, A. C., Reeves, B. D., Biswas, S. K., et al. (2012). Global analysis of viral infection in an archaeal model system. *Front. Microbiol.* 3:411. doi: 10.3389/fmicb.2012.00411

- Makarova, K. S., Yutin, N., Bell, S. D., and Koonin, E. V. (2010). Evolution of diverse cell division and vesicle formation systems in Archaea. *Nat. Rev. Microbiol.* 8, 731–741. doi: 10.1038/nrmicro2406
- Markine-Goriaynoff, N., Gillet, L., Van Etten, J. L., Korres, H., Verma, N., and Vanderpasschen, A. (2004). Glycosyltransferases encoded by viruses. *J. Gen. Virol.* 85, 2741–2754.
- Marciano, D. K., Russel, M., and Simon, S. M. (2001). Assembling filamentous phage occlude pIV channels. *Proc. Natl. Acad. Sci. U.S.A.* 98, 9359–9364. doi: 10.1073/pnas.161170398
- Martin, A., Yeats, S., Janevic, D., Reiter, W.-D., Aicher, W., and Zillig, W. (1984). SAV 1, a temperate u.v.-inducible DNA virus-like particle from the archaeabacterium *Sulfolobus acidocaldarius* isolate B12. *EMBO J.* 3, 2165–2168.
- Meyer, B. H., and Albers, S. V. (2013). Hot and sweet: protein glycosylation in *Crenarchaeota*. *Biochem. Soc. Trans.* 41, 384–3892. doi: 10.1042/BST20120296
- Mochizuki, T., Sako, Y., and Prangishvili, D. (2011). Provirus induction in hyperthermophilic archaea: characterization of *Aeropyrum pernix* spindle-shaped virus 1 and *Aeropyrum pernix* ovoid virus 1. *J. Bacteriol.* 193, 5412–5419. doi: 10.1128/jb.05101-11
- Molineux, I. J., and Panja, D. (2013). Popping the cork: mechanisms of phage genome ejection. *Nat. Rev. Microbiol.* 11, 194–204. doi: 10.1038/nrmicro2988
- Neumann, H., and Zillig, W. (1990). The TTV1-encoded viral protein TPX: primary structure of the gene and the protein. *Nucleic Acids Res.* 18, 195. doi: 10.1093/nar/18.1.195
- Pelve, E. A., Lindas, A. C., Martens-Habbena, W., De La Torre, J. R., Stahl, D. A., and Bernander, R. (2011). Cdv-based cell division and cell cycle organization in the thaumarchaeon *Nitrosopumilus maritimus*. *Mol. Microbiol.* 82, 555–566. doi: 10.1111/j.1365-2958.2011.07834.x
- Pietila, M. K., Laurinmaki, P., Russell, D. A., Ko, C. C., Jacobs-Sera, D., Butcher, S. J., et al. (2013). Insights into head-tailed viruses infecting extremely halophilic archaea. *J. Virol.* 87, 3248–3260. doi: 10.1128/jvi.03397-12
- Pietila, M. K., Roine, E., Paulin, L., Kalkkinen, N., and Bamford, D. H. (2009). An ssDNA virus infecting archaea: a new lineage of viruses with a membrane envelope. *Mol. Microbiol.* 72, 307–319. doi: 10.1111/j.1365-2958.2009.06642.x
- Pina, M., Basta, T., Quax, T. E., Joubert, A., Baconnais, S., Cortez, D., et al. (2014). Unique genome replication mechanism of the archaeal virus AFV1. *Mol. Microbiol.* 92, 1313–1325. doi: 10.1111/mmi.12630
- Pina, M., Bize, A., Forterre, P., and Prangishvili, D. (2011). The archeoviruses. *FEMS Microbiol. Rev.* 35, 1035–1054. doi: 10.1111/j.1574-6976.2011.00280.x
- Pohlschroder, M., Ghosh, A., Tripepi, M., and Albers, S. V. (2011). Archaeal type IV pilus-like structures—evolutionarily conserved prokaryotic surface organelles. *Curr. Opin. Microbiol.* 14, 357–363. doi: 10.1016/j.mib.2011.03.002
- Poranen, M. M., Daugelavicius, R., and Bamford, D. H. (2002). Common principles in viral entry. *Annu. Rev. Microbiol.* 56, 521–538. doi: 10.1146/annurev.micro.56.012302.160643
- Porter, K., Kukkaro, P., Bamford, J. K., Bath, C., Kivela, H. M., Dyall-Smith, M. L., et al. (2005). SH1: a novel, spherical halovirus isolated from an Australian hypersaline lake. *Virology* 335, 22–33. doi: 10.1016/j.virol.2005.01.043
- Porter, K., Tang, S. L., Chen, C. P., Chiang, P. W., Hong, M. J., and Dyall-Smith, M. (2013). PH1: an archaeovirus of *Haloarcula hispanica* related to SH1 and HHIV-2. *Archaea* 2013, 456318. doi: 10.1155/2013/456318
- Prangishvili, D. (2013). The wonderful world of archaeal viruses. *Annu. Rev. Microbiol.* 67, 565–585. doi: 10.1146/annurev-micro-092412-155633
- Prangishvili, D., Arnold, H. P., Gotz, D., Ziese, U., Holz, I., Kristjansson, J. K., et al. (1999). A novel virus family, the Rudivirusidae: structure, virus-host interactions and genome variability of the *Sulfolobus* viruses SIRV1 and SIRV2. *Genetics* 152, 1387–1396.
- Prangishvili, D., Forterre, P., and Garrett, R. A. (2006a). Viruses of the Archaea: a unifying view. *Nat. Rev. Microbiol.* 4, 837–848. doi: 10.1038/nrmicro1527
- Prangishvili, D., Vestergaard, G., Haring, M., Aramayo, R., Basta, T., Rachel, R., et al. (2006b). Structural and genomic properties of the hyperthermophilic archaeal virus ATV with an extracellular stage of the reproductive cycle. *J. Mol. Biol.* 359, 1203–1216. doi: 10.1016/j.jmb.2006.04.027
- Prangishvili, D., and Garrett, R. A. (2005). Viruses of hyperthermophilic Crenarchaeota. *Trends Microbiol.* 13, 535–542. doi: 10.1016/j.tim.2005.08.013
- Prangishvili, D., Koonin, E. V., and Krupovic, M. (2013). Genomics and biology of Rudivirus, a model for the study of virus-host interactions in Archaea. *Biochem. Soc. Trans.* 41, 443–450. doi: 10.1042/BST20120313
- Prangishvili, D., and Krupovic, M. (2012). A new proposed taxon for double-stranded DNA viruses, the order “Ligamenvirales”. *Arch. Virol.* 157, 791–795. doi: 10.1007/s00705-012-1229-7
- Prangishvili, D., and Quax, T. E. (2011). Exceptional virion release mechanism: one more surprise from archaeal viruses. *Curr. Opin. Microbiol.* 14, 315–320. doi: 10.1016/j.mib.2011.04.006
- Quax, T. E., Krupovic, M., Lucas, S., Forterre, P., and Prangishvili, D. (2010). The *Sulfolobus* rod-shaped virus 2 encodes a prominent structural component of the unique virion release system in Archaea. *Virology* 404, 1–4. doi: 10.1016/j.virol.2010.04.020
- Quax, T. E., Lucas, S., Reimann, J., Peau-Arnaudet, G., Prevost, M. C., Forterre, P., et al. (2011). Simple and elegant design of a virion egress structure in Archaea. *Proc. Natl. Acad. Sci. U.S.A.* 108, 3354–3359. doi: 10.1073/pnas.1018052108
- Quax, T. E., Voet, M., Sismeiro, O., Dillies, M. A., Jagla, B., Coppee, J. Y., et al. (2013). Massive activation of archaeal defense genes during viral infection. *J. Virol.* 87, 8419–8428. doi: 10.1128/jvi.01020-13
- Quemin, E. R., Lucas, S., Daum, B., Quax, T. E., Kuhlbrandt, W., Forterre, P., et al. (2013). First insights into the entry process of hyperthermophilic archaeal viruses. *J. Virol.* 87, 13379–13385. doi: 10.1128/jvi.02742-13
- Quemin, E. R., Prangishvili, D., and Krupovic, M. (2014). Hard out there: understanding archaeal virus biology. *Future Virol.* 9, 703–706. doi: 10.2217/FVL.14.52
- Rakonjac, J., Bennett, N. J., Spagnuolo, J., Gagic, D., and Russel, M. (2011). Filamentous bacteriophage: biology, phage display and nanotechnology applications. *Curr. Issues Mol. Biol.* 13, 51–76.
- Rakonjac, J., Feng, J., and Model, P. (1999). Filamentous phage are released from the bacterial membrane by a two-step mechanism involving a short C-terminal fragment of pIII. *J. Mol. Biol.* 289, 1253–1265. doi: 10.1006/jmbi.1999.2851
- Redder, P., Peng, X., Brugger, K., Shah, S. A., Roesch, F., Greve, B., et al. (2009). Four newly isolated fuselloviruses from extreme geothermal environments reveal unusual morphologies and a possible interviral recombination mechanism. *Environ. Microbiol.* 11, 2849–2862. doi: 10.1111/j.1462-2920.2009.02009.x
- Reddy, B. L., and Saier, M. H. Jr. (2013). Topological and phylogenetic analyses of bacterial holin families and superfamilies. *Biochim. Biophys. Acta* 1828, 2654–2671. doi: 10.1016/j.bbamem.2013.07.004
- Rice, G., Tang, L., Stedman, K., Roberto, F., Spuhler, J., Gillitzer, E., et al. (2004). The structure of a thermophilic archaeal virus shows a double-stranded DNA viral capsid type that spans all domains of life. *Proc. Natl. Acad. Sci. U.S.A.* 101, 7716–7720. doi: 10.1073/pnas.0401773101
- Roine, E., and Bamford, D. H. (2012). Lipids of archaeal viruses. *Archaea* 2012, 384919. doi: 10.1155/2012/384919
- Roine, E., Kukkaro, P., Paulin, L., Laurinavicius, S., Domanska, A., Somerharju, P., et al. (2010). New, closely related haloarchaeal viral elements with different nucleic Acid types. *J. Virol.* 84, 3682–3689. doi: 10.1128/jvi.01879-09
- Rossman, J. S., and Lamb, R. A. (2013). Viral membrane scission. *Annu. Rev. Cell Dev. Biol.* 29, 551–569. doi: 10.1146/annurev-cellbio-101011-155838
- Russel, M., and Model, P. (1989). Genetic analysis of the filamentous bacteriophage packaging signal and of the proteins that interact with it. *J. Virol.* 63, 3284–3295.
- Saier, M. H. Jr., and Reddy, B. L. (2015). Holins in bacteria, eukaryotes, and archaea: multifunctional xenologues with potential biotechnological and biomedical applications. *J. Bacteriol.* 197, 7–17. doi: 10.1128/jb.02046-14
- Samson, R. Y., Obita, T., Freund, S. M., Williams, R. L., and Bell, S. D. (2008). A role for the ESCRT system in cell division in archaea. *Science* 322, 1710–1713. doi: 10.1126/science.1165322
- Schleper, C., Kubo, K., and Zillig, W. (1992). The particle SSV1 from the extremely thermophilic archaeon *Sulfolobus* is a virus: demonstration of infectivity and of transfection with viral DNA. *Proc. Natl. Acad. Sci. U.S.A.* 89, 7645–7649. doi: 10.1073/pnas.89.16.7645
- Snyder, J. C., Brumfield, S. K., Kerchner, K. M., Quax, T. E., Prangishvili, D., and Young, M. J. (2013a). Insights into a viral lytic pathway from an archaeal virus-host system. *J. Virol.* 87, 2186–2192. doi: 10.1128/jvi.02956-12
- Snyder, J. C., Samson, R. Y., Brumfield, S. K., Bell, S. D., and Young, M. J. (2013b). Functional interplay between a virus and the ESCRT machinery in archaea. *Proc. Natl. Acad. Sci. U.S.A.* 110, 10783–10787. doi: 10.1073/pnas.1301605110

- Snyder, J. C., Brumfield, S. K., Peng, N., She, Q., and Young, M. J. (2011). *Sulfolobus* turreted icosahedral virus c92 protein responsible for the formation of pyramid-like cellular lysis structures. *J. Virol.* 85, 6287–6292. doi: 10.1128/jvi.00379-11
- Soler, N., Marguet, E., Verbavatz, J. M., and Forterre, P. (2008). Virus-like vesicles and extracellular DNA produced by hyperthermophilic archaea of the order Thermococcales. *Res. Microbiol.* 159, 390–399. doi: 10.1016/j.resmic.2008.04.015
- Stedman, K. M., Deyoung, M., Saha, M., Sherman, M. B., and Morais, M. C. (2015). Structural insights into the architecture of the hyperthermophilic *Fusellovirus* SSV1. *Virology* 474, 105–109. doi: 10.1016/j.virol.2014.10.014
- Suttle, C. A. (2007). Marine viruses—major players in the global ecosystem. *Nat. Rev. Microbiol.* 5, 801–812. doi: 10.1038/nrmicro1750
- Torsvik, T., and Dundas, I. D. (1974). Bacteriophage of *Halobacterium salinarium*. *Nature* 248, 680–681. doi: 10.1038/248680a0
- Typas, A., Banzhaf, M., Gross, C. A., and Vollmer, W. (2012). From the regulation of peptidoglycan synthesis to bacterial growth and morphology. *Nat. Rev. Microbiol.* 10, 123–136. doi: 10.1038/nrmicro2677
- Vaney, M. C., and Rey, F. A. (2011). Class II enveloped viruses. *Cell. Microbiol.* 13, 1451–1459. doi: 10.1111/j.1462-5822.2011.01653.x
- Veesler, D., Ng, T. S., Sendamarai, A. K., Eilers, B. J., Lawrence, C. M., Lok, S. M., et al. (2013). Atomic structure of the 75 MDa extremophile *Sulfolobus* turreted icosahedral virus determined by CryoEM and X-ray crystallography. *Proc. Natl. Acad. Sci. U.S.A.* 110, 5504–5509. doi: 10.1073/pnas.1300601110
- Visweswaran, G. R., Dijkstra, B. W., and Kok, J. (2011). Murein and pseudomurein cell wall binding domains of bacteria and archaea—a comparative view. *Appl. Microbiol. Biotechnol.* 92, 921–928. doi: 10.1007/s00253-011-3637-0
- Votteler, J., and Sundquist, W. I. (2013). Virus budding and the ESCRT pathway. *Cell Host Microbe* 14, 232–241. doi: 10.1016/j.chom.2013.08.012
- Wais, A. C., Kon, M., Macdonald, R. E., and Stollar, B. D. (1975). Salt-dependent bacteriophage infecting *Halobacterium cutirubrum* and *H. halobium*. *Nature* 256, 314–315. doi: 10.1038/256314a0
- Weissenhorn, W., Poudevigne, E., Effantin, G., and Bassereau, P. (2013). How to get out: ssRNA enveloped viruses and membrane fission. *Curr. Opin. Virol.* 3, 159–167. doi: 10.1016/j.coviro.2013.03.011
- White, R., Chiba, S., Pang, T., Dewey, J. S., Savva, C. G., Holzenburg, A., et al. (2011). Holin triggering in real time. *Proc. Natl. Acad. Sci. U.S.A.* 108, 798–803. doi: 10.1073/pnas.1011921108
- Young, R. (2013). Phage lysis: do we have the hole story yet? *Curr. Opin. Microbiol.* 16, 790–797. doi: 10.1016/j.mib.2013.08.008

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

Copyright © 2015 Quemin and Quax. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Fluorescence microscopy visualization of halomucin, a secreted 927 kDa protein surrounding *Haloquadratum walsbyi* cells

Ralf Zenke¹, Susanne von Gronau², Henk Bolhuis³, Manuela Gruska⁴, Friedhelm Pfeiffer² and Dieter Oesterhelt^{2*}

¹ Imaging Facility, Max-Planck-Institute of Biochemistry, Martinsried, Germany, ² Department of Membrane Biochemistry, Max-Planck-Institute of Biochemistry, Martinsried, Germany, ³ Yerseke Marine Microbiology, Royal Netherlands Institute for Sea Research, Yerseke, Netherlands, ⁴ Department of Molecular Structural Biology, Max-Planck-Institute of Biochemistry, Martinsried, Germany

OPEN ACCESS

Edited by:

Sonja-Verena Albers,
University of Freiburg, Germany

Reviewed by:

Carsten Sanders,
Kutztown University, Germany
Reinhard Rachel,
University of Regensburg, Germany

*Correspondence:

Dieter Oesterhelt,
Department of Membrane
Biochemistry, Max-Planck-Institute of
Biochemistry, Am Klopferspitz 18,
82152 Martinsried, Germany
oesterhe@biochem.mpg.de

Specialty section:

This article was submitted to Microbial Physiology and Metabolism, a section of the journal Frontiers in Microbiology

Received: 16 January 2015

Accepted: 13 March 2015

Published: 30 March 2015

Citation:

Zenke R, von Gronau S, Bolhuis H, Gruska M, Pfeiffer F and Oesterhelt D (2015) Fluorescence microscopy visualization of halomucin, a secreted 927 kDa protein surrounding *Haloquadratum walsbyi* cells. *Front. Microbiol.* 6:249. doi: 10.3389/fmicb.2015.00249

At the time of its first publication, halomucin from *Haloquadratum walsbyi* strain HBSQ001 was the largest archaeal protein known (9159 aa). It has a predicted signal sequence, making it likely to be an extracellular or secreted protein. Best BLAST matches were found to be mammalian mucins that protect tissues to dehydration and chemical stress. It was hypothesized that halomucin participates in protection against desiccation by retaining water in a hull around the halophilic organisms that live at the limits of water activity. We visualized *Haloquadratum* cells by staining their intracellular polyhydroxybutyrate granules using Nile Blue. Halomucin was stained by immunofluorescence with antibodies generated against synthetic peptides derived from the halomucin amino acid sequence. Polyhydroxybutyrate stained cells were reconstructed in 3D which highlights not only the highly regular square shape but also the extreme flatness of *Haloquadratum*. Double-staining proves halomucin to be extracellular but to be only loosely associated to cells in agreement with its hypothesized function.

Keywords: halomucin, halophilic archaea, polyhydroxybutyrate, cell shape, *Haloquadratum*, immunofluorescence, protein secretion

Introduction

Extreme hypersaline ecosystems with NaCl concentrations approaching saturation harbor a rich microbial community which is dominated by halophilic archaea of the family Halobacteriaceae (Oren, 2002, 2012). The dominant organism found at the highest salinities is a square gas-vesicle containing archaeon that was first detected by Walsby (1980). In many laboratories attempts have been undertaken to grow these spectacular organisms in pure culture but even in 2002, more than two decades after the initial description, Oren writes that “At present, microbiologists still dream of growing and characterizing the square, flat, halophilic Archaeon, first described in 1980...” (Oren, 2002).

At that time, a pleomorphic halophilic archaeon (strain 801030/1) had been cultivated which included square cells but lacked the characteristic gas vesicles described by Walsby. This isolate is now being known as *Haloarcula quadrata* (Oren et al., 1999; Oren, 2002). It is motile and contains

flagella that form a right-handed helix as previously described in an environmental sample (Alam et al., 1984).

After attempts to grow Walsby's square archaeon had failed for more than 2 decades, two groups independently reported the cultivation of Walsby's square archaeon in 2004 (Bolhuis et al., 2004; Burns et al., 2004). The obtained isolates are so similar in their genetic makeup that they are now recognized as strains from the same species, *Haloquadratum walsbyi* (Burns et al., 2007), a name which highlights the hypersalinity tolerated by the organism, its peculiar shape, and honors the scientist who first described the organism (Walsby, 1980). *Hqr. walsbyi* can grow in media which are not only saturated in salt (3.3 M NaCl) but in addition contain more than 2 M MgCl₂ (Bolhuis et al., 2004). This has been referred to as "life at the limits of water activity" (Bolhuis et al., 2006). Despite these extreme conditions, *Haloquadratum* reaches very high cell densities (Walsby, 1980; Oren, 2002) but its growth is exceedingly slow (Bolhuis et al., 2004; Burns et al., 2004).

The Spanish isolate is strain HBSQ001, the Australian isolate is strain C23 and is the type strain of the organism (Burns et al., 2007). The two strains show many similarities but some strain-specific features were detected, including a different number of the cell surface layers (Burns et al., 2007). The genome sequence proved highly similar (98.6% sequence identity, including intergenic regions, and a complete lack of genome rearrangements) (Dyall-Smith et al., 2011) even though the sampling sites (Spain, Australia) represent near-maximal distance on Earth. This confirmed the close similarity already inferred from ecological data and was interpreted as "limited diversity in a global pond" (Dyall-Smith et al., 2011).

Two fascinating aspects of *Hqr. walsbyi* are its square shape and its extreme flatness (about 0.1–0.5 μm) (Walsby, 1980; Stoeckenius, 1981; Kessel and Cohen, 1982; Bolhuis et al., 2004; Burns et al., 2007). If an organism attempts to increase the surface to volume ratio, it has two options: (i) reduction in size (which is only possible up to certain limits) or (ii) to become flat, thus avoiding a limitation of cell size and leading to square cells upon cell-division (Bolhuis, 2005). *Haloquadratum* cells range in size from 1.5 to 11 μm (Walsby, 1980) but extremely large cells of more than 40 × 40 μm have been observed with no indication of cell division structures (Bolhuis et al., 2004). Tomography images show that *Haloquadratum* has two types of intracellular structures: gas vesicles (e.g., gvpA, HQ_1782A) and polyhydroxybutyrate (PHB) granules. The latter can be visualized with the fluorescent dye Nile Blue A (Ostle and Holt, 1982). The biosynthetic enzymes have been identified in the genome (phaBCE: HQ_2309A-HQ_2311A) (Bolhuis et al., 2006; Dyall-Smith et al., 2011). It were actually the gas vesicles that alerted Walsby of the square cells which otherwise would have been easily overlooked (Walsby, 1980).

One of the highlights found in the genome of the Spanish isolate of *Hqr. walsbyi* was halomucin, the largest archaeal protein known at that time (9159 residues, 927 kDa). Among the best matches in databases were the sequences of vertebrate mucin and it was proposed that halomucin forms a water enriched capsule around the cells. This prediction is based on the presence of a typical Sec-type signal sequence at the N-terminus of the protein. Interestingly, the ortholog from strain C23 (Hqrw_1107)

is significantly shorter (7836 aa) and lacks the two copies of the CTLD (C-type lectin domain) (InterPro:IPR001304) that are found in the protein from strain HBSQ001 (Dyall-Smith et al., 2011).

Mucins form important structures in prokaryotes and eukaryotes including mammals where they are involved in protection of cells or cell tissues to desiccation or chemical stress (Tabak et al., 1982). Mucins also have been implicated in the protection of mammalian tissues against viral attack and pathogenic bacteria and may be applied in purified form as a broad-range antiviral supplement to personal hygiene products, baby formula or lubricants to support the human immune system (Liegels et al., 2012).

Secretion of such a huge protein through the Sec core seems challenging. We wanted to confirm that halomucin is a secreted protein and wanted to investigate the type of association between halomucin and its cells. We addressed this by double-staining where cells were visualized by Nile Blue staining of their cytoplasmic polyhydroxybutyrate granules and halomucin was visualized by immunostaining. We found it to be an extracellular protein which is only loosely attached to cells.

Materials and Methods

For staining of polyhydroxybutyrate granules, a Nile Blue stock solution in ethanol was prepared (1 mg/ml). *Hqr. walsbyi* strain HBSQ001 was grown in HAS medium (with CaCl₂ reduced to 0.3 mM) to stationary phase as described (Bolhuis et al., 2004). 100 μl of cells were mixed with 100 μl HAS medium (Bolhuis et al., 2004) and 6 μl of the Nile Blue stock solution and incubated for at least 1 h. Cells were visualized in a Leica SP2 confocal microscope, excitation (Nile Blue): 633 nm, detection: 645–750 nm.

Polyclonal antibodies against halomucin were generated in chicken to allow subsequent usage under high-salt conditions. Antibodies were generated using synthetic peptides as antigen (Peptide synthesis at MPI of Biochemistry, Core Facility). The synthetic peptides with the unique amino acid sequences, NELSVDTSAQPIDDLSA and RGAAPAWLGVVSG-PAATA, corresponded to position 1458–1474 and 8774–8791 of HQ_1081A, respectively. Polyclonal antibodies were generated by Davids Biotechnologie, Regensburg, Germany. Prior to use, polyclonal antibodies were diluted (1:5, 1:25, 1:100) with HAS medium.

For fluorescent microscopy, a FITC-conjugated rabbit anti-chicken-IgY (Promega #G2691, Promega, Madison, WI) secondary antibody was used. This antibody was diluted (1:6) with HAS medium. 100 μl *Hqr. walsbyi* cells were mixed with 100 μl diluted polyclonal antibody and incubated for 1 h at room temperature. Then, 20 μl of the secondary antibody dilution were added. For double staining, 6 μl of the Nile Blue stock solution were added. Cells were visualized by Leica Confocal Software.

Results and Discussion

Electron tomography imaging of a single square cell of *Hqr. walsbyi* strain HBSQ001 revealed the presence of gas vesicles

(GV) and PHB granules (**Figure 1**, reproduced from Bolhuis et al., 2006). The contours of the cells were visualized by staining the ubiquitously present polyhydroxybutyrate granules with Nile Blue and imaged by fluorescence microscopy (**Figure 2**). This image shows the head-on and side-on views of a single cell, confirming the square shape and the extreme flatness of *Hqr. walsbyi* (Walsby, 1980; Stoeckenius, 1981; Bolhuis, 2005; Burns et al., 2007). **Figures 2A,B** are stills from a short video clips (Video S1) where the 3D reconstructed cells are animated to represent one complete 360° turn of the cell. Several such video clips are available in the Supplementary Material (Videos S1–S4). Once more, these images confirm the extremely unusual cell shape of *Hqr. walsbyi*. Cells from this organism are distinctively square, while those from other species like *Haloarcula quadrata*, which do form quadratic cells, appear overall more irregular (Oren et al., 1999). *Hqr. walsbyi* cells are extremely flat (ca 0.2 μm) whereas the length in the other two dimensions are at least 5 μm. The extreme flatness of the cells may imply the complete absence of cell turgor but the forces that cause cell edges to be as straight as they are observed are currently enigmatic since no genes could be identified in the genome that might express structural proteins involved in maintaining the cell structure. It has been argued that the corners are rather a secondary consequence of the flat morphology (Bolhuis, 2005). In prokaryotic cells many essential processes take place at their surface. These include amongst others nutrient and oxygen uptake, light driven generation of a trans-membrane proton gradient and extrusion of end products. The large flat cells of *Hqr. walsbyi* guarantee the perfect surface to volume ratio and allow cells to become very large without suffering the drawbacks of limited diffusion speed of essential nutrients as in spherical cells.

Although the majority of the cells are distinctively square, sometimes deviations in cell shape can be observed as shown in **Figure 3**. The cell in **Figure 3A** contains round, circular, stain-free regions that appear as holes. Similar “holes” or electron dense regions have been observed before (Stoeckenius, 1981). Currently there is no satisfactory explanation for these structures but these might be caused by either damage of the fragile cells, presence of a high density material like DNA or regions where the opposing cell membranes are in contact, thereby not leaving any space for gas vesicles or PHB granules. In **Figure 3B** the cell appears rectangular rather than quadratic and is most likely an example of an elongating cell before cell division. This cell contains a more extended stain-free region but still seems intact as it retains the PHB granules in the cytoplasm. In **Figure 3C** only three of the four edges form a perfect straight corner whereas the fourth corner has a more “edged” shape. Also these edged parts have been observed previously but may simply express the fragility of these square cells, especially when treated with stains and undergoing washing steps before microscopy.

It has been predicted that halomucin is a secreted protein based on the detection of a Sec-type signal sequence at the N-terminus of the protein. However, a protein that huge may be a challenge for any secretion system and true evidence for the extracellular location of halomucin was lacking. That the electron tomographical image in **Figure 1** does not reveal any features that can be interpreted as extracellular mucin is most likely

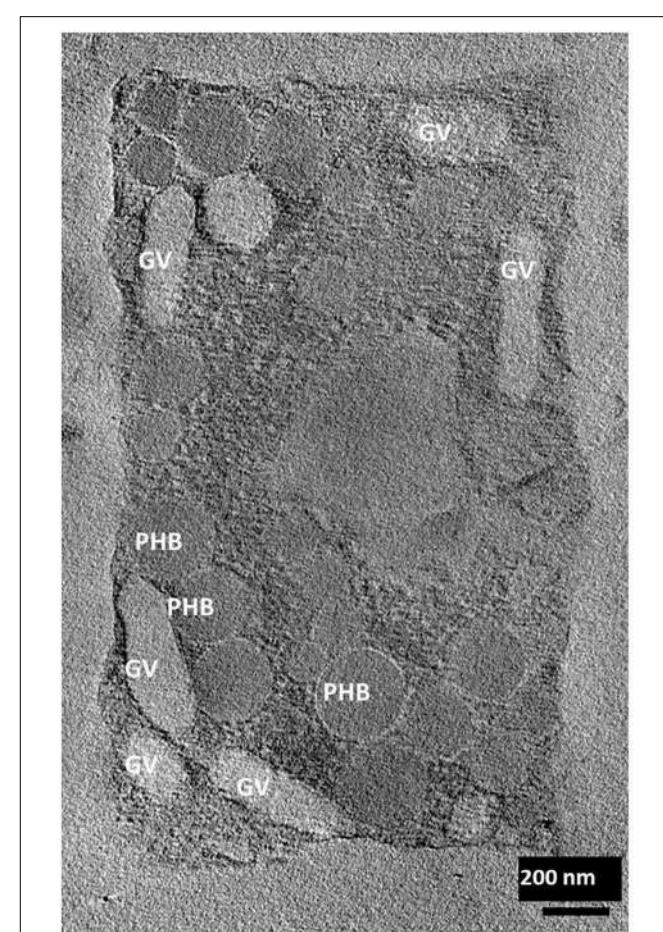


FIGURE 1 | Electron tomographic image of a single square cell of *Hqr. walsbyi* (reproduced from Bolhuis et al., 2006). Gas vesicles (GV) and polyhydroxybutyrate granules (PHB) are indicated (image by M. Gruska while working with H. Engelhardt).

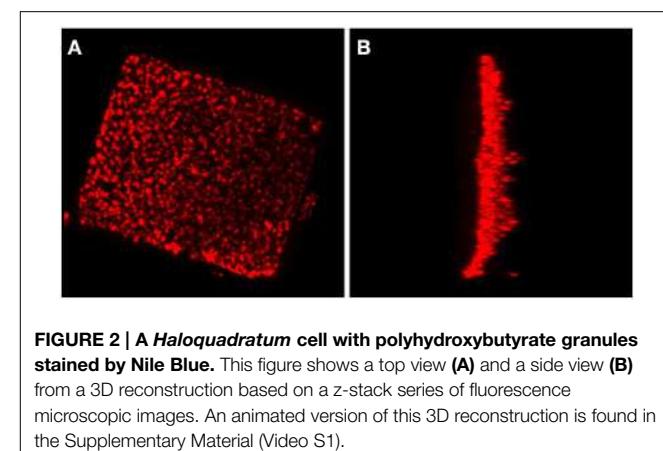


FIGURE 2 | A *Haloquadratum* cell with polyhydroxybutyrate granules stained by Nile Blue. This figure shows a top view (A) and a side view (B) from a 3D reconstruction based on a z-stack series of fluorescence microscopic images. An animated version of this 3D reconstruction is found in the Supplementary Material (Video S1).

caused by the fact that the preparations required for this type of images is often destructive to the extracellular features. Of all the images of the square archaeon that are published since its discovery in 1980, only one electron microscopy image reveals

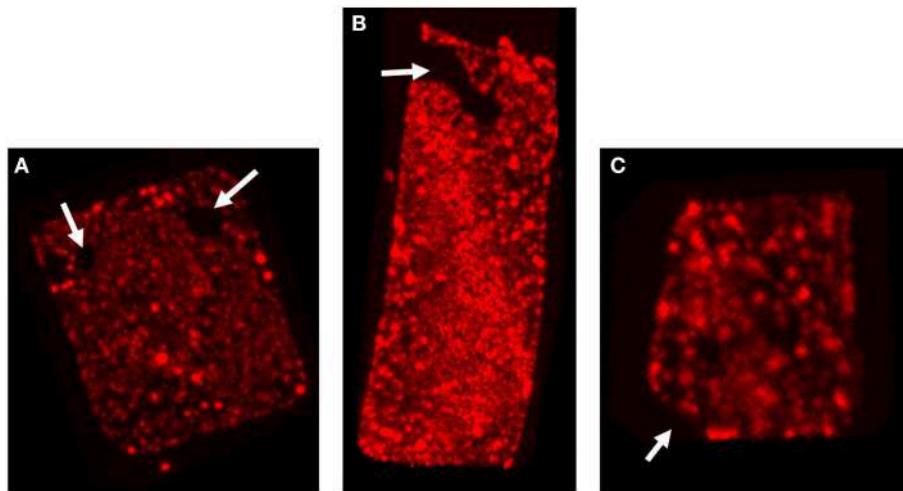


FIGURE 3 | Irregular cell shapes in *Hqr. walsbyi*. The three images show the top view of different Nile Blue stained cells. The white arrows illustrate the different irregularities as discussed in the text: **(A)** “holes” in the cell, **(B)** a rectangular cell, also containing a

more extended stain-free region, and **(C)** a cell where one corner has a more “edged” shape. Each of the images corresponds to one of the animations provided in the Supplementary Material (Videos S2–S4).

an extracellular matrix (Kessel and Cohen, 1982) and only after the identification of halomucin in the genome of *Hqr. walsbyi*, it was speculated that this matrix might be the illusive halomucin (Bolhuis et al., 2004). Using the polyclonal antibodies generated against two synthetic peptides derived from the halomucin sequence we could for the first time successfully identify halomucin (Figure 4). The applied double-staining technique clearly identified the halomucin outside the cells proving that indeed this huge protein is secreted over the cell membrane. As proteins are normally transported in their unfolded state through the Sec pore of the membrane, this apparently is also true for the extremely long halomucin protein. Considering an average translocation rate of 270 amino acid residues per minute and an expense of one ATP per 50 amino acids as determined for the SecEYGA translocation system of *Escherichia coli* (Tomkiewicz et al., 2006), the translocation of one 9159 amino acids long halomucin protein would take about 34 min and require the hydrolysis of about 183 ATP molecules. This would be quite a burden on the protein translocation system but possibly only a few halomucin proteins are sufficient to exert their extracellular function thus preventing blockage of the secretory system. Halomucin was proposed to function as water enriched capsule around the cells although alternative functions such as involvement in generating a defensive barrier against cations or halophages might also be possible. Halomucin as a phage resistance barrier makes sense since phages are ubiquitous in salterns and the genome of *Hqr. walsbyi* shows sufficient evidence of phage related genes and DNA insertions in the past.

There are similarities in amino acid sequence and domain organization to mammalian mucins (Bolhuis et al., 2006) which are known to protect various tissues against desiccation or harsh chemical conditions. Halomucin may be important for *Hqr. walsbyi* to grow at the limits of water activity in a way that is reminiscent of the mucous lungfish cocoon, which enables it to

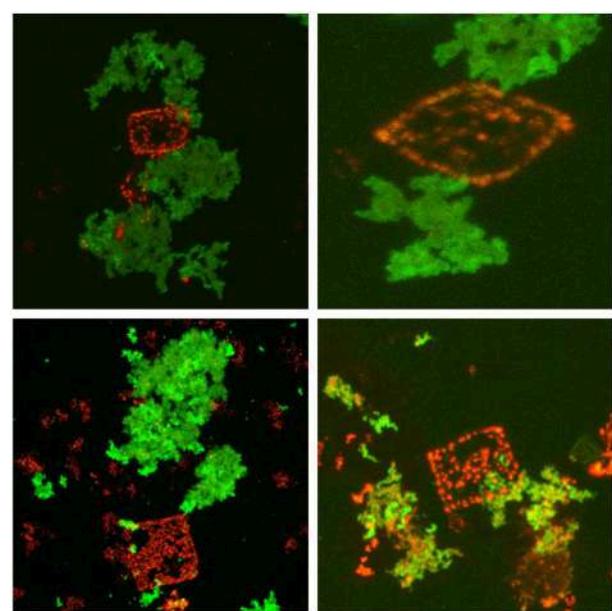


FIGURE 4 | Immunofluorescence stained halomucin loosely associated with the Nile Blue stained *Hqr. walsbyi* cells. Immunofluorescence staining appears green while Nile Blue staining of polyhydroxybutyrate granules appears red and marks *Haloquadratum* cells. Halomucin is clearly an extracellular protein which is only found loosely attached to cells.

survive for months outside of water (Chew et al., 2004). However, the sequence databases have grown considerably since the initial publication of the halomucin sequence. When comparing halomucin to these updated databases, a number of proteins of prokaryotic origin give a higher blast score than the initially identified mammalian mucins (Dyall-Smith et al., 2011).

The halomucin of the two isolated *Haloquadratum* strains show distinct differences, with the protein from strain C23 being significantly shorter (7836 aa) than the ortholog from strain HBSQ001 (9159 aa). This size reduction has been attributed to deletion of two internal gene segments of 2.4 kb (codons 283–1075) and 1.3 kb (codons 5078–5583) (Dyall-Smith et al., 2011). The protein regions selected for antibody production are present in both strains but only one (pos 877–8791) is identical in the two strains while the other (pos 1458–1474) is in a highly variable region (8 of 17 residues differ). A distinct difference is the lack of a pair of CTLD (C-type lectin domain) (InterPro:IPR001304) domains in strain C23, caused by the 2.4 kb deletion. It has been speculated that halomucin is modified by sialic acid (Bolhuis et al., 2006), a post-translational modification which is largely restricted to higher eukaryotic organisms. The corresponding pair of sialic acid biosynthesis proteins (neuAB, HQ_3518A/HQ_3519A) is, however, restricted to strain HBSQ001. The strain-specific occurrence of sialic acid biosynthesis genes and of the halomucin CTLD domains may indicate a functional correlation.

Upon double-staining, halomucin becomes visible in the form of large clusters which seem neither closely associated with the cells nor completely independent. At the point of closest association, halomucin may be connected/bound to the cell surface. It cannot be excluded that detachment of a halomucin cluster from the cell tears part of the cell surface away, leading to the observed “holes.” However, this would imply a distinct capacity of the cell for resealing of the cytoplasmic membrane as PHB granules are retained in the cytoplasm. It is more likely that in the natural environment, which is devoid of strong external disturbances,

this loose association is sufficient to exert its protective functioning.

It should be noted that we applied the antibodies to intact cells without fixation or membrane permeabilization. Thus, an intracellular pool of halomucin, if present, would have escaped detection.

In conclusion, we have experimentally confirmed halomucin to be a secreted protein which is found outside of but only loosely connected to the cell. In addition, we provide a set of images and video films that highlight the spectacular shape of this organism which dominates the most salty ecological niches known to support biological growth on earth.

Author Contributions

DO designed the work. HB isolated and cultivated this strain of the organism and analyzed data. SG performed cell staining. RZ carried out microscopy, image analysis, and 3D reconstruction. MG generated the tomographic image of *Haloquadratum*. FP analyzed data. FP, HB, and DO wrote the manuscript.

Acknowledgments

We thank Harald Engelhardt who was critically involved in obtaining tomographic images under high-salt conditions.

Supplementary Material

The Supplementary Material for this article can be found online at: <http://www.frontiersin.org/journal/10.3389/fmicb.2015.00249/abstract>

References

- Alam, M., Claviez, M., Oesterhelt, D., and Kessel, M. (1984). Flagella and motility behaviour of square bacteria. *EMBO J.* 3, 2899–2903.
- Bolhuis, H. (2005). “Walsby’s square archaeon,” in *Adaptation to Life at High Salt Concentrations in Archaea, Bacteria, and Eukarya*, eds N. Gunde-Cimerman, A. Oren, and A. Plemenitas (Dordrecht: Springer), 185–199.
- Bolhuis, H., Palm, P., Wende, A., Falb, M., Rampp, M., Rodriguez-Valera, F., et al. (2006). The genome of the square archaeon *Haloquadratum walsbyi*: life at the limits of water activity. *BMC Genomics* 7:169. doi: 10.1186/1471-2164-7-169
- Bolhuis, H., te Poele, E. M., and Rodriguez-Valera, F. (2004). Isolation and cultivation of Walsby’s square archaeon. *Environ. Microbiol.* 6, 1287–1291. doi: 10.1111/j.1462-2920.2004.00692.x
- Burns, D. G., Camakaris, H. M., Janssen, P. H., and Dyall-Smith, M. L. (2004). Cultivation of Walsby’s square haloarchaeon. *FEMS Microbiol. Lett.* 238, 469–473. doi: 10.1111/j.1574-6968.2004.tb09790.x
- Burns, D. G., Janssen, P. H., Itoh, T., Kamekura, M., Li, Z., Jensen, G., et al. (2007). *Haloquadratum walsbyi* gen. nov., sp. nov., the square haloarchaeon of Walsby, isolated from saltern crystallizers in Australia and Spain. *Int. J. Syst. Evol. Microbiol.* 57, 387–392. doi: 10.1099/ijss.0.64690-0
- Chew, S. F., Chan, N. K., Loong, A. M., Hiong, K. C., Tam, W. L., and Ip, Y. K. (2004). Nitrogen metabolism in the African lungfish (*Protopterus dolloi*) aestivating in a mucus cocoon on land. *J. Exp. Biol.* 207, 777–786. doi: 10.1242/jeb.00813
- Dyall-Smith, M. L., Pfeiffer, F., Klee, K., Palm, P., Gross, K., Schuster, S. C., et al. (2011). *Haloquadratum walsbyi*: limited diversity in a global pond. *PLoS ONE* 6:e20968. doi: 10.1371/journal.pone.0020968
- Kessel, M., and Cohen, Y. (1982). Ultrastructure of square bacteria from a brine pool in Southern Sinai. *J. Bacteriol.* 150, 851–860.
- Lieleg, O., Lieleg, C., Bloom, J., Buck, C. B., and Ribbeck, K. (2012). Mucin biopolymers as broad-spectrum antiviral agents. *Biomacromolecules* 13, 1724–1732. doi: 10.1021/bm3001292
- Oren, A. (2002). Molecular ecology of extremely halophilic Archaea and Bacteria. *FEMS Microbiol. Ecol.* 39, 1–7. doi: 10.1111/j.1574-6941.2002.tb00900.x
- Oren, A. (2012). Taxonomy of the family Halobacteriaceae: a paradigm for changing concepts in prokaryote systematics. *Int. J. Syst. Evol. Microbiol.* 62, 263–271. doi: 10.1099/ijss.0.038653-0
- Oren, A., Ventosa, A., Gutierrez, M. C., and Kamekura, M. (1999). *Haloarcula quadrata* sp. nov., a square, motile archaeon isolated from a brine pool in Sinai (Egypt). *Int. J. Syst. Bacteriol.* 49, 1149–1155. doi: 10.1099/00207713-49-3-1149
- Ostle, A. G., and Holt, J. G. (1982). Nile Blue A as a fluorescent stain for poly-beta-hydroxybutyrate. *Appl. Environ. Microbiol.* 44, 238–241.
- Stoeckenius, W. (1981). Walsby’s square bacterium: fine structure of an orthogonal prokaryote. *J. Bacteriol.* 148, 352–360.
- Tabak, L. A., Levine, M. J., Mandel, I. D., and Ellison, S. A. (1982). Role of salivary mucins in the protection of the oral cavity. *J. Oral Pathol.* 11, 1–17. doi: 10.1111/j.1600-0714.1982.tb00138.x
- Tomkiewicz, D., Nouwen, N., van Leeuwen, R., Tans, S., and Driessens, A. J. M. (2006). SecA supports a constant rate of preprotein translocation. *J. Biol. Chem.* 281, 15709–15713. doi: 10.1074/jbc.M600205200

Walsby, A. E. (1980). A square bacterium. *Nature* 283, 69–71. doi: 10.1038/283069a0

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

Copyright © 2015 Zenke, von Gronau, Bolhuis, Gruska, Pfeiffer and Oesterhelt. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

ADVANTAGES OF PUBLISHING IN FRONTIERS



FAST PUBLICATION

Average 90 days
from submission
to publication



COLLABORATIVE PEER-REVIEW

Designed to be rigorous –
yet also collaborative, fair and
constructive



RESEARCH NETWORK

Our network
increases readership
for your article



OPEN ACCESS

Articles are free to read,
for greatest visibility



TRANSPARENT

Editors and reviewers
acknowledged by name
on published articles



GLOBAL SPREAD

Six million monthly
page views worldwide



COPYRIGHT TO AUTHORS

No limit to
article distribution
and re-use



IMPACT METRICS

Advanced metrics
track your
article's impact



SUPPORT

By our Swiss-based
editorial team