Philipps Universität Marburg

“Decipering the subunit interaction in the crenarchaeal archaellum”

Dissertation

zur Erlangung des Doktorgrades

der Naturwissenschaften

(Dr. rer. nat.)

dem Fachbereich Biologie

der Philipps-Universität Marburg

vorgelegt

von

MSc.- Biotechnologe

Tomasz Paweł Neiner

Geboren in Krapkowice

Marburg (Lahn), Oktober 2014

Die Untersuchungen zur vorliegenden Arbeit wurden von August 2011 bis August 2014 am Max-Planck-Institut für terrestrische Mikrobiologie unter der Leitung von PD Dr. Sonja-Verena Albers durchgeführt.

Vom Fachbereich Biologie der Philipps-Universität Marburg

als Dissertation angenommen am:

Erstgutachter:	Prof. Dr. Sonja-Verena Albers
Zweitgutachter:	Prof. Dr. Hans-Ulrich Mösch
Drittgutachter:	Prof. Dr. Martin Thanbichler
Viertgutachter:	Prof. Dr. Lars-Oliver Essen

Tag der mündlichen Prüfung: 17.12.2014

Die während der Promotion erzielten Ergebnisse sind zum Teil in folgenden Originalpublikationen veröffentlicht:

Lassak K, **Neiner T**, Ghosh A, Klingl A, Wirth R, Albers SV
Molecular analysis of the crenarchaeal flagellum
Mol. Microbiol. 2012 Jan; 83 (1):110-24

Reindl S, Ghosh A, Williams GJ, Lassak K, **Neiner T**, Henche AL, Albers SV, Tainer JA.
Insights into FlaI functions in archaeal motor assembly and motility from structures, conformations, and genetics.
Mol Cell. 2013 Mar 28;49(6):1069-82

Banerjee A, **Neiner T**, Tripp P, Albers SV
Insights into subunit interactions in the Sulfolobus acidocaldarius archaellum cytoplasmic complex.
FEBS J. 2013 Dec;280(23):6141-9

Neiner T, Banerjee A, D'Imprima E, Reindl S, Ghosh A, Arvai AS, Vonck J, Tainer JA and Albers SV
Structural and functional characterization of FlaH: *an archaellum clock?*
In preparation

Table of Contents

1. Introduction

1.1. Motility

An energy consuming process allowing active movement is defined as motility. This ability to move is an important feature allowing organisms to adapt to incessantly changing environments, which is required for their survival. During million of years of evolution microorganisms developed a variety of different motility mediating mechanisms. Less complex eukaryotes, like protozoa, can produce several structures, which enable them to overcome short distances. They can move using wavy tails called flagella, hair-like structures called cilia, or foot-like structures called pseudopodia (Chubb *et al.*, 2002). However considering the "simpler" physiology of prokaryotic organisms, the diversity of motility mechanisms observed in this domain appears much more impressive. Thus, bacterial motility has been puzzling scientists ever since the first microscopes enabled a deeper look into the microbial universe and nowadays it still stays one of the most intensive studied topics in the field of microbiology.

The most common and, so far, extensively studied bacterial motility organelle is the flagellum, which propels the cells with its long, wavy filament in a rotary fashion (Silverman & Simon, 1974). Flagella are mediating two different types of bacterial motility. They are powering the swimming ability of individual cells in liquid environments (Berg & Anderson, 1973, Macnab, 2003), as also the multi cellular motion across viscous environments or surfaces, termed swarming (Kearns, 2010). Numerous bacteria are capable to move across surfaces through alternating extension and retraction of filaments known as type IV pili, this flagella-independent type of motility is called twitching (Mattick, 2002). An additional type of bacterial motility is gliding, which is mediated by focal-adhesion complexes that bind to the surface and translocate the cell body along the cell axis (Nan & Zusman, 2011). An interesting variation of gliding was described in *Mycoplasma* species, which developed tiny "legs" allowing slow movement in a continuous manner (Miyata, 2008).

Although the past decades of extensive research led to numerous achievements in the understanding of how bacteria move, the analogous topic in Archaea, the second group of prokaryotes, was barely explored. Recently several cell surface structures in archaea have been reported. Moreover one of these structures could be identified as the archaeal flagellum, due to similarities to the bacterial counterpart. This archaeal organelle is also a several cell length long wavy filament that enables swimming via its rotation, analogously to the bacterial flagellum (Shahapure *et al.*, 2014a). However more detailed studies revealed a unique structure, much closer related to the bacterial type IV pilus, than to the flagellum, therefore a new name was postulated and the archaeal flagellum was formally renamed as "Archaellum" (Jarrell & Albers, 2012). Still our knowledge about the archaellum and its assembly mechanisms remains relatively scarce in comparison to these of the motility apparatus of bacteria or eukaryotes.

1.2. Bacterial flagellum

The flagellum is the most common bacterial motility apparatus, widespread in Gram-positive and Gram-negative bacteria, which it is used mostly for swimming in aqueous environments (Berg & Anderson, 1973) and also in some cases for swarming across solid surfaces (Kaiser, 2007). Certain bacteria, like *Proteus mirabilis* or *Escherichia coli* can use the same flagella

system to mediate both types of motility, increasing their number on the cell surface (Merino *et al.*, 2006). On the contrary *Vibrio parahaemolyticus* or *Aeromonas hydrophila* possess a secondary flagella system, which is expressed only in response to growth in viscous environments enforcing swarming motility (Merino et al., 2006). Interestingly, *Pseudomonas* species use the same flagella system for swimming and swarming but they are able to exchange individual components of the flagella motor depending on the type of motility performed (Toutain *et al.*, 2005). The flagellum consists of three substructures: the filament, the hook and the basal body (Fig.1). The filament reaches several cell lengths, is ca. 20 nm in diameter and acts as a propeller. The basal body anchors the flagellum in the cell surface, whereas the hook connects the filament with the basal body. The flagellar basal body of Gram-negative bacteria consists of a rod with several rings: the L (lipopolysaccharide) ring in the outer membrane; the P (peptidoglycan) ring in the surface of the cytoplasmic membrane; the MS (membranous and supra-membranous) ring and the C ring, which is part of the motor (Fig. 1)(Thomas *et al.*, 2006). In Gram-positive bacteria the L and P rings are missing.

Usually the flagellar filaments reach lengths several sizes longer than the cell dimension and are estimated to be composed of around flagellins 20 000 subunits. In *Bacillus subtillis* the complete assembly process takes 60 minutes, of which the filament assembly alone requires approximately 30 minutes. Flagellum assembly requires the incorporation of more than ten flagellin subunits per second which is an extremely energetically expensive process, since it requires the production of several dozen of different proteins that are involved directly either in the assembly or its regulation (Guttenplan *et al.*, 2013). For this reason, it is not surprising that the biosynthesis of this highly complex molecular machine is strictly regulated already at the transcriptional level, so that the structural components are expressed in a hierarchical manner according to the current assembly status (Aldridge & Hughes, 2002, McCarter, 2006). First, in response to environmental factors, the flagellar class I master regulators, FlhC or FlhD, are activated. They in turn lead to the expression of the class II genes, encoding structural components of the hook and basal body. At this stage the remarkably flagellar self-assembly process begins. First, around 26 copies of FliF aggregate to form the MS ring, which is the base and the attachment point for further subcomponents (Chevance & Hughes, 2008). In the next step three proteins: FliG. FliM and FliN are forming the cup-shaped C-ring, which works later as the functional rotor of the flagellum (Fig. 1) (Blair, 2006). The C-ring is also responsible for the control of the direction of the filament rotation by switching between clockwise and counterclockwise rotation under the influence of CheY, a component of the chemotaxis system (Bren & Eisenbach, 1998). Subsequently two transmembrane proteins MotA and MotB, which are also anchored into the peptidoglican layer, build up the stationary part of the flagellum, the so called stator complex. In parallel the type III secretion machinery is formed inside the C-ring. The membrane proteins FlhA, FlhB, FliO, FliP, FliQ and FliR align with the central channel in the MS ring and form the molecular export gate of the type III secretion system (T3SS) (Fig. 1)(Evans *et al.*, 2014). Underneath the gate three components, FliI, FliJ and FliH, build the export ATPase complex. FliI is a homohexameric ATPase, FliJ is proposed to interact with one of the export gate components FlhA and FliH locates this complex in the C-ring through interaction with FliN (Chevance & Hughes, 2008, Evans et al., 2014). After the T3SS is set up, the biosynthesis of the rod can be initiated. Those components are getting refolded by the export ATPase complex and then pressed through the T3SS export gate into a 2 nm wide channel starting in the MS-ring to be finally incorporated on the top of the growing structure. According to this scheme, six subunits of FlgB, FlgC and FlgF respectively are secreted in order to build up the proximal part of the rod, followed by 26 subunits of FlgG forming the distal part of the rod (Chevance & Hughes, 2008). Interestingly, the P and L ring components, FlgI and FlgH respectively, are exported via the Sec

pathway, independently from the flagella associated T3SS (Cohen & Hughes, 2014). After the rod penetrates the outer membrane, reaching the destination length of around 25 nm and the assembly of the P and L - rings is completed, the synthesis of the hook is activated.

The hook is formed by approximately 120 FlgE subunits and reaches a length of around 55 nm, what is under strict control of FliK (Samatey *et al.*, 2004, Evans et al., 2014). During the biosynthesis of the hook-basal body complex, the sigma factor FliA encoded by a class II gene is repressed by the anti-sigma factor FlgM. However as soon as the hook formation is accomplished, FlgM is secreted and FliA can initiate the expression of the class III genes, such as those encoding the flagellins (Aldridge & Hughes, 2002). Once the hook has reached its mature size, the control protein FliK interacts with FlhB and turns off the hook component recognition and switches the export gate specificity toward the flagellins. Subsequently, FlgD, the hook cup protein, dissociates from the top of the structure and is replaced by the filament cup protein FliD, initiating the filament assembly (Yonekura *et al.*, 2000). Before FliD can enter the flagellar export gate, its N-terminus is linked to the C-terminal end of the preceding subunit, which has already entered the export channel (Evans *et al.*, 2013). This clever energy saving system accelerates significantly the filament assembly, omitting the need of additionally energy input to speed up the passive diffusion of single filament subunits. This sophisticated mechanism also assists to keep the biosynthesis of the flagellum at a constant rate, even though the assembly takes place several microns away from the cell surface (Evans et al., 2014).

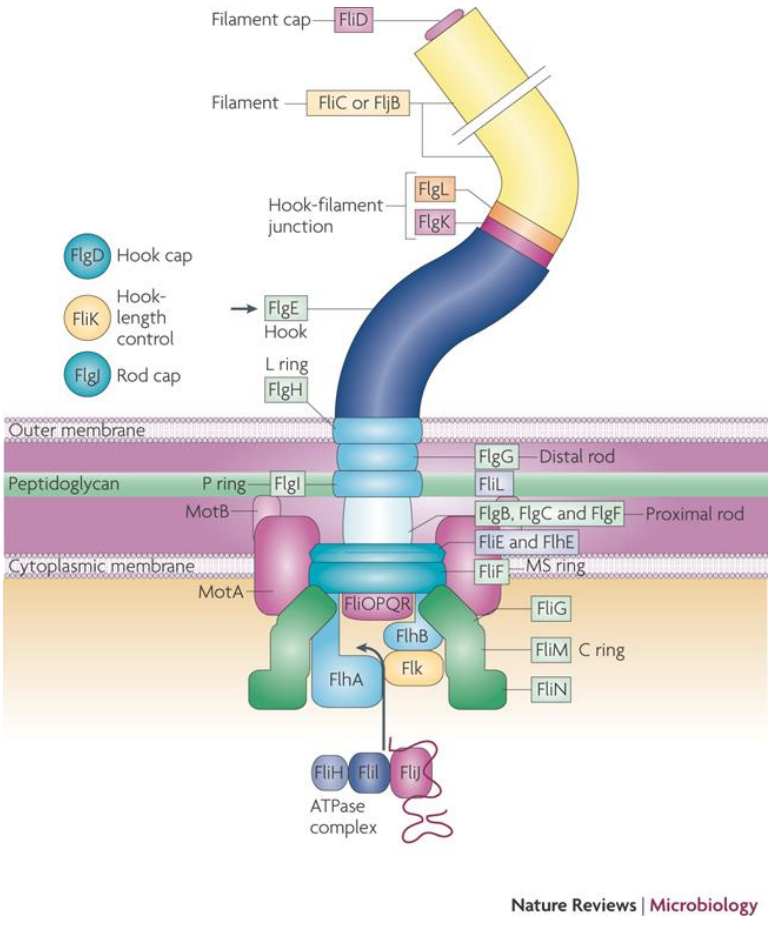


Fig. 1. Model of the structure and assembly of the bacterial flagellum. Reprint from Chevance et al., 2008.

Despite the great complexity of the flagellum, the highly effective motor consists of only five proteins: FliG, FliM and FliN that form the rotor (Yamaguchi *et al.*, 1986) and the membrane proteins MotA and MotB that form the ion-conducting stator complex (Dean *et al.*, 1984, Khan *et al.*, 1988). MotB has a peptidoglycan-binding domain, which fixes the stator complex stably to the peptidoglycan layer. MotA interacts with FliG and, due to conformational changes induced by ion passage, pushes FliG turning the flagellar rotor (Kojima & Blair, 2001, Kim *et al.*, 2008). The nature of the translocated ions impacts on the rotation speed, when driven by proton force the flagella rotate up to 20 000 rpm or even up to 100 000 rpm in case of Na+ driven motors, enabling swimming velocity of 160 μm per second (Magariyama *et al.*, 1994).

Another interesting characteristic is that bacterial species display different numbers and distribution of flagella on the cell surface. These features appear not to be random and preliminary studies indicated that a strict regulatory mechanism is required. So far two major players within this regulatory pathway could be identified: FlhF and FlhG, but their roles remain still an enigma (Bulyha *et al.*, 2011, Kazmierczak & Hendrixson, 2013, Guttenplan et al., 2013).

1.3. Type IV Pili

The type IV pili (T4P) were found to be highly abundant among Gram-negative bacteria therefore they were previously assigned exclusively to this group of organisms. However recently, it has been shown that T4P are also present in Gram-positive bacteria, especially among Clostridia species. For this reason, the T4P is one of the most widespread and universal cell surface structures among the entire domain of bacteria (Varga *et al.*, 2006, Melville & Craig, 2013). Interestingly, T4P related structures are very abundant in Archaea (Pohlschroder *et al.*, 2011). Up to date multiple archaeal surface structures, including pili and the archaeal motility apparatus, the archaellum, have been reported do share homology in terms of architecture and assembly with the T4P of Gram-negative bacteria (Lassak *et al.*, 2012a).

T4P are involved in an astonishing variety of cellular functions, such as formation of microcolonies and biofilm, DNA uptake or cell signaling (O'Toole & Kolter, 1998, Chen & Dubnau, 2004). T4P can have a crucial role in pathogenesis. For example, in the human pathogen *Neisseria gonorrhoeae* T4P facilitate the adhesion of the bacterial cell to the eukaryotic host-cell (Craig *et al.*, 2004). Furthermore, due to the unique ability to undergo cycles of filament extension and retraction, the T4P are also able to drive a flagella independent type of bacterial motility called twitching (Mattick, 2002). In this case the T4P are often located at the cell poles and extend reaching lengths exceeding 10 μm (Mattick, 2002, Jarrell & McBride, 2008b). The T4P filament has usually a diameter of 5-9 nm and is composed of multiple copies of a single pilin protein, PilA (Giltner *et al.*, 2012). The elongated pilus is capable to adhere to surfaces and subsequently pull the cell forward by retraction, generating thereby forces of up to 150 pN, what makes it the strongest molecular motor characterized so far (Clausen *et al.*, 2009).

The pilin protein PilA shows a number of unusual features, such as methylated N-terminus, a conserved hydrophobic N-terminal part consisting of 25 residues and a C-terminal disulphide bond (Craig et al., 2004). The core assembly machinery of the T4P consists of five components (Fig. 2): the previous described pilins; a polytopic membrane protein, which form the anchoring platform for the filament (Crowther *et al.*, 2004); an outer membrane secretin; an energy supplying assembly ATPase and often also a second disassembly ATPase (Jakovljevic *et al.*, 2008).

The type IV pilins are synthesized as preproteins, containing a type III signal peptide, which is subsequently recognized and cleaved by a prepilin peptidase PilD. The N-terminus of the processed pilin forms a long, highly hydrophobic α-helix, which enables them to remain in the membrane before they get incorporated into the filament, while the C-terminus of the pilin resembles a β-sandwich structure. During the assembly, the densely packed α-helices become surrounded by the β-sheet domains, which lead to the incredible stability of the filament. In Gram-negative bacteria, the outer membrane protein PilQ forms large ring-like cylindrical pore, which crosses the lipid bilayer (Collins *et al.*, 2005, Siewering *et al.*, 2014). The T4P assembly mechanism is ATP driven and the ATPases are the best characterized component of this machinery. PilB is involved specifically in pilus assembly, whereas, PilT is responsible for disassembly or retraction of the pilus. They belong to the type II/IV secretion ATPase superfamily and interact with the core polytopic membrane protein PilC (Tripathi & Taylor, 2007, Takhar *et al.*, 2013). ATPases form hexameric rings, which change their conformation in response to ATP hydrolysis (Savvides, 2007). Therefore, in the assembly model PilB can be found as a closed ATP-bond form and an opened ADP-bond form. These changes are accompanied by energy release, which push the pilins approximately 10 Å upwards across the inner membrane into the growing filament, providing space for the next subunit (Craig *et al.*, 2006, Melville & Craig, 2013). The T4P system evolved an interesting pilin recycling mechanism. During the retraction, the pilus is disassembled and the primary components are temporary stored in the membrane to be subsequently reused for the assembly of the next pilus (Skerker & Berg, 2001). This approach is significantly limiting the need to synthesize huge amounts of pilins for the rapid pili assembly.

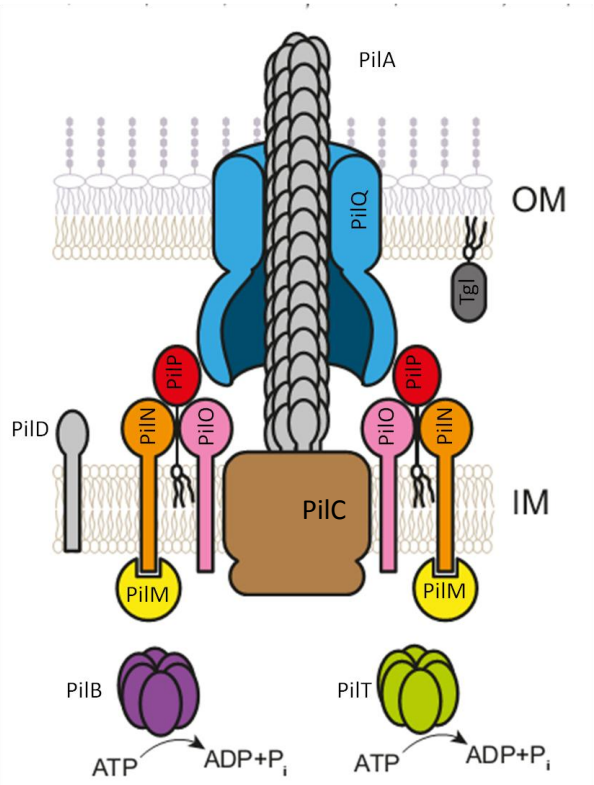


Fig. 2. Model of type IV pili assembly mechanisms. PilA prepilins are processed by the specific prepilin peptidase PilD and incorporated into the base of the growing pilus structure. The energy for the pilin assembly is obtained by the action of the polytopic membrane domain PilC and the assembly ATPase PilB. The pilus filament is extruded through the outer membrane via a ring-like cylindrical pore complex formed by PilQ oligomers. Abbreviations: OM, outer membrane; IM - inner membrane; Reprint from Friedrich et al. 2014.

1.4. Archaea – the third domain of life

Already from ancient times humanity has tried to name and describe its surrounding environment. Thus the birth of taxonomy dates back to the times of ancient Greece and its great philosophers, like Aristotle. He had classified living organisms based on clearly visible features; therefore the initial taxonomy recognized only plants and animals as two major groups of existing life. Microorganisms were discovered scarcely hundreds of years after the great antique civilizations had already fallen, in times where the technological development enabled to overcome the optical limitations of the human eye. Antonie van Leeuwenhoek is considered the father of modern microbiology because he was the first one that observed bacteria using a microscope of his own design (Porter, 1976). Further progress in the field of microscopy led to differentiation between eukaryotes and prokaryotes. Though eukaryotes were over the next century's organized on taxonomical kingdoms and subkingdoms it was constantly believed that prokaryotes includes only one specific group of organism called bacteria. Only with the achievements of the late 20th century this viewpoint could be changed. The discovery of DNA and later development of methods enabling the sequencing and understanding the information pool revolutionized the concept of taxonomy. This modern phylogenetic approach for biological classification resulted in identification of a new group of organisms, named later as Archaea.

In 1960s Carl Woese examined the small ribosomal RNA subunits of different microorganisms in order to analyze their phylogenetic relationships (Pechman & Woese, 1972). Unexpectedly, the 16S rRNA of *Methanobacterium* species revealed a completely different oligonucleotide signature, unrelated to either those obtained from bacteria or from eukaryotes (Fox *et al.*, 1977). Further sequencing analysis of two other thermoacidophilic genera, namely *Thermoplasma* and *Sulfolobus* showed the same unique oligonucleotide signature observed in methanogens (Woese *et al.*, 1978, Woese *et al.*, 1984). At this point, the existence of a third group of organisms differing significantly from the so far known bacteria and eukaryotes became obvious.

In the following years, numerous archaeal strains were identified, which were mostly isolated from unusual environments, like the solfataric hot springs in Japan, Island or Yellowstone National Park, the NaCl supersaturated Dead Sea or geothermally heated marine sediments in Italy (Mullakhanbhai & Larsen, 1975, Brock *et al.*, 1972, Fiala & Stetter, 1986). To survive these extreme conditions Archaea evolved a unique cell membrane. Isoprenoid chains are linked to a *sn*-glicerol moiety via ether binding, whereas eukaryal and bacterial membranes are mostly composed of fatty acid chains ester-linked to glycerol (Kates, 1992, Albers & Meyer, 2011). Although there are some rare examples of ether-linked lipids reported also in bacteria, the *sn1* stereochemistry of the glycerol backbone remains a unique feature of the archaeal membrane (Weijers *et al.*, 2006, Albers & Meyer, 2011). Moreover the membranes of most thermophilicacidophilic archaea are formed by a tetra-ether monolayer, where the ends of the hydrophobic isoprenoid tails are chemically linked with each other (Sprott *et al.*, 1991, Macalady *et al.*, 2004). This tetra-ether based architecture preserves the membrane rigidity at high temperatures and prevents ion permeability, maintaining intracellular neutral pH in often extremely acidic environments (Yamauchi *et al.*, 1993, Macalady et al., 2004).

Another interesting feature of the archaeal cell envelop is the completely lack of structures typical for bacteria, like murein (Kandler & Hippe, 1977, Kandler & Konig, 1978). Only few examples of archaeal pseudomurein have been reported so far, but these structures deviate significantly from the bacterial counterparts (Kandler & Konig, 1998). Instead archaea species developed their own unique cell envelope structure: the S-layer, which based on its simple composition and widespread distribution could even be considered the evolutionary youngest cell

wall structure (Albers & Meyer, 2011). Archaeal S-layers are mostly composed of multiple copies of a single type of highly glycosylated protein, which can oligomerize into two-dimensional crystalline structures. These glyco-protein layers surround the cells, determining their shape and protecting these organisms from adverse environmental agents, while allowing access to all crucial nutrients (Arbing *et al.*, 2012).

Archaea resemble bacteria in terms of morphology, both domains are represented by unicellular organisms lacking the cell nucleus, whose genetic information is encoded on circular chromosomes of similar size (Kletzin, 2007). However certain molecular features in archaea are functionally closer related to eukaryotic equivalents rather than bacterial ones. A good example is the archaeal transcription machinery, which relies on DNA-dependent RNA-polymerase type II-like protein, eukaryotic-like transcription factors and TATA-box-like transcription initiation sites (Huet *et al.*, 1983, Bell & Jackson, 2001, Allers & Mevarech, 2005). On the other hand the transcription regulatory pathway and the subsequent translation process show more similarities to those from bacteria (Dennis, 1997, Bell & Jackson, 1998).

Beside the similarities to Bacteria and Eukarya, genetic analysis revealed a set of several hundred genes unique for Archaea, finally confirming the three domain classification proposed by Woese (Woese *et al.*, 1990, Graham *et al.*, 2000). The list of archaea unique features, including cell wall composition, ribosome structure or motility machineries, grew over the last decades to such an extent, that there is presently no doubt that Archaea form a separate group of organisms, which is related to neither Eukarya nor Bacteria. Moreover the attribution of Archaea to colonize only the extreme habitats also had to be corrected, over the years multiple examples of mesophilic archaea have been found, as bacteria species, in diverse habitats like soil and oceans (Chaban *et al.*, 2006a). Today the domain of Archaea counts already thousands of species and has been divided into five distinct subkingdoms: Euryarchaeota, Korarchaeota, Crenarchaeota, Thaumarchaeota and Nanoarchaeota (Olsen & Woese, 1997, Huber *et al.*, 2002, Elkins *et al.*, 2008, Brochier-Armanet *et al.*, 2008).

1.5. The cell surface structures in Archaea

In the last decades the unique cell envelope composition of Archaea and the associated cell appendages became an attractive topic for research. This increased interest resulted in the identification of a variety of diverse cell surface structures all over the domain of Archaea, among which also entirely unique structures could be discovered. One of this newly unique cell appendages discovered are the cannulae, which have been observed in the *Pyrodictium* genus (Rieger *et al.*, 1995).These hyperthermophiles inhabit hydrothermal marine environments with temperatures between 80 to even 110^{o}C (Stetter *et al.*, 1983). *Pyrodictium* species grow in networks connected with extracellular tubules named cannulae. These tubular structures are 25nm in diameter and show extreme heat resistance (Nickell *et al.*, 2003). So far, three glycoproteins, CanA, CanB, CanC, have been proposed to form these exceptional cell appendages (Nickell et al., 2003). It has been observed, that the cannulae do not enter the cytoplasm of the connected cells, but only the pseudoperiplasmic space (Nickell et al., 2003). The exact function of these structures have not been elucidated yet, but they might have a role in nutrient exchange and adhesion (Ng *et al.*, 2008). Another example of a unique archaeal cell surface structure is the hamus, which was found in the euryarchaeon SM1 isolated from cold (-10^{o}C) sulphidic springs (Rudolph *et al.*, 2001). The SM1 cells are small, 0.6 μm in diameter cocci with approximately 100 hami attached

peritrichously to the cell envelop (Moissl *et al.*, 2005). These filaments have a diameter of around 8 nm, have a length of between 1 and 3 µm and exhibit a highly interesting architecture. Every 46 nm three prickles protrude from the filament, which ends with a tripartite hook facing back to the cell (Moissl et al., 2005). SM1 cells are living in mixed archaeal/bacterial communities and these filamentous hami allow them to interact with their bacterial partners (Moissl *et al.*, 2002).

Many of the widespread and highly conserved archaeal cell surface structures show similarities to the bacterial type IV pili system (Ng et al., 2008, Pohlschroder et al., 2011, Lassak et al., 2012a). Based on the highly conserved class III signal sequence, the bioinformatical tool "FlaFind" was developed to identify potential pilin encoding genes *in situ*. FlaFind search revealed that type III signal containing pilin/archaellin-like proteins are extremely wide spread in Archaea, suggesting a strong preference for the T4P based model in the assembly of many archaeal cell appendages (Szabo *et al.*, 2007a). Thus using this approach multiple structures exhibiting an impressive functional variety could be identified. These T4P related cell appendages are involved in biofilm formation, cell aggregation as also cell motility and are currently the most intensive studied surface structures in Archaea.

1.6. The Type IV Pili in Archaea

In the last years, T4P-like structures have been the focus of many researches. However, the lack of genetic tools has limited those studies to only a few archaea. Until nowadays, the largest advances in this field have been achieved in *Sulfolobus* species and *Methanococcus maripaludis* as model organism for crenarchaeota and euryarchaeota, respectively (Lassak et al., 2012a). *Sulfolobus* species hold three distinct T4P-like systems. All *Sulfolobus* spp. analyzed so far possess UV-inducible pili and archaella (Lassak et al., 2012a). Furthermore, *S. acidocaldarius* has a second T4P related system, the so called Aap-pilus, which can be found solely in this particular organism (Henche *et al.*, 2012a). Interestingly, *Sulfolobales* and the recently reported *Haloferax volcanii* have only one type III signal recognizing peptidase PibD, which is responsible for the cleavage of all the produced pre-pilin or pre-archaellin involved in diverse surface structures (Szabo *et al.*, 2006, Tripepi *et al.*, 2010). In contrast, *M. maripaludis* has two independent peptidases each one devoted to a different system, the pre-archaellin processing FlaK and the pre-pilin peptidase EppA (Ng *et al.*, 2009). The best studied so far T4P related surface structures in Archaea are briefly discussed in the following section.

1.6.1. The Ups-pilus

An interesting example of archaeal adaptation of the T4P model is the Ups-pilus (**U**V inducible type IV **p**ilus of ***S****ulfolobales*). *Sulfolobus* species evolved this T4P related system as part of a defense mechanism against UV stress (van Wolferen *et al.*, 2013). After exposure to UV light *Sulfolobales* produce numerous very short pili, which mediate cellular aggregation, enabling DNA exchange between the clustered cells, representing thereby a unique DNA-repair mechanism (Ajon *et al.*, 2011, Frols *et al.*, 2008). Because the DNA-damaging agent bleomycin could also activate the Ups-pili biosynthesis leading to cellular aggregation, it is believed that not the UV-light alone, but the appearance of DNA double strand breaks is the initiation trigger of this system (Frols et al., 2008). The Ups-pili mediated cellular aggregation is highly specific and occurs only between the same *Sulfolobus* species. This specificity is achieved due to the ability of the pilins to recognize strain-specific S-layer glycan structures (van Wolferen personal

communication). The Ups-pilus is composed of five proteins encoded in a genomic locus: the hypothetical protein with unknown function UpsX; the secretion ATPase UpsE and the polytopic membrane protein UpsF, homologues to bacterial T4Pcomponents PilB and PilC respectively; finally two class III signal containing pilins UpsA/B (van Wolferen et al., 2013).

1.6.2. The Aap-pilus

The **a**rchaeal **a**dhesive **p**ilus can be found exclusively in *S. acidocaldarius*, where it represents the most abundant cell surface structure (Henche et al., 2012a). The Aap-pili are mainly involved in surface adhesion and biofilm formation, wherein they are important for the formation of tower-like structures (Henche *et al.*, 2012b).

Similarly to the Ups-pilus, the Aap-pilus is built of five elementary components encoded in one gene cluster. Beside the T4P referring genes encoding the ATPase AapE, the polytopic membrane protein AapF and two pilins AapA/B, a potential iron-sulfur oxidoreductase AapX is encrypted within this genomic cluster (Henche et al., 2012a). All the five genes are essential for the biosynthesis of the pilus, which despite significant similarities with the bacterial T4P system shows however an entirely different structure (Henche et al., 2012a). The Aap-pili are 11 nm in diameter and display three stranded helices, showing 138^{o} rotation and rise of 5.7 Å per subunit (Henche et al., 2012a). Both pilins are structurally involved in the filament assembly, but in a different extent, AapB is the major pilin, while AapA has only a secondary role (Henche et al., 2012a).

1.6.3. The Archaellum

Archaella are, until now, the most intensively studied cell surface structures in Archaea. These usually 12-14 nm thick whip-like structures are found in all motile archaea (Thomas *et al.*, 2001a, Ng et al., 2008). It has been shown that the archaellum is responsible for the swimming motility in multiple cren- and euryarchaeota species and this movement is generated in a rotary fashion, as in the case of bacterial flagella (Ghosh & Albers, 2011a, Lassak *et al.*, 2012b, Shahapure *et al.*, 2014b). Paradoxically despite the obvious functional similarities between the archaellum and the flagellum, these two motility organelles are structurally completely nonrelated and no homologous components have been found. Surprisingly, the archaellar architecture resembles another bacterial surface structure, the T4P. Indeed both, the archaellum and the T4P present remarkably homologous in terms of structure and assembly (Jarrell & Albers, 2012).

All the archaella encoding genes are usually clustered in one genomic locus (*fla* locus), except for the pre-archaellin peptidases, which shows always a distinct genomic location (Ng *et al.*, 2006). The archaellum gene cluster could be confirmed in an increasing number of crenarchaeota and euryarchaeota species due to the growing amount of archaeal genome sequences available (Desmond *et al.*, 2007). Moreover, archaellum encoding genes have recently also been identified in the newly proposed phylum of Thaumarchaeota (Spang *et al.*, 2012). However, the *fla* locus differ significantly between the archaeal subkingdoms. This genomic locus contains between seven to 13 different genes, where the latter corresponds more to euryarchaeal systems, while crenarchaeota display the most minimalistic known archaellum system (Jarrell & Albers, 2012).

Generally the first genes of the *fla* operon encode the filament subunits, the so called archaellins (*flaA* and/or *flaB*). In euryarchaeal species, like *H. salinarum*, up to five different archellins can be encoded within the *fla* locus, whereas most crenarchaeota carry only a single archaellin gene (Jarrell & Albers, 2012). In *M. voltae* a hook region could be visualized and interestingly a deletion of FlaB3, which is one of the archaellins resulted in assembly of a functional, but hookless filament (Bardy *et al.*, 2002). An interesting phenomenon was found in *H. volcanii*, which has two archaellin encoding genes, *flgA1* and *flgA2*. It was shown using mass spectrometry analysis that FlgA1 is the major, whereas FlgA2 is the minor subunit of the archaellum of *H. volcanii* (Tripepi *et al.*, 2013). Deletion of FlgA1 leads as expected to an abolished archaellation, whereas the deletion of FlgA2 displays surprisingly a hypermotile phenotype, what could never been reported for any other archaellin deletion mutant so far (Tripepi et al., 2013).

The archaellin genes are followed by *flaCDE*, which products are believed to be involved in chemotaxis and might be responsible for switching the rotation direction of the archaellar motor (Schlesner *et al.*, 2009). Interestingly, these genes are completely absent in the subkingdom of crenarchaeota, where also no chemotaxis were ever reported so far, what would be consistent with the proposed function of FlaCDE. Crenarchaea encode FlaX, which is a unique archaellum component found exclusively in this subkingdom (Ghosh & Albers, 2011a). FlaX from *S. acidocaldarius* forms highly stable ring-like structures, with a diameter of around 30 nm (Banerjee *et al.*, 2012a). It has been also demonstrated that FlaX is essential for archaella biosynthesis in *S. acidocaldarius* and its cellular stability depends on the presence of archaellar core components FlaHIJ, which might suggest that these proteins interact together in order to form the archaellar basal body (Lassak et al., 2012b, Banerjee et al., 2012a).

The *flaHIJ* genes are highly conserved in all the archaella systems. FlaH is a RecA-like protein, showing homology to the bacterial circadian clock KaiC. This protein contains a typical Walker A, but an incomplete Walker B motif and both are crucial for ATP-binding. Therefore FlaH is predicted to be an ATPase-like protein with suggested nucleotide dependent regulatory role rather than an ATPase function (Ghosh & Albers, 2011a). FlaI and FlaJ are homologues to the T4P components PilB and PilC, respectively (Fig. 3)(Peabody *et al.*, 2003a). FlaI was classified as a member of the "secretion superfamily ATPase" protein group involved in bacterial type II/IV secretion, type IV pili and archaella assembly (Ghosh *et al.*, 2011a). *In silico* analysis has predicted FlaJ as a polytopic membrane protein with seven to nine transmembrane domains and two positive charged cytoplasmic loops (CytoI and CytoII)(Ghosh & Albers, 2011a). Because of the high conservation and their similarity to the bacterial T4P system, FlaHIJ are predicted to form the core of the archaellum, providing thereby the assembly platform for the other components (Banerjee et al., 2013). Recent studies could show that the three archaellar components, FlaX, FlaH and FlaI, of *S. acidocaldiarius* are strongly interacting with each other and it has been therefore postulated that they might form a triple subcomplex within the archaellum assembly system (Banerjee et al., 2013). Considering the size of the FlaX ring-like structures, it seems reasonable to believe that they might act as a molecular scaffold. Moreover the space left inside the ring would be sufficient for incorporation of the FlaI hexomer, which might subsequently connect this FlaXI complex with the polytypic membrane protein FlaJ. FlaH has been shown to interact with both FlaI and FlaX, therefore it could be placed at the interface between FlaX and FlaI. Thus, these FlaXHI subcomplex in association with FlaJ are postulated to be the motor containing basal body of the archaellum (Banerjee et al., 2013). However FlaX is only present in the subkingdom of crenarchaeota, suggesting that a distinct architecture of the archaellar basal body in euryarchaeta, which might involve the chemotaxis related proteins FlaCDE, must be present. Although the function of FlaH is not known yet, it has been

hypothesized to modulate FlaI, directing the energy from assembly into torque generation (Ghosh & Albers, 2011a).

FlaF and *flaG* are also highly conserved in all known *fla* operons, although in crenarchaeota they appear in an inverted order. Nevertheless, still the functions of FlaF and FlaG remain unknown (Chaban *et al.*, 2007, Lassak et al., 2012b). However both are predicted membrane associated proteins containing each a single transmembrane domain. Furthermore recently obtained preliminary results suggest a mutual interaction between these proteins and that FlaF can recognize and bind the S-layer associated glycan tree (personal communication with Ankan Banerjee and Patrik Tripp). Thus the current archaella assembly model assumes that FlaF and FlaG might form a pore-like structure in the cell wall, enabling the filament elongation.

The knowledge about the structure of the archaellum is limited to two phylogenetically distant organisms, inhabiting completely different ecological niches: the crenarchaeon *S. shibatae* and the euryarchaeon *H salinarum.* Despite the differences between these organisms, their archaella are structurally comparable, but totally distinct from the bacterial flagellum. The inner core of the archaellar filament is densely packed lacking an inner channel, which is contrary to the bacterial flagella system (Cohen-Krausz & Trachtenberg, 2002, Cohen-Krausz & Trachtenberg, 2008). The inner core is formed by the hydrophobic α-helices on the N-terminus of the archaellins, as it is in bacterial T4P. Since the archaellin leader domain is high conserved, this dense package can be a common feature of all archaella systems in the entire third domain of life (Cohen-Krausz & Trachtenberg, 2002, Cohen-Krausz & Trachtenberg, 2008). The C-terminal parts of the archaellins are surrounding the solid core structure in a right-handed 3-start helix fashion (Trachtenberg *et al.*, 2005). The size of the C-terminus of the archaellins differs among different archaeal species, which results in the varying diameters of archaella.

The similarities of the archaellum to its bacterial counterpart are restricted exclusively to the performed functions, therefore it seems very unlikely that archaella assemble in a flagella related fashion. Moreover no homology between structural elements of these parallel systems could ever be observed. Beside the external structural pattern of the filaments both appendages differ also significantly in size, as flagella are usually 20 nm in diameter, whereas archaella are much thinner with diameters between 12 and 14 nm (Bardy *et al.*, 2003, Jarrell & McBride, 2008b). Moreover, archaella filaments are densely packed lacking the central channel, which is a crucial feature of the type III secretion system based assembly model, required for the export and subsequently incorporation of the flagellins at the distal tip of the growing flagella (Chevance & Hughes, 2008, Cohen-Krausz & Trachtenberg, 2008). So far no detailed archaellar basal body structure has been reported, however electron micrographs of isolated archaella demonstrated a knob-like architecture, excluding the ring-based organization typical for the bacterial flagella (Kalmokoff *et al.*, 1988, Southam *et al.*, 1990b, Kupper *et al.*, 1994).

The similarities to the bacterial T4P system, like the conserved N-terminus of the archaellins, dense filament package as also its size and numerous subunit homologies implied, that archaellum and type IV pili might be evolutionary related (Trachtenberg & Cohen-Krausz, 2006, Jarrell & Albers, 2012). The archaellins contain a class III signal peptide typical for the T4P pre-pilins. Furthermore it has been shown, that these archaeal signal sequences can be recognized and the carrying archaellins subsequently processed by a dedicated pre-archaellin peptidase FlaK/PibD (Bardy & Jarrell, 2002, Albers *et al.*, 2003, Szabo et al., 2006). Both FlaK and PibD belong to the family of integral membrane aspartic acid proteases containing a doubled of essential aspartates, whereas the second one is part of an conserved GxHyD motif (Hy-hydrophobic amino acid) (Henche *et al.*, 2014). Interestingly, these enzymes are homologues to the bacterial pre-pilin peptidase PilD. FlaK of *Methanococcus voltae* is processing exclusively pre-archaellins, whereas the typical for *Sulfolobales* PibD shows a wider substrate spectrum and provides a universal

peptidase for all the T4P-like surface structures, including pili and archaella (Albers et al., 2003, Bardy & Jarrell, 2003). Thus, deletion of PibD in *S. acidocaldarius* resulted in a completely "naked" cell, whereas deletion of FlaK in *M. maripaludis* impacts only the archaella, demonstrating the crucial role of the pre-subunits activation for the archaella assembly.

FlaI and FlaJ are further examples of T4P-like assembly model adaptation in archaea. FlaJ is predicted to be the integral membrane protein considered as the archaellar assembly platform (Ghosh & Albers, 2011a). FlaI of *S. acidocaldarius* has been characterized as a functional metal ion dependent ATPase with an optimum pH of 6.5 and temperature of 75^{o}C, which corresponds to the physiological conditions within the cell in the natural habitat (Ghosh et al., 2011a). Furthermore, the positive impact of archaeal lipids on the ATPase activity of FlaI could be demonstrated, suggesting the involvement of the membrane in regulation of the archaellum (Ghosh et al., 2011a). Recently, also the assembly/rotation cycle has been uncoupled and archaellated, but not motile, *in vivo* mutants have been obtained, proving thereby the bifunctional role of FlaI, which generates energy for both, assembly and rotation (Reindl *et al.*, 2013). Moreover the crystal structure of *S. acidocaldarius* FlaI has been solved and ATP-dependent hexamerisation demonstrated (Reindl et al., 2013). Further structural analysis revealed that the crown grooves of the hexameric FlaI are negatively charged, which led to the assumption that these sites might interact with the highly positively charged cytoplasmic loops of FlaJ, forming a stable FlaJ-FlaI core complex (Banerjee et al., 2013).

Another interesting feature of archaella is the characteristic N-linked glycosylation of the archaellins (Voisin *et al.*, 2005). This post-translational modification of the structural subunits was observed in numerous species and seems to be widespread among archaea (Ng et al., 2008). In contrast post-translational N-linked glycan tree attachment has never been reported for the bacterial flagella or for the bacterial T4P. Although glycosylation itself could be confirmed for T4P and in few cases also for the flagellum, the glycan tree showed always O-linked attachment (Logan, 2006).

Summarizing all the available information the assembly of the archaellum might precede like this: 1) The filament subunits are produced as pre-archaellins and are inserted into the membrane. 2) Then the membrane located pre-archaellins undergo N-linked glycosylation and their signal sequences get cleaved by T4P-like archaellin peptidase FlaK/PibD. 3) The polytopic membrane protein FlaJ forms the assembly platform, recruiting FlaH and FlaI into the core complex. 4) FlaF together with FlaG forms a pore-like structure into the S-layer. 5) The processed archaellins are then incorporated at the base of the growing structure, driven most likely by FlaI mediated ATP hydrolysis.

In the last years more attention has been devoted to the regulation of archaella assembly leading to the discovery of a highly complex regulatory network in *S. acidocaldarius*, what given the relatively simple archaellum in this organism, was quite unexpected. However the biosynthesis of motility organelles and locomotion itself are highly energy consuming processes so it seems economically reasonable to restrict them until they are really necessary.

Under laboratory cultivation conditions Aap-pili have been identified as the predominating cell appendages, whereas archaella appearance was rather rare. Significant increase of the archaellation was first observed in cells entering the stationary grow phase, indicating that depletion of certain nutrients might promote the archaella biosynthesis (Szabo et al., 2006, Lassak et al., 2012). Subsequently, a direct impact of the nitrogen sources limitation on the induction of archaella expression could been confirmed (Lassak et al., 2012b). The *fla* operon of *S. acidocaldarius* contains two independent promoters, one located upstream of the archaellin encoding gene *flaB* and a second one preceding *flaX*. Both are highly up-regulated under

starvation conditions (Lassak et al., 2012b). Further studies have identified two transcriptional activators: ArnR and ArnR1. Both contain similar helix-turn-helix motives and both have been proven to bind to the *flaB* promoter (Lassak *et al.*, 2013). *ArnR* and *arnR1* are encoded close to the *fla* operon and the deletion of these genes inhibits archaellation. Parallel to these findings a set of repressors of archaella expression have been identified: the forkhead-associated (FHA) domain containing ArnA and the von Willebrand (vWA) domain containing ArnB (Reimann *et al.*, 2012). Both have been shown to strongly interact with each other and repress archaella expression in a phosphorylation dependent manner (Reimann et al., 2012). It was demonstrated for *S. tokodai* that the ArnA homologue (ST0829) binds directly to the *flaX* promoter *in vitro* (Wang *et al.*, 2010, Duan & He, 2011), what unfortunately could not be reproduced in *S. acidocaldarius.*

Depleting environmental goods triggers archaella biosynthesis also in species belonging to euryarchaeota, such as *M. jannashii* and *M. maripaludis*. Both mentioned species show sensitivity to H_2 and leucine limitations, in consequence of which they attempt to change the habitat using they archaella mediated swimming ability (Mukhopadhyay *et al.*, 2000, Hendrickson *et al.*, 2008, Xia *et al.*, 2009). However, molecular regulators involved in archaella synthesis initiation or repression have never been reported for this subkingdom. Moreover the Arn regulatory factors identified in crenarchaeota are missing in euryarchaeota, suggesting a completely distinct archaellum regulatory network for this subkingdom (Lassak et al., 2013).

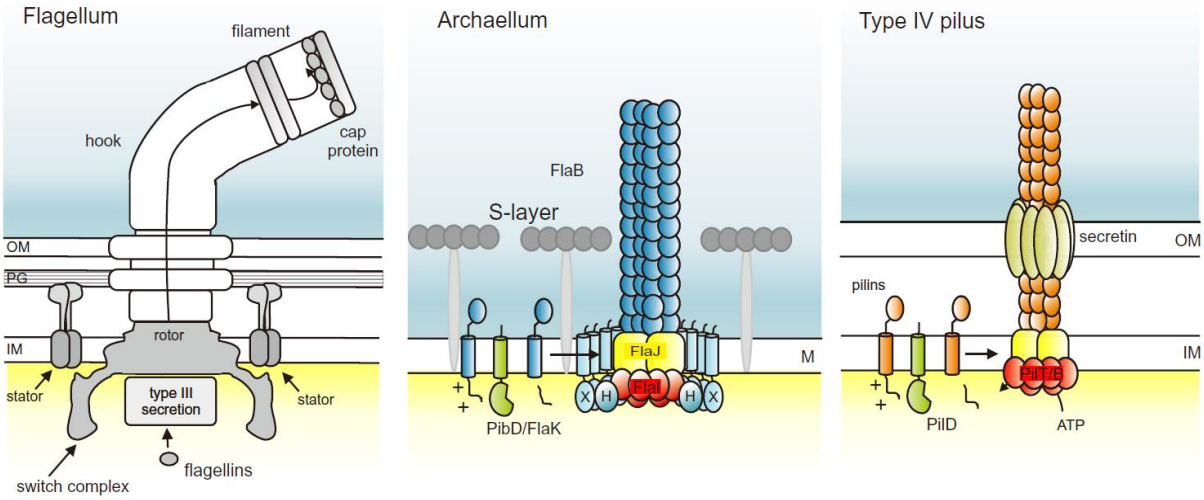


Fig. 3. Schematic comparison of the crenarchaeal archaellum to the flagellum and type IV pili of Gram-negative bacteria. The flagellum is composed of several dozens of different components, which assemble at the distal tip of the growing structure in a type III secretion system based manner. The archaellum and the T4P on the other hand, assemble by insertion of previously processed by PibD/FlaK or PilD archaellins or pilins, respectively, at the base of the structure. In both, the archaellum and T4P two interacting components: the integral membrane protein FlaJ or PilC and the ATPase FlaI or PilB respectively provide the platform and the energy for the assembly of these structures. Recently FlaX and FlaH have been shown to be archaellar motor complex components (Banerjee *et al.*, 2013). Homologues proteins are marked with the same color. Abbreviations: OM, outer membrane; PG, peptidoglycan; IM - inner membrane; M, membrane; Reprint from Shahapure et al., 2014.

1.7. Sulfolobus acidocaldarius as model organism

Sulfolobus acidocaldarius DSM639 was first isolated in 1972 from a volcanic hot-spring at Yellowstone National Park by Thomas Brock (Brock et al., 1972). It was the first characterized representative of the order *Sulfolobales*, belonging to the Creanarchaeota kingdom (Fuchs *et al.*, 1996). *S. acidocaldarius* is a hyperthermoacidophile, which grows optimally at 75°C, pH around 2 and under strictly aerobic conditions (Chen *et al.*, 2005). It is able to grow chemolithoautotrophically oxidizing elemental sulfur, however it can also grow heterotrophically using organic substrates like peptides and sugars (Shivvers & Brock, 1973, Grogan, 1989). The cell shape of S. *acidocaldarius* is defined as highly irregular coccus of around 1 µm in diameter (Mcclure & Wyckoff, 1982).

Nowadays, multiple species are classified as members of the genus *Sulfolobus* and all of them were found in sulfataric hot-springs all over the world. For example, *S. solfataricus* strains P1 and P2 were isolated from a solfataric field in Italy; *S. shibatae* and *S. tokodaii* in Japan; *S.islandicus* in Iceland and *S. yangmingensis* in China (Chaban et al., 2006a).

S. solfataricus and *S. tokodai* are next to *S. acidocaldarius* the best phenotypically characterised *Sulfolobus* species. However, due to its high genomic stability, *S. acidocaldarius* became the preferred model organism for crenarchaeal studies (Grogan *et al.*, 2001). *S. solfataricus* genome contains more than 300 insertion sequence elements (IS) and miniature inverted-repeat transposable elements (MITEs), whereas *S. acidocaldarius* contains only four partial IS-elements, which are not active (Grogan & Hansen, 2003, Brugger *et al.*, 2004). The biggest impediment for genetic manipulation in Archaea in general is the inefficiency of commonly used selection markers like antibiotics. The creation of an uracil auxotrophic strain allowed to overcome these limitations and was proved to be an useful selection marker for gene insertion or deletion in *S. acidocaldarius* (Jonuscheit *et al.*, 2003, Albers & Driessen, 2008, Wagner *et al.*, 2009). Moreover the identification of inducible promoters and the discovery of the not-integrative plasmid pRN1 from *S. islandicus* led to the development of multicopy shuttle vectors, facilitating in-trans gene complementation and overexpression in *S. acidocaldarius* (Berkner *et al.*, 2007, Berkner *et al.*, 2010, Wagner *et al.*, 2012).

The constantly expanding genetic tool box, the availability of fully genome sequence, high genetic stability, easy cultivation and the high diversity of multiple cell appendages make *Sulfolobus acidocaldarius* a perfect model organism to study the role, assembly and regulation of crenarchaeal surface structures.

1.8. The circadian clock protein KaiC

In order to keep pace with the cyclic environmental changes arising from the rotation of the planet around its axis, numerous organisms developed peculiar molecular timing systems. These biological circadian clocks allow them to control their physiological behavior in a progressive day-night cycle, in nearly perfect 24-hours rhythms. Up to date several different types of molecular circadian clocks have been discovered in all domains of life. However the timekeeping mechanisms seem to differ between phylogenetic kingdoms, thus no universal "timer" could be found, suggesting that circadian oscillators evolved repeatedly during evolution (Rosbash, 2009, Edgar *et al.*, 2012).

The simplest and at the same time best studied biological oscillator was found in representatives of Cyanobacteria. Although the life cycle of the model organism *Synechococcus elongatus* counts only six hours, more than 30% of the genes are expressed in a circadian fashion evidencing the importance of developing a hardly infallible molecular clock tracking 24-hours rhythms (Ito *et al.*, 2009).Interestingly the cyanobacterial circadian clock is composed of only three proteins: KaiA, KaiB and KaiC. In this system the time is tracked by the phosphorylation status of KaiC, whereas KaiA and KaiB, accelerate its auto-phosphorylation and auto-dephosphorylation activity, respectively. All the components have already been biochemically and structurally well characterized. However, the biggest advantage is that the cyanobacterial circadian clock can be easily reconstituted *in vitro* by mixing the components in presence of ATP, enabling tracking of cyclic phosphorylation and dephosphorylation events for up to 10 days (Nakajima *et al.*, 2005)

The key player within this system is KaiC, which is composed of two highly similar domains: the N-terminal called CI and the C-terminal CII domain, which is probably a result of internal gene duplication (Dvornyk *et al.*, 2003). Both KaiC domains contain WalkerA and imperfect WalkerB motifs and show homology to the *RecA/Rad51/DCM1* family of ATP-dependent recombinases (Leipe *et al.*, 2000). Structural analysis revealed that KaiC functions as a hexamer, resembling in shape a double doughnut, where a ring formed by six CII domains is placed on a similarly built CI ring (Hayashi *et al.*, 2003, Pattanayek *et al.*, 2004). Despite the high homology, the two KaiC domains differ significantly in the performed function. The CI domain is responsible for the ATP-dependent hexamerisation of KaiC (Hayashi *et al.*, 2004) showing an extremely slow ATPase activity, hydrolyzing only 15 ATP molecules per cycle (Terauchi *et al.*, 2007). The KaiCII domain contains the major phosphorylation sites, Serine 431 and Threonine 432, which undergo cyclic phosphorylation and dephosphorylation in a defined order: S/pT - pS/pT – pS/T (Rust *et al.*, 2007). Though KaiC exhibits ATPase, auto-kinase and auto-phosphatase activities, the circadian rhythms can be fully reconstructed only under the control of the two other components: KaiA and KaiB. First, homo-dimeric KaiA binds specifically to the C-terminal tail of the CII domain, unraveling the so-called A-loop and unlocking the kinase activity of KaiC (t=0h)(Vakonakis & LiWang, 2004, Pattanayek *et al.*, 2006). This leads then to the phosphorylation of Threonine 432 in all CII subunits of the KaiC hexamer in an allosteric manner (t=6h). Followed by the next phosphorylation round where the situation occurs again for Serine 431 resulting in hyper-phosphorylated state in all six KaiCII subunits (t=12h)(Egli *et al.*, 2013). This double phosphorylated pTpS state is privileged by KaiB, which binds as an homo-dimer to the surface of the KaiCII domain forming a ternary KaiABC complex, where KaiA is sequestered and thereby cannot promote KaiC phosphorylation any longer (Pattanayek *et al.*, 2011, Pattanayek *et al.*, 2013). In parallel KaiB stimulates the dephosphorylation of KaiC, which occurs via an unusual ATP-synthesis mechanism and is accompanied by KaiC subunit exchange (t=18h)(Egli *et al.*, 2012, Kageyama *et al.*, 2006, Ito *et al.*, 2007). Also the dephosphorylation of KaiC take place in a strict order, starting with the dephosphorylation of pT432 in all six KaiCII subunits, followed by dephosphorylation of pS431 till the hypo-phosphorylated state of KaiC is restored, ends the cycle (Nishiwaki *et al.*, 2007).

Despite multiple studies performed the complete mechanism of KaiC circadian rhythmicity is still not fully understood. It is so far unclear what exactly enables the 24 hour time keeping, what is the role of the KaiCI ATPase activity and whether the phosphorylation/dephosphorylation events appear as a result or are they a part of the oscillatory mechanism? Nevertheless the stability of the KaiABC oscillator is truly astonishing and makes this system a perfect object for studies.

KaiC homologues are also widespread among Archaea and can be found in almost all species belonging to the third domain of life, including *Sulfolobales* (Dvornyk et al., 2003). The majority

of these KaiC homologues are smaller single domain versions, different than those in Cyanobacteria. The function of these KaiC-like proteins remains still unclear; however given the often very extreme habitats of archaea the need of circadian oscillators is questionable. Moreover the existents of such time keeping system could not be so far ever confirmed in Archaea. Recent studies suggested however, that a set of genes in *H. salinarum* and *H. volcanii* might be regulated in a light/dark dependent manner (Whitehead *et al.*, 2009, Maniscalco *et al.*, 2014). Both examined organism live in habitats near the earth surface, susceptible to the effect of the daily rhythms of the planet. Furthermore both species contain KaiC homologues; however those involvement in the regulation of day/night dependent gene expression could not be demonstrated in those studies (Whitehead et al., 2009, Maniscalco et al., 2014). Considering in addition the lack of any homologues of the two other bacterial circadian clock components KaiA and KaiB in archaea, it appears to be unlikely that these archaeal KaiC-like proteins are involved in any time-keeping functions. Assuming the existence of an archaeal circadian oscillatory system it seems a unique archaeal mechanism rather than the cyanobacterial KaiABC model. Therefore it has been suggested that the highly abundant archaeal KaiC homologues might exhibit regulatory roles, distinct from the bacterial archetype, but perhaps functionally related. Furthermore, recently, it has been suggested that these single domain KaiC homologues may be involved in archaeal DNA repair based on the structural similarities to RadA (Haldenby *et al.*, 2009). It could be also shown that the KaiC homologue Sto0579 of *S. tokodai* can interact with two major DNA repair system players, RadA and the single-strand binding protein SSB and modulate their functions (Sheng *et al.*, 2008). As it becomes more and more obvious that the archael KaiC homologues are not analogues in function it was postulated to remove the misleading link and reclassify all the single RecA domain proteins, currently annotated as KaiC into aRadC (Haldenby et al., 2009).

An interesting example of an archaeal KaiC homologue is FlaH, which as a structural component of the archaellum has obviously nothing to do with either circadian rhythms, or DNA repair. FlaH together with FlaI and FlaJ are believed to form the archaellar core complex, where FlaJ is anchoring it in the membrane and FlaI delivers energy for assembly. However nothing is known about the role of FlaH and whether its function is somehow KaiC related. FlaH shows similarity to the CII domain of KaiC, which is the functional part of this protein, carrying the auto-phosphorylation and auto-dephosphorylation activity. It was long time postulated that FlaH may be involved in modulating the activity of FlaI, which is the only functional ATPase in the archaellum. Therefore, it is tempting to think that this modulation happens on the basis of a KaiC related working model. Answering this questions and elucidating the role of FlaH is one of the main goals of the presented here work.

2. Scope of the thesis

Motility is a common feature of all three domains of life. However in the last decades the main focus has been set on the eukaryotic cilium and the bacterial flagellum, whereas the archaellum remained almost completely unexplored. Single studies limited to phylogenetic, genetic and physiologycal analyses have been performed almost exclusively on representatives of Euryarchaeota, while the assembly, architecture and biophysical aspects of the archaellum have never been investigated. Interestingly the archaellar gene clusters in Crenarchaeota show a simplified genomic organization compared to Euryarchaeota. The "*fla*" operons in Crenarchaeota lack the chemotaxis involved *flaCDE* genes and show a limited archaellin number, what makes them privileged in terms of archaella structural studies. Since *S. acidocaldarius* has an archaellum involving merely seven different components and offers at the same time an advanced genetic tool box it seemed to be a perfect model organism for studying the crenarchaeal archaellum, what is the main objective of the presented here work.

A number of genetic studies already confirmed the essentiality of the *fla* genes for the archaellar assembly and/or functionality in euryarchaeal model organisms (Patenge *et al.*, 2001, Chaban et al., 2007). In the chapter 3.1 analogously work testing the impact of single *fla* genes deletions on the archaella assembly and function in *S.acidocaldarius* is presented. Moreover analysis concerning the genetic arrangement of the *fla* operon and the physiological conditions inducing archaella biosynthesis are also shown in this chapter.

From seven Fla proteins of *S. acidocaldarius* only two, FlaI and FlaH, do not possess any transmembrane domain, therefore both are from the beginning predicted as archaeal basal body components. While FlaH resembles a single RecA domain with defected Walker B motif, FlaI appears to be the only functional ATPase within the archaellum system. FlaI shows a two domain architecture, including well conserved Walker A/B motifs, typical for ATPases powering the bacterial type II secretion system and the assembly of type IV pili. In chapter 3.2 a detailed structural analysis of FlaI was performed, showing a unique cross-subunit interaction based hexamerisation pattern and its the bifunctional mode of action, as it is involved in archaellar assembly and rotation. Recent studies revealed that another component of the crenarchaeal archaellum FlaX, forms oligomeric ring like structures, indicating a possible role as stator complex for FlaX within this system (Banerjee et al., 2012a). Section 3.3 shows that in *S. acidocaldarius* both the archaellar cytoplasmic components FlaH and FlaI can interact with FlaX with high affinities. Moreover a stable ternary FlaXc-FlaH-FlaI complex could be confirmed *in vitro*, suggesting that FlaX may act as a membrane bound cytoplasmic scaffold during the crenarchaeal archaellum assembly process. Chapter 3.4 is devoted to the structural analysis of FlaH, demonstrating strong affinities of this protein towards nucleotides, showing additionally the involvement of both Walker motifs in ATP binding, which turned out to be crucial for the archaellation of *S. acidocaldarius* cells. Furthermore in chapter 3.4 attempts are made to clarify the involvement of FlaH in the archaella assembly and function. The intriguing homology to bacterial circadian clock protein KaiC might imply a phosphorylation dependent regulatory role for FlaH. Indeed a time dependent autophosphorylation activity *in vitro* has been demonstrated for FlaH of *S. acidocaldarius*.

3. Results

The following part is composed of four independent studies, all focusing on different research objectives to expand of our knowledge about the archaeal motility organelle, the archaellum. All manuscripts presented here are either published (3.1-3.3) or prepared for submission (3.4). Each chapter is preceded with a short summary and a statement highlighting the individual contributions of all involved co-authors.

3.1 Molecular analysis of the crenarchaeal archaellum

Summary:

Motility is an inseparable attribute of life, which has been already intensively studied in Bacteria and Eukarya, but was quite neglected in Archaea, where the current knowledge is limited to phylogenetic, genetic and physiological analyses of a few euryarchaeal species. However, nothing was known about motility in the second large subkingdom of Archaea, the crenarchaeota. Using *S. acidocaldarius* as model organism, we have performed the first detailed study about crenarchaeal motility. In this work we conclusively linked the genes clustered in the *fla* operon of this thermoacidophile with its archaeal flagellum, recently renamed as the archaellum (Jarrell & Albers, 2012). Constructing in-frame deletion mutants of all seven *fla* genes of *S. acidocaldarius* we have demonstrated, that all of them are essential for the assembly of archaella. Since the archaellum is not the only surface structure of *S. acidocaldarius*, we have introduced the same *fla* deletions in a non-piliated Δ*aapF* strain. Phenotypical analysis of these Δ*aapF*/Δ*fla* mutants confirmed the relevance of *fla* genes for the archaella assembly, since all the double deletion strains revealed to be non-archaellated and non-piliated. Moreover we have defined physiological conditions inducing the archaellum biosynthesis by monitoring the FlaB expression levels with immunoblot analysis under different conditions. As result of this investigation, we have determined that the limitation of nitrogen sources has a direct impact on the activation of archaella production, as we have observed increased FlaB expression after transferring the cells to tryptone lacking medium. Using RT-PCR we have showed that the *fla* locus is divided into two transcriptional units, where *flaB* and *flaX-flaJ* are transcribed separately. Combining q-RT-PCR and reporter gene assay we have confirmed two independent promoter within the *fla* operon, one located upstream of *flaB* and a second one upstream of *flaX*. Both promoters are getting activated in response to starvation conditions, although the transcription level of *flaB* is significantly higher than this of *flaX-flaJ*. FlaB is the archaellin, which is needed in a much higher number in relation to the other archaellar components in order to build up the several cell length filament.

Moreover, the Δ*aapF*/Δ*fla* deletion strains have been examined for their swimming behavior using semi-solid plate assay and a high-temperature phase-contrast microscope. All the double deletion strains, which were shown to be not archaellated, revealed to be also immotile, delivering the final prove that the archaellum is responsible for the swimming motility of crenarchaeota.

During the phenotypical analysis we have observed that the non-piliated background strain Δ*aapF* shows significantly stronger archaella expression in response to starvation than the wild type MW001, resulting also in higher swimming velocities. This phenomenon cannot be explained so far, may however indicate a crosstalk between the regulatory systems of the different cell surface appendage systems of *S. acidocaldarius.*

Contributions:

Construction of deletion strains, qRT-PCR and RT-PCR analysis, reporter gene assays, investigation of physiological condition promoting archaella expression and thermo microscopy analysis were performed by Kerstin Lassak. Tomasz Neiner constructed double deletion strains and complementation strains, performed cellular fractionation analysis, established and performed motility assay of *Sulfolobus* strains on semi-solid Brock plates. Andreas Klingl and Abhrajoti Ghosh have done the Electron Microscopy analysis. Reinhard Wirth provided Thermo Microscopy facilities and supervised swimming motility analysis. Kerstin Lassak, Abhrajyoti Ghosh and Sonja-Verena Albers wrote the manuscript, which was revised by all the authors.

Molecular Microbiology (2012) 83(1), 110–124

doi:10.1111/j.1365-2958.2011.07916.x
First published online 20 November 2011

Molecular analysis of the crenarchaeal flagellum

Kerstin Lassak,[1] Tomasz Neiner,[1] Abhrajyoti Ghosh,[1] Andreas Klingl,[2] Reinhard Wirth[3] and Sonja-Verena Albers[1]*

[1]*Molecular Biology of Archaea, Max Planck Institute for terrestrial Microbiology, Karl-von-Frisch-Strasse 10, 35043 Marburg, Germany.*
[2]*Cell Biology and LOEWE Research Centre for Synthetic Microbiology, Philipps University Marburg, Karl-von-Frisch Strasse 8, 35032 Marburg, Germany.*
[3]*Department of Microbiology & Archaea Centre, University of Regensburg, Universitätsstrasse 31, 93053 Regensburg, Germany.*

Summary

The ability to move towards favourable conditions provides fundamental advantages to organisms. Interestingly, flagella as motility structures evolved independently in the bacterial and the archaeal kingdom. Whereas bacterial flagella have been intensively studied, our knowledge regarding the archaeal counterpart is mostly restricted to *Euryarchaeota* rather than crenarchaeal flagella. We therefore investigated the flagellar assembly system of the crenarchaeal model organism *Sulfolobus acidocaldarius in vivo*. Promoter studies and qRT-PCR analyses of the flagella gene cluster provided evidence that the expression of the *fla* genes was induced by tryptone starvation. Moreover, we confirmed presence of a secondary *fla* promoter within the *flaB* gene that regulates the transcription of downstream genes *flaX-J*. Markerless in-frame deletions for all *fla* genes encoded in the *fla* gene cluster were constructed. Western blot analysis of all *fla* deletion strains suggested hierarchical protein interactions during the archaeal flagella assembly. Moreover, functional analysis by thermomicroscopy revealed non-motile cells for each of the mutant strains. Electron micrographs demonstrated that lack of motility coincided with the loss of flagellar assembly. Thus we demonstrated that all seven *fla* genes are essential for crenarchaeal flagellum assembly and function.

Accepted 30 October, 2011. *For correspondence. E-mail albers@mpi-marburg.mpg.de; Tel. (+49) 6421178426; Fax (+49) 6421178429.

Introduction

Swimming motility is one of the common characteristics of prokaryotes encompassing the domains of Bacteria and Archaea. In general, this movement is powered by a nano-macromolecular complex on the cell surface, called 'flagellum' (Kearns, 2010). Precise spatial and temporal regulation of flagellar gene expression with respect to differing environmental conditions is crucial for correct flagella assembly and hence for survival (Adams *et al.*, 1985; Ryan and Shapiro, 2003; Anderson *et al.*, 2010).

Present knowledge of flagellum assembly components and their function has been derived mostly from studies of bacterial species whereas understanding of the archaeal flagellum is still in its infancy. Over the last two decades, researchers have described archaeal flagella (Weiss, 1973; Jarrell and McBride, 2008) and many new insights into the archaeal flagella assembly system have emerged in recent years. The most striking finding revealed that the archaeal flagella assembly system shows homology to the bacterial type IV pili assembly system (Albers and Pohlschröder, 2009; Ghosh and Albers, 2011), which is required for twitching motility, cell–cell interaction and pathogenesis in Gram-negative bacteria (Craig and Li, 2008). This fact has renewed interest in understanding structure, assembly mechanisms and evolutionary relevance of archaeal flagella systems in comparison with bacterial macromolecular surface structures. In addition, advancement of archaeal genetic tools has further facilitated studying flagellum assembly components and their requirements in flagellation. However, our current knowledge is mostly restricted to the subkingdom of *Euryarchaeota*. Studies on euryarchaea, such as *Halobacterium salinarum*, *Methanocaldococcus jannaschii*, the three *Methanococci M. voltae*, *M. maripaludis* and *M. thermolithotropicus* and comparatively little on the crenarchaeon *Sulfolobus solfataricus*, have revealed several unique features of archaeal flagella (Patenge *et al.*, 2001; Thomas and Jarrell, 2001; Thomas *et al.*, 2002; Chaban *et al.*, 2007a; Szabó *et al.*, 2007a).

Although archaeal flagella are functionally analogous to those present in Bacteria, in many aspects both surface structures are different (Ng *et al.*, 2006). Typically, the archaeal flagellum is thinner than its bacterial counterpart with 10–14 nm in Archaea (Koval and Jarrell, 1989) compared with approximately 20 nm in Bacteria (Jones and

Table 1. Strains and plasmids used in this study.

Strain or plasmid	Genotype	Source/Reference
Strains		
DH5α	*Escherichia coli* K-12 cloning strain	Gibco
DSM 639	Wild-type *Sulfolobus acidocaldarius*	DSMZ
MW001	DSM 639 Δ*pyrE*	M. Wagner and S.V. Albers, pers. comm.
MW013	MW001 ΔSaci_1178 (Δ*flaB*)	This work
MW026	MW001 ΔSaci_1178 by introducing stop codons (*flaB*:taa)[a]	This work
MW014	MW001 ΔSaci_1177 (Δ*flaX*)	This work
MW015	MW001 ΔSaci_1176 (Δ*flaG*)	This work
MW016	MW001 ΔSaci_1175 (Δ*flaF*)	This work
MW017	MW001 ΔSaci_1174 (Δ*flaH*)	This work
MW018	MW001 ΔSaci_1173 (Δ*flaI*)	This work
MW019	MW001 ΔSaci_1172 (Δ*flaJ*)	This work
MW156	MW001 ΔSaci_2318 (Δ*aapF*)[b]	Henche *et al.* (2011)
MW451	MW001 ΔSaci_2318 ΔSaci_1178 (Δ*aapF flaB*:taa)	This work
MW452	MW001 ΔSaci_2318 ΔSaci_1177 (Δ*aapF* Δ*flaX*)	This work
MW453	MW001 ΔSaci_2318 ΔSaci_1176 (Δ*aapF* Δ*flaG*)	This work
MW454	MW001 ΔSaci_2318 ΔSaci_1175 (Δ*aapF* Δ*flaF*)	This work
MW455	MW001 ΔSaci_2318 ΔSaci_1174 (Δ*aapF* Δ*flaH*)	This work
MW456	MW001 ΔSaci_2318 ΔSaci_1173 (Δ*aapF* Δ*flaI*)	This work
MW151	MW001 ΔSaci_2318 ΔSaci_1172 (Δ*aapF* Δ*flaJ*)	This work
Plasmids		
pSVA406	Gene targeting plasmid, pGEM-T Easy backbone, *pyrEF* cassette of *S. solfataricus*[c]	M. Wagner and S.V. Albers, pers. comm.
pΔ2*pyrEF*	Gene targeting plasmid, pBluescript backbone, *pyrEF* cassette of *S. solfataricus*	Ghosh and Albers (2011)
pSVA1613	In-frame deletion of *flaB* cloned into pSVA406 with ApaI, PstI	This work
pSVA1622	*flaB* with stop codons cloned into pSVA406 with ApaI, PstI	This work
pSVA1602	In-frame deletion of *flaX* cloned into pSVA406 with ApaI, PstI	This work
pSVA1604	In-frame deletion of *flaG* cloned into pSVA406 with ApaI, PstI	This work
pSVA1603	In-frame deletion of *flaF* cloned into pSVA406 with ApaI, PstI	This work
pSVA1605	In-frame deletion of *flaH* cloned into pSVA406 with ApaI, PstI	This work
pSVA1606	In-frame deletion of *flaI* cloned into pSVA406 with ApaI, PstI	This work
pSVA179	In-frame deletion of *flaJ* cloned into pΔ2*pyrEF* with PstI, BamHI	This work
pSVA180	In-frame deletion of *aapF* cloned into pΔ2*pyrEF* with PstI, BamHI	This work
pCMal*lacS*	pRN1-based shuttle vector with *lacS* reporter gene of *S. solfataricus*[d]	Berkner *et al.* (2010)
pSVA1614	pC*lacS*, promoterless vector for negative control in promoter activity assay	This work
pSVA1600	*flaB* promoter replacing Mal promoter, cloned into pCMal*lacS* with SacII, NcoI	This work
pSVA1601	*flaX* promoter replacing Mal promoter, cloned into pCMal*lacS* with SacII, NcoI	This work
pSVA2101	*flaB* ORF cloned into pCMal*lacS* with NcoI, EagI	This work
pSVA2102	*flaX* ORF cloned into pCMal*lacS* with NcoI, EagI	This work
pSVA2103	*flaG* ORF cloned into pCMal*lacS* with NcoI, EagI	This work
pSVA2104	*flaF* ORF cloned into pCMal*lacS* with NcoI, EagI	This work
pSVA2105	*flaH* ORF cloned into pCMal*lacS* with NcoI, EagI	This work
pSVA351	*flaI* ORF cloned into pCMal*lacS* with NcoI, EagI	This work
pSVA2106	*flaJ* ORF cloned into pCMal*lacS with* NcoI, EagI	This work

a. *flaB*:taa: *flaB* deletion strain constructed by replacing nine base pairs of *flaB* coding sequence with three stop codons.
b. *aapF* encodes the membrane protein of the Aap pilus assembly machinery.
c. *pyrEF* encode the enzymes orotate phosphoribosyltransferase and orotidine-5-mono-phosphate decarboxylase respectively.
d. *lacS*: encodes β-D-galactosidase of *S. solfataricus*.

Aizawa, 1991). Moreover, the flagellins of both domains share no sequence homology. Archaeal flagellins are synthesized as type IV pilin-like preproteins, whereas bacterial flagellins are secreted via a type III secretion system (Faguy *et al.*, 1994). Before assembly preflagellins are processed by the preflagellin peptidase PibD or FlaK, which resembles the prepilin peptidase PilD of the bacterial type IV pilus assembly system (Bardy and Jarrell, 2002; 2003; Szabó *et al.*, 2007b). The mechanism of force generation also differs for bacterial and archaeal flagella. Rotation of the bacterial flagellar filament is driven by proton motive force, rather than ATP hydrolysis that powers both euryarchaeal flagella and eukaryotic flagella (Summers and Gibbons, 1971; Larsen *et al.*, 1974; Manson *et al.*, 1977; Streif *et al.*, 2008).

Flagellum assembly in all sequenced, flagellated Archaea is driven by the *fla* locus, which encodes 7–13 flagella related genes (Ghosh and Albers, 2011). Genetic analysis of euryarchaea, e.g. of *H. salinarum* and *M. maripaludis*, has revealed that all these flagellar proteins are essential for flagellation (Patenge *et al.*, 2001; Thomas *et al.*, 2001; 2002; Chaban *et al.*, 2007a).

The *fla* gene cluster from Sulfolobales encodes seven genes (Ghosh and Albers, 2011). In contrast to all other sequenced flagellated Archaea, *Sulfolobus* spp. exhibit only one flagellin. The flagellin gene *flaB* (Saci_1178) is located upstream of a set of flagella accessory genes *flaX, flaG, flaF, flaH, flaI* and *flaJ* (Saci_1177–1172). In addition, the gene *pibD* encoding the preflagellin peptidase is located at a distal position in the chromosome.

The *fla* genes *flaH*, *flaI* and *flaJ* are present in the genome of all sequenced, flagellated Archaea and are believed to be involved in the export of the flagellins. FlaJ is predicted to have 7–9 transmembrane domains. This polytopic membrane protein shows homology to GspE/PilC, the inner membrane components in type II secretion and type IV pilus assembly systems (Peabody *et al.*, 2003). FlaJ has been shown to be essential for flagellation in both Euryarchaea and *S. solfataricus* (Patenge *et al.*, 2001; Chaban *et al.*, 2007a; Szabó *et al.*, 2007a). FlaI of *Sulfolobus acidocaldarius* shows ATP dependent oligomerization *in vitro*, indicating its possible role in flagella assembly similar to PilB involved in type IV pili assembly (Ghosh *et al.*, 2011). FlaH shows homology to secretion ATPase family members, but has an incomplete ATP-binding motif and no motif for ATP hydrolysis was detected. It has been postulated that FlaH interacts with FlaI and modulates its ATPase activity (Jarrell *et al.*, 2007).

For both Eury- and Crenarchaea it has been proposed that FlaI, FlaH and FlaJ generate a transmembrane secretory platform, which allows both secretion of flagellins across the membrane for flagella assembly and interaction with other components in the assembly process (Thomas *et al.*, 2001; Ghosh and Albers, 2011). However, exact functions of all these components in archaeal flagellum assembly remain elusive.

An open intriguing question concerns the mechanism how archaeal flagellar movement results in motility. In general, motility allows organisms to adapt to changing environmental conditions.

Numerous bacterial flagella have been elucidated as rotating filaments by dark-field light microscopy and fluorescence labelling (Berg and Anderson, 1973; Turner *et al.*, 2000; Scharf, 2002; Schmitt, 2003), whereas the mode of action for archaeal flagella is still poorly understood. For only two euryarchaeal species – *H. salinarum* and *Haloquadratum walsbyi* – it was demonstrated that clockwise rotating flagella lead to forward movement whereas counterclockwise rotation of flagella results in backward movement (Weiss, 1973; Alam *et al.*, 1984). Swimming has been shown for the euryarchaeon *Pyrococcus furiosus* (Näther *et al.*, 2006). Motile cells on semi-solid plates were observed for both cren- and euryarchaea like *S. solfataricus* (Szabó *et al.*, 2007a) and *M. voltae* (Jarrell *et al.*, 1996). Moreover, it has been shown that *S. acidocaldarius* escapes from lethal hot spots with temperatures above 90°C (Lewus and Ford, 1999). Thus temperature sensitive motility is one unique survival mechanism of this crenarchaeon. However, so far it has never been proven for *S. acidocaldarius* that motility is flagella mediated. Therefore, we have generated systematic in-frame deletions of each of the *fla* genes in *S. acidocaldarius*. By combining EM analysis and thermomicroscopic analysis we have demonstrated that all *fla* genes are required for flagella assembly and motility. Additionally, we could show that peptide starvation induces expression of *fla* genes, which were under control of two promoters of the *fla* gene cluster.

Results and discussion

Transcriptional analysis of the fla *locus of* S. acidocaldarius

Genes of the *fla* locus of *S. acidocaldarius* include seven genes Saci_1178–1172 called *flaB*, *flaX*, *flaG*, *flaF*, *flaH*, *flaI* and *flaJ*. As all genes are oriented in the same direction, these genes may form an operon (Fig. 1A).

Interestingly, the genes encoding *flaB* and *flaX* are separated by an intergenic region of 75 bp encompassing a TATA box as well as a B recognition element (BRE), suggesting the presence of two distinct transcriptional units. This is further supported by deep sequencing analysis predicting a secondary transcription start in the *flaB* open reading frame (ORF) of *S. solfataricus* (Wurtzel *et al.*, 2010).

To test, whether there is a transcriptional unit including *flaB*, we mapped the putative operon. RT-PCRs using primers spanning the intergenic regions between the *fla* genes as well the adjacent genes (Saci_1179; Saci_1171) were performed (Fig. 1B). To confirm mRNA presence and cDNA integrity internal regions were also amplified. As expected, Saci_1179 and Saci_1171 do not belong to the putative *fla* operon as no PCR products for the intergenic regions were obtained (Fig. 1B, lanes c). By contrast, PCR products could be amplified for *flaX*, *flaF*, *flaG*, *flaH*, *flaI*, *flaJ* and moreover for *flaB*, indicating an operon structure including all seven *fla* genes of *S. acidocaldarius*.

However, this approach cannot rule out the presence of a weak terminator region downstream of the *flaB* ORF. Therefore, in a parallel approach we performed qualitative RT-PCR on cDNA synthesized with *flaJ*, *flaG*, *flaH* and *flaX* specific primers, respectively. If there are two separate transcriptional units, we would expect PCR products only for *flaX*, *flaG*, *flaF*, *flaH*, *flaI* and *flaJ*, but not for *flaB*. Indeed, whereas PCR products could be amplified for all six *fla* genes *flaX-flaJ* (Fig. 1C), no PCR product could be obtained for *flaB*. In contrast, we could amplify a product

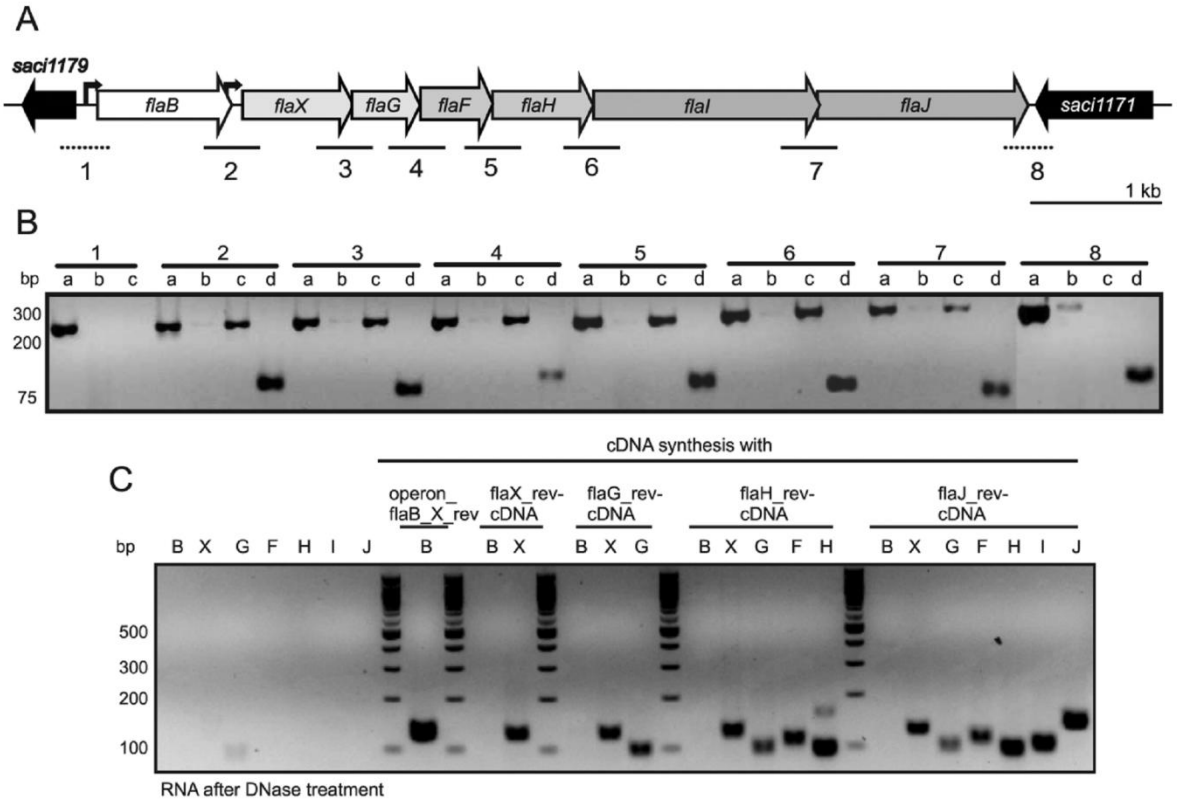


Fig. 1. Transcriptional analysis of the *fla* gene cluster of *S. acidocaldarius*.
A. Gene organization of the *fla* gene locus. Genes are shown as arrows. Black, bent arrows represent confirmed promoter elements. Black bars below the gene locus reflect specific RT-PCR products shown in 1B. Dashed lines illustrate the regions in which no RT-PCR products were obtained.
B. RT-PCR analysis of the *fla* gene locus. Expected PCR products 1–8 correspond to the numbering in Fig. 1A. Genomic DNA as positive control and RNA treated with DNase as negative control were loaded in lanes (a) and (b) respectively. Integrity of cDNA was tested with primers amplifying regions within the *fla*-ORFs shown in lane (d). RT-PCR products loaded in lane (c) indicate the transcription of the *fla* genes as an operon.
C. Transcript mapping of the *fla* gene locus. For first strand cDNA synthesis *fla* gene specific primers indicated above the gel were used. Presence of *fla* transcripts was checked with primers amplifying regions within the *fla*-ORFs *flaB* (B), *flaX* (X), *flaG* (G), *flaF* (F), *flaH* (H), *flaI* (I) and *flaJ* (J) respectively. As negative controls, PCRs using RNA after DNase treatment as template were performed (lanes 1–7).

for *flaB* after cDNA synthesis using another, 5′ located *flaX* primer, which was previously used to amplify the intergenic *flaB-flaX* region. This is in line with the data obtained for the operon mapping. We conclude that an independent transcriptional unit for *flaB* exists, which is not strictly terminated at the 3′ end of the *flaB* ORF. Transcriptional readthroughs in such an operon presumably facilitate a precise regulation of the ratio of the different flagellar components (Thomas and Jarrell, 2001). Thereby the gene encoding the structural subunit is located at the 5′ end of the *fla* gene cluster and is highly transcribed, whereas 3′ end located genes encode less transcribed flagellar core components, which are needed in small amounts. Especially for methanogenic Archaea multiple polycistronic mRNA transcripts were observed within the *fla* locus. All of them are originated from a promoter located upstream of *flaB1* – the first gene of the *fla* locus (Thomas and Jarrell, 2001; Thomas *et al*., 2002).

Construction of deletion mutants of all genes of the S. acidocaldarius fla *gene cluster*

To investigate the importance of the Fla proteins for flagella assembly and function, markerless in-frame deletion mutants of all seven *fla* genes were constructed (Table 1) in *S. acidocaldarius* MW001 (M. Wagner and S.V. Albers, per. comm.). The mutant genotypes were confirmed by PCR and sequence analysis using primers encompassing the flanking regions of the target gene (Fig. S1A). Moreover, as the genetic organization of the *fla* genes might

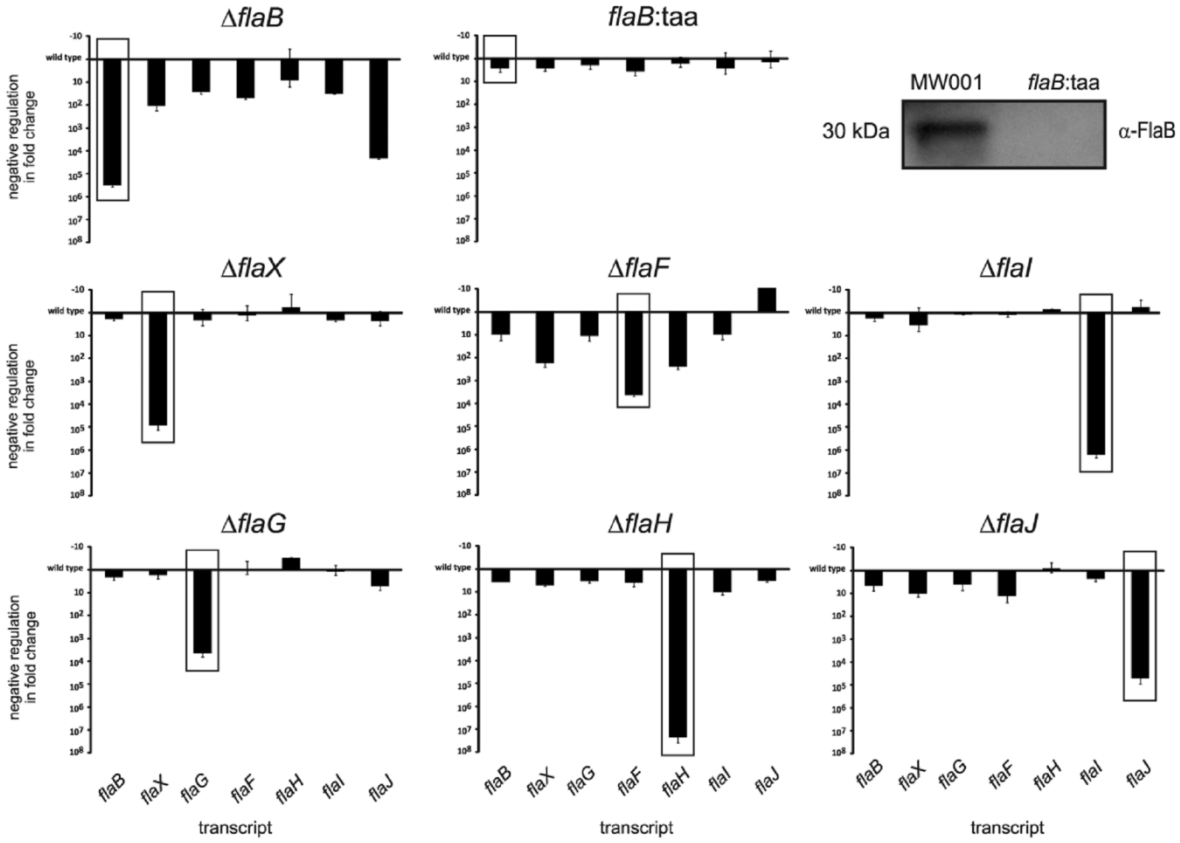


Fig. 2. qRT-PCR analysis of the effects of *fla* in-frame deletions on neighbouring gene expression. Total RNA isolated from Δ*fla* deletion mutants was harvested and corresponding cDNAs were subjected to qRT-PCR analysis using primers specific for indicated *fla* genes (shown underneath the plot). Relative transcript expression levels of each target were normalized to an internal control gene *secY*. The values reflect the fold change in expression compared with MW001, which is designated as baseline. Each bar marked with a box represents the deleted gene of the corresponding mutant. The means and standard deviations of biological replicates are shown. Abolishment of FlaB production for *flaB*:taa was confirmed by Western blot analysis of membrane fractions from strains MW001 and *flaB*:taa.

result in polar effects on downstream genes, qRT-PCR analysis of the mutant strains was performed and compared with the wild type (Fig. 2). Whereas no polar effects were found in Δ*flaX*, Δ*flaG*, Δ*flaH*, Δ*flaI* and Δ*flaJ*, deletion of *flaF* partially influenced transcription of downstream genes. Only Δ*flaB* resulted in a dramatic decrease in downstream gene expression. Therefore another strategy was used to generate a new *flaB* deletion strain, termed *flaB*:taa. In this strain, three stop codons were introduced after 30, 36 and 45 bp of the start codon of the gene leaving the remainder of the *flaB* sequence intact. This strain *flaB*:taa demonstrated no polar effects and restored wild-type transcription levels for all seven *fla* genes including *flaB* (Fig. 2). Western blot analysis of the membrane fraction of *flaB*:taa ensured, that translation of FlaB was indeed abolished. Thus, successful generation of the seven *fla* deletion strains allowed further investigation of flagella assembly and function.

Physiological characterization of flagella biosynthesis

To test, whether the annotated flagellin protein FlaB of *S. acidocaldarius* actually polymerizes and thereby forms the crenarchaeal flagellum, surface structures from sheared cells of MW001 were isolated. In this study, the deletion strain *flaB*:taa was used as a negative control. Due to the fact that it is difficult to distinguish between flagella (10–14 nm in diameter) and pili (8–10 nm in diameter), a deletion strain ΔSaci_2318, called Δ*aapF*, was also included. AapF encodes the membrane protein of the Aap pilus assembly machinery and is essential for pilus biogenesis (Henche *et al.*, 2011).

After density gradient centrifugation (Fig. 3A) pure flagella could be visualized in the Δ*aapF* strain by EM analysis (Fig. 3C–F), whereas for MW001 predominantly pili were detected (Fig. S2). Moreover, so far uncharacterized filaments with a diameter of 5 nm, called threads, were

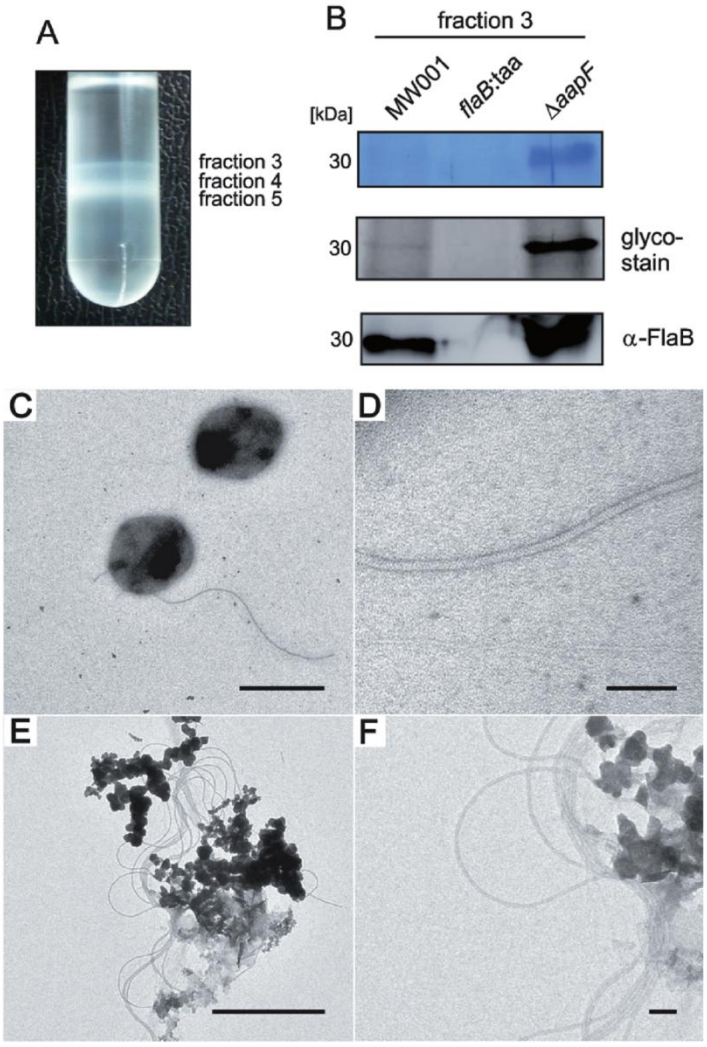


Fig. 3. Characterization of purified flagella of MW001 and *flaB*:taa and Δ*aapF* deletion strains.
A. Cells of Δ*aapF* were grown at 75°C in Brock basal medium supplemented with 0.1% (w/v) tryptone before CsCl (0.5 mg ml^{-1}) gradient centrifugation. An enriched flagella fraction 3 was used for further analysis.
B. Coomassie Brilliant Blue R-250 staining (first row) and a carbohydrate stain (second row) revealed a glycosylated protein for MW001 and Δ*aapF*. It was identified as FlaB of *S. acidocaldarius* by immunoblotting using FlaB specific antibodies (third row).
C–F. TEM analysis of negatively stained flagella of Δ*aapF* proved the identity of the flagellum preparation showing the expected wavy pattern. (D) shows, besides flagella, another type of filament with a diameter of ~5 nm, called thread. Size bars represent 1 μm (C, E) or 100 nm (D, F).

observed in Δ*aapF* (Fig. 3D). SDS-PAGE of the purified fraction 3 of flagella extractions revealed a protein band at the size of flagellin monomer (predicted molecular weight of 29 kD) for both MW001 and Δ*aapF,* but was absent in case of *flaB*:taa (Fig. 3B). Mass spectrometry and Western blotting identified the band and the corresponding signal as flagellin FlaB (Fig. 3B, third panel). Thus, monitoring FlaB after purification of flagella via density gradient is a good indicator for filament assembly.

Surprisingly, we observed an approximately 20-fold increase in FlaB signal intensity for the Δ*aapF* deletion strain in comparison with MW001. This led us to hypothesize, that Δ*aapF* is not only non-piliated, but also hyperflagellated. In fact, further TEM analysis of Δ*aapF* revealed both an increased number of flagella per cell and a higher amount of flagellated cells overall (Fig. S2).

Very recently, an altered migration behaviour of FlaB on SDS-PAGE was observed in a deletion mutant unable to produce UDP-sulfoquinovose (Meyer *et al.*, 2011). Latter sugar is part of the N-glycan decorating the S-layer protein and the cytochrome $b_{558/566}$ of *S. acidocaldarius* (Zähringer *et al.*, 2000; Peyfoon *et al.*, 2010). This led to the assumption that the mature FlaB carries the same glycan composition. Indeed, by performing a glycostain we received a signal at the corresponding size of the flagellin monomer (Fig. 3B, second panel). In the case of euryarchaeal flagellins, N-linked glycosylation is also reportedly widespread (Thomas *et al.*, 2001; Näther *et al.*, 2006).

EM analysis of exponentially growing *S. acidocaldarius* MW001 cells revealed few flagellated cells. Due to the possibility that flagella biosynthesis is growth phase dependent as it has been observed for other microorgan-

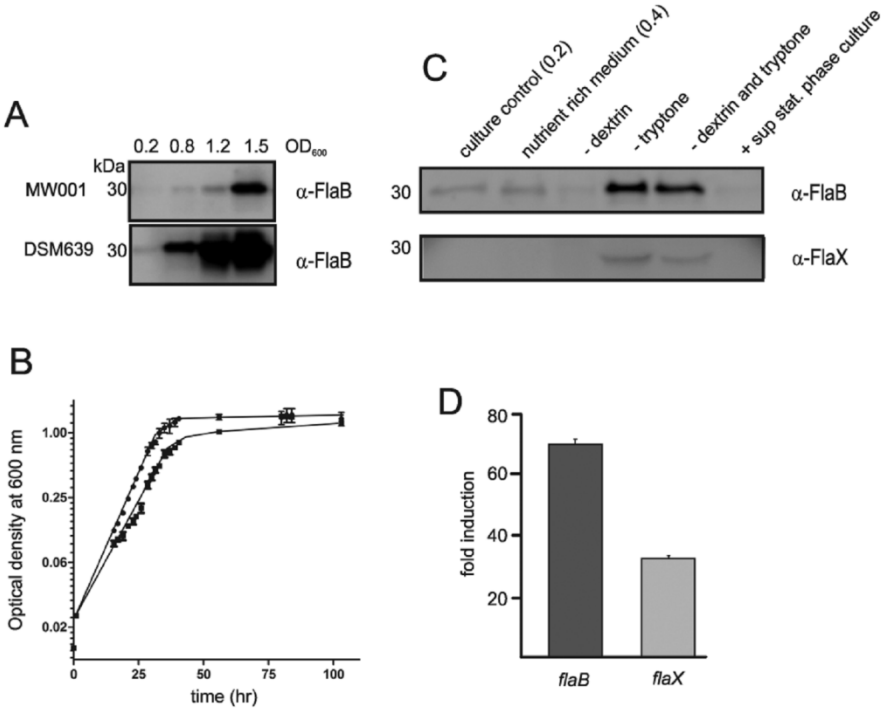


Fig. 4. Physiological regulation of flagellar components.
A. Whole cell lysates of both *S. acidocaldarius* DSM 639 and MW001 were harvested at different growth phases and analysed by immunoblotting using anti-FlaB antibodies. Equal amounts of cells were loaded into each lane in all experiments.
B. Growth of *S. acidocaldarius* DSM 639 (circles) and MW001 (rectangles) was measured at an optical density of 600 nm. Growth curves represent means of three independent experiments.
C. Whole cell lysates of MW001 were prepared before and 4 h after shifting cells into either rich medium or different minimal media or after adding a supernatant of a stationary phase culture. Immunoblotting was performed with antibodies directed against FlaB or FlaX.
D. RNA was isolated from MW001 cells after 4 h incubation in either full medium or tryptone depleted medium. Transcriptional regulation of *flaB* and *flaX* in starved cells was determined by qRT-PCR in comparison with cells grown in full medium.

isms (Heuner *et al.*, 1999; Maurer *et al.*, 2005), two wild-type strains *S. acidocaldarius* DSM 639 and MW001 were grown in Brock medium and samples were taken at different stages of the growth. Whole cell lysates of both strains were analysed for FlaB accumulation by immunoblotting (Fig. 4A). The doubling times of MW001 and DSM 639 did not differ significantly (Fig. 4B). However, MW001 seems to have minor growth disadvantages, which might be caused by the *pyrE* deletion. In pre-exponentially grown cells (Fig. 4A, $OD_{600} = 0.2$) very low expression levels of FlaB could be detected. In contrast, samples taken from exponential (Fig. 4A, $OD_{600} = 0.8$) to stationary phase cultures (Fig. 4A, $OD_{600} = 1.5$) revealed a robust expression of FlaB as determined by immunoblot. Thus, for MW001 the signal intensity increased approximately 20-fold at $OD_{600} = 1.5$ compared with $OD_{600} = 0.8$.

To determine whether induction of FlaB production during entrance into stationary phase was induced by a secreted quorum sensing signal, a pre-exponentially growing MW001 culture was resuspended in spent medium of a late stationary culture ($OD_{600} = 2.8$). However, the FlaB signal intensity remained unaffected in Western blot (Fig. 4C, last lane). As a putative quorum sensing signal was not an inducer, we tested whether nutrient limitation might be a trigger. Therefore we cultivated MW001 cells to pre-exponential phase ($OD_{600} = 0.2$) and then transferred the culture in medium lacking either the nitrogen source tryptone and/or the carbon source dextrin respectively. Interestingly, only tryptone depletion led to a *c.* 50-fold increase in signal intensity for FlaB compared with cells grown with tryptone (Fig. 4C, lane 4). To examine whether FlaB induction after starvation also occurs in other proteins of the flagella assembly system, we also included FlaX in our study. As already observed for FlaB, immunoblotting revealed FlaX accumulation again only in case of tryptone starvation (Fig. 4C, lane 4).

To exclude a general stress response, further stresses such as increased salinity and pH were also tested (data not shown). However, in this study only starvation was found to induce flagellar biogenesis (Fig. 4C). Previously,

temperature dependent motility of *S. acidocaldarius* was reported (Lewus and Ford, 1999). Thus we assumed that flagella are necessary for *S. acidocaldarius* to escape from nutrient limited environments and high temperature hot spots. To further test if FlaB and FlaX induction is a result of transcriptional regulation, we measured also *flaB* and *flaX* expression by qRT-PCRs. Total RNA was isolated from cells that were tryptone depleted for 4 h. A 70-fold and a 30-fold induction of *flaB* and *flaX* transcription were observed respectively. This result clearly showed that induction of flagella biogenesis upon starvation is regulated at the transcriptional level (Fig. 4D).

In a previous section, we showed two distinct transcriptional units of the *fla* gene cluster and a transcriptional readthrough for the *flaB* ORF, presumably due to a weak terminator sequence. Taking this into account, there are two possible explanations for the starvation dependent expression pattern. On the one hand, two putative promoters preceding *flaB* and *flaX* could be regulated independently from each other. On the other hand, the observed induction of *flaX* expression could be a result only of the induced *flaB* promoter. Especially latter idea would allow an adjustment of the optimal ratio of the different Fla proteins. As a large amount of energy is required to synthesize flagella, a strictly regulated flagella biogenesis might provide survival advantages.

Moreover, it was surprising that the flagella production was 20-fold induced in the pili deletion mutant Δ*aapF* in comparison to MW001. This suggests a regulatory interplay between expressions of different surface appendages of *S. acidocaldarius*. Further analyses will shed light on signal transduction pathways and their cross-talk regulating filament assembly with respect to changing environmental conditions. In this context, a recent study by Duan and Guo He unveiled that the regulatory FHA domain-containing protein ST0829 binds the *flaX* promoter of *Sulfolobus tokodaii* in a phosphorylation dependent manner (Duan and He, 2011). It therefore seems that the crenarchaeal flagellar assembly system is subject to complex regulation rather than to a simple on–off switch mechanism.

Cellular analysis of flagellar assembly in Δfla *strains*

For euryarchaeal flagella assembly systems, proteins homologous to FlaB, FlaF, FlaG, FlaH, FlaI and FlaJ were found to be essential for flagella biosynthesis (Patenge *et al.*, 2001; Chaban *et al.*, 2007b). To test whether all proteins encoded by the *fla* gene cluster of *S. acidocaldarius* are also involved in flagellar assembly, electron microscopic analyses of all seven *fla* deletion strains were performed.

To identify unambiguously flagella on the cell surface, the *aapF* in-frame deletion was also introduced into the seven existing *fla* mutant strains (Fig. S1B) to ensure pili production cannot confound flagella biogenesis analysis. Electron microscopic images of negatively stained cells of all seven double deletion mutants showed that all Δ*fla* mutants lack flagella (Fig. 5B–H). Thus we concluded that for flagella assembly all seven *fla* genes are essential. This includes also *flaX,* which is unique for crenarchaea. FlaX shares no similarity to any flagellar protein, but rather shows weak similarity to bacterial methyl-accepting proteins involved in sensing extracellular signals (Ghosh and Albers, 2011).

The Δ*aapF* Δ*fla* double deletion strains were complemented by supplying the corresponding *fla* gene copy in *trans* under control of a maltose inducible promoter (Fig. S3). By introducing *flaB* and *flaG* in *trans*, the wild-type phenotype could only partially rescued as both strains produced only short, incomplete flagella segments on the cell surface. However, for the other five complementation strains the wild-type phenotype could be restored as complete flagella were observed by TEM analysis. This clearly showed that the mutant phenotypes were only caused by the deletion of the corresponding *fla* gene.

In comparison with the limited number of proteins known to be involved in archaeal flagella assembly, more than 50 gene products are reportedly necessary to produce functional flagella in Gram-negative bacteria (Macnab, 2003; 2004). Thus our work might encourage research to further elucidate associated proteins involved in transport, folding and assembling of core flagellar proteins to a macromolecular complex.

Functional characterization of the flagella

In *M. voltae* and *S. solfataricus* it has been reported that flagella are essential for surface motility using soft agar and gelrite plates respectively (Chaban *et al.*, 2007b; Szabó *et al.*, 2007a). However, for only two euryarchaea, *H. salinarum* and *H. walsbyi*, it has been demonstrated that flagella are used for swimming motility in liquid medium (Alam and Oesterhelt, 1984; Alm and Mattick, 1997). Additionally, experimental data suggest that the same holds true for *P. furiosus* (Näther *et al.*, 2006). Accordingly, we aimed to study swimming motility and the influence of the flagella-biosynthesis genes on it in *S. acidocaldarius*. Therefore, we analysed swimming behaviour of the wild-type strain MW001 in comparison with Δ*aapF* on surfaces using semi-solid gelrite plates lacking tryptone. An increased swimming radius of ~7.5 mm was observed for Δ*aapF* after 48 h of incubation at 75°C, whereas no swimming radius was detected in any of the double deletion strains (Fig. S4).

In a parallel experiment, swimming velocities were determined under non-inducing and starvation induced conditions using a high temperature phase-contrast

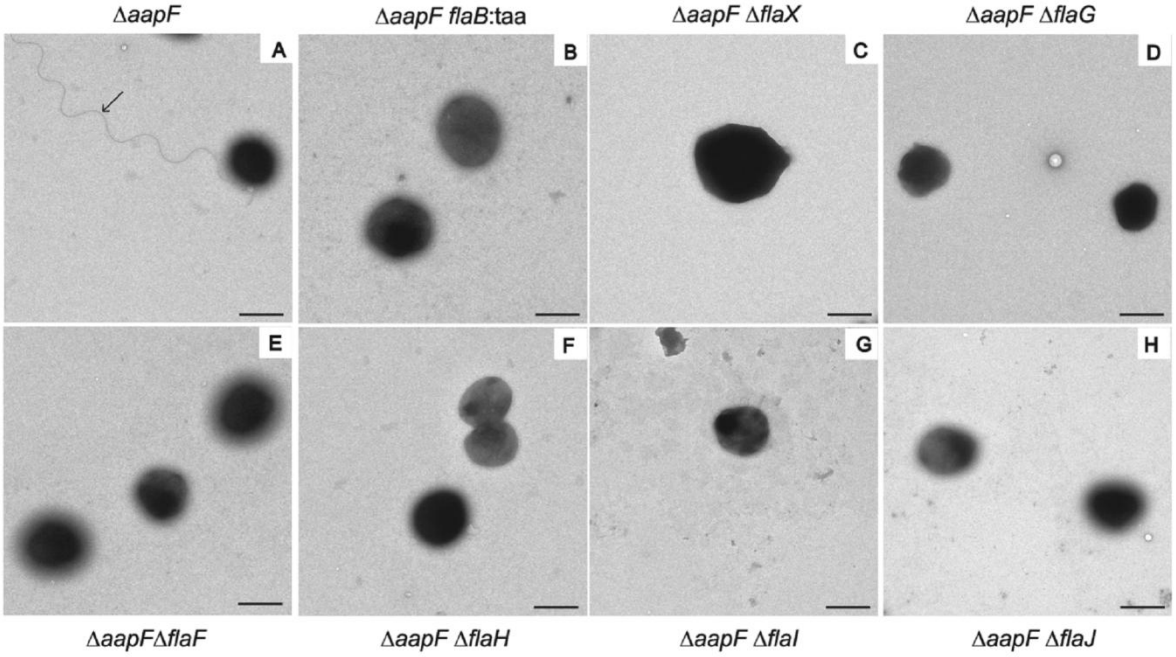


Fig. 5. TEM analysis of double deletion strains Δ*aapF* Δ*fla*. For each strain 50 cells were analysed. Representative electron micrographs of Δ*aapF* (A), and of all double deletion strains Δ*aapF* Δ*fla* (B–H) are shown. Flagella, highlighted with an arrow, were exclusively seen on the surface of Δ*aapF*, whereas all double deletion strains lack both pili and flagellar filaments. Scale bars represent 800 nm.

microscope at 75°C. Under non-inducing conditions both MW001 cells and Δ*aapF* cells were virtually non-motile. Upon tryptone depletion, MW001 cells were motile with a velocity of 30 μm s^{-1} (see Movie S1) (Fig. 6), whereas cells of the pili mutant Δ*aapF* revealed hyper-motile behaviour with an average speed of 48 μm s^{-1} (Fig. 6). Thus hyper-flagellation of Δ*aapF* presumably leads to higher velocity (see Movie S2).

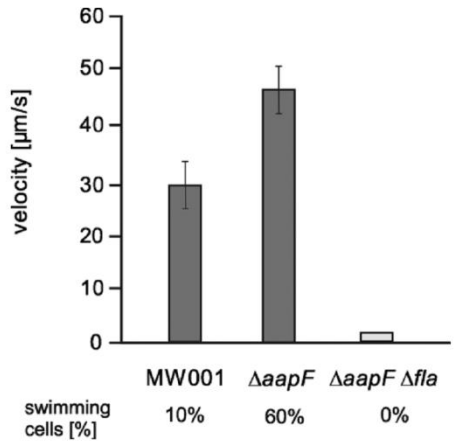


Fig. 6. Distributions of swimming velocities of MW001, Δ*aapF* and double deletion strains Δ*aapF* Δ*fla* differ significantly. Cells were incubated at 75°C for 4 h in tryptone depleted medium before thermomicroscopy. Velocities were calculated for 50 individual cells of each strain. Percentage of swimming cells for each strain is indicated below the graph.

Importantly, *flaB*:taa deletion mutant was non-motile (Fig. 6). This demonstrated for the first time the direct involvement of flagella in swimming motility for crenarchaea. Swimming motility was also not observed for any other Δ*aapF* Δ*fla* double deletion strain (see Movie S3 as an example for Δ*aapF flaB*:taa).

Taken together, we postulate that nitrogen starvation as well as the previous described temperature shock (Lewus and Ford, 1999) enhance swimming motility of *Sulfolobus* cells and thereby allow the cell population to escape from unfavourable environments. Previous studies revealed that bacterial swimming speeds of *Pseudomonas aeruginosa* PAK and *Vibrio alginolyticus* are in the same range as those we measured for *S. acidocaldarius* (Liu *et al.*, 1990; Doyle *et al.*, 2004). To date few studies have focused on archaeal swimming velocities. For the euryarchaeon *Halobacterium halobium* a velocity of 3 μm s^{-1} was determined (Weiss, 1973). However, *P. furiosus*, the 'rushing fireball', assembles up to 50 flagella on its surface and can reach velocities up to 110 μm s^{-1} (B. Herzog and R. Wirth, pers. comm.). This high velocity might be due to the fact that latter euryarchaeon has to escape rapidly from hot spots exceeding its optimal growth temperature, which is in the range of 90°C–95°C (Martins and Santos, 1995).

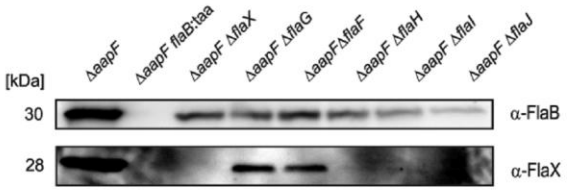


Fig. 7. Western blot analysis of Δ*aapF* Δ*flaB-J* strains. Whole cell lysates of Δ*aapF* and of the double deletion strains Δ*aapF* Δ*flaB-J* were resolved via SDS-PAGE and the expression of FlaB and FlaX was detected with anti-FlaB and anti-FlaX peptide antibodies. Numbers [kDa] refer to the sequence-derived molecular weights of FlaB and FlaX respectively.

Presence of Fla proteins in the different fla *deletion strains*

The qRT-PCR analysis of all *fla* deletion mutants has demonstrated that no polar effects of nearby genes could be observed (Fig. 2). To analyse whether *fla* gene products are expressed at the protein level, cultures of MW001 and of all seven Δ*aapF* Δ*fla* deletion strains were grown in tryptone depleted medium and whole cell lysates were subjected to SDS-PAGE. Subsequently the presence of FlaB and FlaX proteins in the *fla* deletion strains was analysed by Western blotting using antibodies directed against FlaB and FlaX. The presence of the other Fla proteins could not be probed as none of the other produced antibodies recognized their specific targets sensitively enough. Several trials to obtain better antibodies for these proteins failed so far.

For all double deletion mutants except Δ*aapF flaB*:taa, FlaB was stably expressed (Fig. 7). However, for the mutant strains *flaB*:taa, Δ*flaH*, Δ*flaI* and Δ*flaJ*, no signal was obtained by Western blot analysis with anti-FlaX antibody. These results suggest that the lack of core flagella biosynthesis proteins FlaH FlaI and FlaJ leads to destabilization of FlaX. If these interacting partners are missing, the protein appears to be rapidly degraded. However, studies need to be performed to unravel the sequential series of highly ordered events that are required for assembly of the crenarchaeal flagellum.

Promoter studies

In a previous section, we demonstrated that all seven *fla* genes of *S. acidocaldarius* are organized in two transcriptional units with two putative promoters. To confirm the *fla* promoters and to analyse their starvation dependent regulation, we cloned the regions 170 bp upstream of *flaB* and 175 bp upstream of *flaX* into the reporter vector pCMal*lacS* (Berkner *et al*., 2010) (Fig. 8A). The resulting plasmids, pSVA1600 and pSVA1601, were transformed in MW001 and in Δ*aapF*. Moreover, pSVA1614, a promoterless version of pCMal*lacS*, was included as a negative control. To determine promoter activity, exponentially growing cells harbouring the different plasmids were both inoculated in rich medium or in tryptone depleted medium. After 4 h incubation LacS activity was measured and values were background subtracted. Under non-inducing conditions, 75 Miller Units were obtained for MW001 and 74 for Δ*aapF* for the *flaB* promoter. As expected, *flaB* promoter activities increased upon tryptone depletion approximately twofold for MW001 and sevenfold for Δ*aapF* (Fig. 8B). Moreover, the strains containing the *flaX* promoter fusions also revealed LacS activity. Interestingly, tryptone depletion did not lead to any significant increase of LacS activity for the *flaX* promoter.

These data suggest a constitutive *flaX* promoter that regulates the basal expression level of the genes *flaX*, *flaG*, *flaF*, *flaH*, *flaI* and *flaJ*, which together encode flagellar accessory and core components. In contrast, the gene encoding the flagellin subunit is under control of the starvation inducible *flaB* promoter. The constitutive

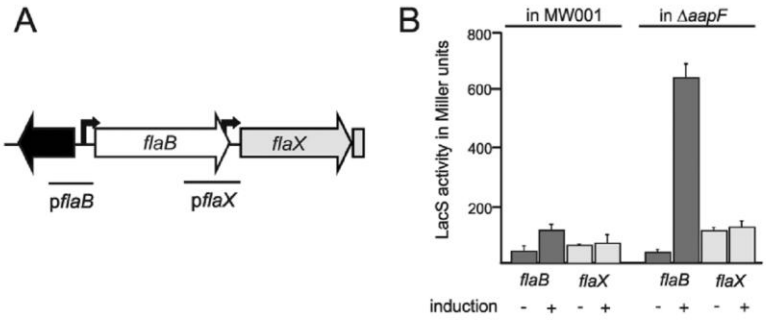


Fig. 8. Promoter activity assay of *fla* promoters.
A. Depiction of part of the *fla* gene cluster. The promoters used for the construction of the reporter gene assay plasmids are underlined.
B. LacS activity of either MW001 or δ*aapF* transformed with reporter gene assays constructs containing either the *flaB* or *flaX* promoter upstream of the *lacS* gene. Transformants were shifted for 4 h in either full medium (–) or starvation medium (+). Strength of the different promoters was measured in a β-galactosidase assay. Activities are given in Miller units.

maintenance of the flagellar core complex allows a rapid assembly of the flagellum under various environmental conditions depending on the availability of flagellin subunits. Moreover, preservation of the core complex is comparatively less energy consuming than a constitutive assembly of the filament, as for latter structure thousands of processed subunits have to assemble permanently in an energy dependent manner. Thus cells can preserve energy by a temporary nature of flagellin biosynthesis upon induction. This observation is in line with our conclusion that biosynthesis of flagella in *S. acidocaldarius* is triggered upon environmental stresses like nitrogen starvation or extreme temperature hot spots.

Regarding both the constitutive activity of the *flaX* promoter obtained by the reporter gene assays, and the starvation dependent 30-fold increase of *flaX* transcription level obtained by qRT-PCR analysis (Fig. 4D), we hypothesize an influence of the *flaB* promoter on the downstream *fla* genes. This interplay of the two *fla* promoters allows a complex regulatory network for *fla* gene expressions as well as an adjustment of the optimal ratios of the different Fla proteins.

Concluding remarks

The present study is the first detailed analysis of regulation, assembly and function of a crenarchaeal flagellum. We could demonstrate that all seven genes of the *fla* locus are essential for biosynthesis and function of the crenarchaeal flagellum, as observed for their euryarchaeal counterparts (Jarrell *et al.*, 1996; Tarasov *et al.*, 2000; Patenge *et al.*, 2001; Thomas *et al.*, 2002). Moreover, our data indicate a central role of the unique crenarchaeal protein FlaX as the presence of the flagellar core components FlaH, FlaI and FlaJ seems to be crucial for the stability of this protein. However, the function of FlaX as well as for the other accessory proteins except FlaI remains elusive. Future investigations focusing on FlaX could be the key to understand why this protein is essential only for the crenarchaeal subkingdom.

Interestingly, some archaeal *fla* genes share homology to genes coding for proteins involved in type IV pili assembly of Gram-negative bacteria (Ghosh and Albers, 2011; Pohlschroder *et al.*, 2011). In addition, both filaments structurally resemble each other in lacking a central channel (Trachtenberg *et al.*, 2005; Trachtenberg and Cohen-Krausz, 2006). Accordingly, a similar assembly mechanism is suggested. However, the mode of action of the crenarchaeal flagellum is not understood yet. Therefore an intriguing question remains: How does a type IV pilus like filaments generate torque to rotate and thereby propel the archaeon forward through its environment?

Experimental procedures

Strains and growth conditions

Sulfolobus acidocaldarius DSM 639 was grown aerobically at 75°C in Brock's basal salts medium (Brock *et al.*, 1972) which contains per litre $(NH_4)_2SO_4$, 1.3 g; KH_2PO_4, 0.28 g; $MgSO_4 \times 7H_2O$, 0.25 g; $CaCl_2 \times 2H_2O$, 0.07 g; $FeCl_3 \times 6H_2O$, 0.020 g, $MnCl_2 \times 4H_2O$, 1.80 mg; $Na_2B_4O_7 \times 10H_2O$, 4.5 mg; $ZnSO_4 \times 7H_2O$, 0.22 mg; $CuCl_2 \times 2H_2O$, 0.05 mg; $NaMoO_4 \times 2H_2O$, 0.03 mg; $VOSO_4 \times 2H_2O$, 0.03 mg; $CoSO_4 \times 7H_2O$, 0.01 mg. Moreover trace elements are added per litre: $Na_2B_4O_7 \times 10H_2O$, 4.5 mg; $ZnSO_4 \times 7H_2O$, 0.22 mg; $CuCl_2 \times 2H_2O$, 0.05 mg; $NaMoO_4 \times 2H_2O$, 0.03 mg; $VOSO_4 \times 2H_2O$, 0.03 mg; $CoSO_4 \times 7H_2O$, 0.01 mg and $MnCl_2 \times 4H_2O$, 1.8 g. Medium was adjusted to pH 3.5 with sulphuric acid and supplemented with 0.1% (w/v) tryptone (Roth) or NZAmine AS (Sigma) and 0.2% (w/v) dextrin. For the uracil auxotrophic strain *S. acidocaldarius* MW001, containing a 322 bp deletion in the *pyrE* gene (M. Wagner and S.V. Albers, pers. comm.), the growth medium was supplemented with 10 μg ml^{-1} uracil. To prepare plates, Brock medium was solidified by adding a final concentration of 0.6% (w/v) Gelrite (Sigma) and $MgCl_2$ and $CaCl_2$ to 10 mM and 3 mM respectively. Plates were then incubated for 5 days at 75°C. For the propagation of plasmids *E. coli* strain DH5α was used.

Construction of plasmids for in-frame deletions and in trans *complementations*

In-frame deletions in the genome of *S. acidocaldarius* MW001 were constructed by deleting the entire open reading frame. As *flaX* overlaps with *flaG*, *flaF* with *flaH* and *flaI* with *flaJ*, truncated versions of these genes were generated, preserving the overlapping base pairs to retain the full sequence of the adjacent *fla* gene.

In cases the target gene had an overlapping sequence with the upstream gene, the deletion construct was made, keeping the overlapping region intact in order to maintain transcription of the upstream genes as well as the downstream gene.

In general, 500–550 bp of the up- and downstream flanking regions of *flaB, flaX, flaG*, *flaF*, *flaH, flaI* and *flaJ* were amplified by PCR using the corresponding primer pairs (Δ*fla*-us-fw/Δ*fla*-us-rev and Δ*fla*-ds-fw/Δ*fla*-ds-rev) listed in Table S1. By overlap extension PCR the flanking regions were fused to the in-frame deletion product. These PCR products were cloned into the suicide vector pSVA406 yielding pSVA1613 (*flaB*), pSVA1602 (*flaX*), pSVA1604 (*flaG*), pSVA1603 (*flaF*), pSVA1605 (*flaH*), pSVA1606 (*flaI*) and pSVA1179 (*flaJ*) respectively (Table 1). For *flaB*:taa, another deletion plasmid pSVA1622 was constructed including three stop codons after 30, 36 and 45 base pairs of the original open reading frame. These three stop codons replaced nine base pairs of *flaB*.

For complementation studies, the *fla* ORFs were amplified using the primers *fla*-compl-fw-NcoI and *fla*-compl-rev-EagI respectively. The resulting PCR fragments were cloned into the NcoI-EagI restricted shuttle vector pCmal*lacS* (Berkner *et al.*, 2010), which carries a maltose promoter.

The final deletion and complementation constructs were confirmed by sequence analysis. Enzymes were obtained

from Fermentas. Kits for extracting and purifying DNA were obtained from Promega and used according to the manufacturer's instructions.

Electroporation of S. acidocaldarius

Before transformation into *S. acidocaldarius*, suicide- and shuttle plasmids were methylated to prevent restriction by the *SuaI* restriction enzyme. For that purpose *E. coli* ER1821 (New England Biolabs) bearing the additional plasmid pM.EsaBC4I (New England Biolabs) was transformed with plasmid DNAs.

Methylated deletion plasmids were electroporated in electro-competent wild-type cells MW001 using a Gene Pulser Xcell (Bio-Rad) with a constant time protocol with input parameters 1.5 kV, 25 μF, 600 Ω in 1 mm cuvettes. Before plating on uracil lacking and NZAmine containing plates, cells were regenerated for 30 min at 75°C in twofold recovery solution (1% sucrose, 20 mM β-alanine, 1.5 mM malate buffer, pH 4.5, 10 mM $MgSO_4$).

Accordingly, methylated complementation plasmids were electroporated into electro-competent Δ*aapF* Δ*fla* mutant cells and were streaked on uracil lacking and NZAmine containing plates.

Generation of deletion mutants and complementation strains

After 5 days of growth on plates insertion of the deletion plasmids into the genome was verified by colony PCR by amplification of the *pyrEF* cassette using the primers 928 and 936. For construction of the eight markerless in-frame deletion strains Δ*flaB, flaB*:taa, Δ*flaXY*, Δ*flaG*, Δ*flaF*, Δ*flaH,* Δ*flaI and* Δ*flaJ,* clones harbouring the gene targeting plasmid were streaked on counter selection medium. The second recombination step was enforced by the presence of 5-FOA and uracil. Uracil auxotrophic colonies could be obtained on plates after 5 days incubation at 75°C. All in-frame deletion strains were verified by PCR and DNA sequencing using primer pairs *fla*-fw-seq and *fla*-rev-seq.

Cells harbouring a complementation plasmid were transferred into medium containing NZAmine and 0.4% (w/v) maltose to induce *fla* gene expression.

Flagella extraction

Seven hundred millilitres of exponentially grown cultures (OD_{600} = 0.8) were pelleted (30 min, 7000 *g*), resuspended in 15 ml Brock's basal salts medium and sheared in a blender (Russell Hobbs; Model No. 14234–56) for four times 15 s. Cells and debris were removed by low-speed centrifugation (30 min, 6000 *g*). Sheared filaments were recovered by high-speed centrifugation 90 min at 339 000 *g*, and resuspended in 0.1 ml Brock's basal salts medium. Further purification by CsCl density gradient (0.5 g ml^{-1}) centrifugation for 42 h (250 000 *g*) resulted in a white fraction, which was isolated and dialysed extensively against Brock's basal salts medium (pH 3.5). Presence of isolated flagella was further confirmed by TEM analysis, SDS-PAGE, Western blotting and MS analysis. Glycosylation of the flagellins was tested on samples separated on SDS-PAGE gels using the Pro-Q Emerald 300 Glycoprotein Gel and Blot Stain Kit (Invitrogen).

Promoter activity assay

To study promoter activity, translational fusions were constructed leaving the native ribosomal binding site unaffected. Therefore, 170 bp upstream of *flaB* and 175 bp upstream of *flaX* were amplified using the primers *fla*-promoter-fw-SacII and *fla*-promoter-rev-NcoI respectively. The resulting PCR fragments were cloned into the SacII-NcoI restricted vector pCmal*lacS* (Berkner *et al.*, 2010). The resulting plasmids, pSVA1600 and pSVA1601, were transformed in MW001 and in Δ*aapF* respectively.

Promoter activity was analysed by measuring the cleavage of o-nitrophenyl-β-D-galactopyranosid (ONPG) by LacS of *S. solfataricus*. The *lacS* gene encodes a β-D-galactosidase with specificities for various β-D-glycosidic compounds (Pisani *et al.*, 1990; Grogan, 1991). In detail, 2 ml of 5×10^8 cells were resuspended in 200 μl Z-buffer (60 mM $Na_2HPO_4 \times 7H_2O$, 40 mM $NaH_2PO_4 \times H_2O$, 10 mM KCl, 1 mM $MgSO_4 \times 7H_2O$, 50 mM β-mercaptoethanol) (Miller, 1972) and lysed by adding 10 μl of $CHCl_3$. The assay was performed in 96-well plates with each reaction being composed of 180 μl Z-buffer, 10 μl cell lysate and 10 μl ONPG. All given strains were assayed in triplicates. The production of ONP was measured at 410 nm (Jonuscheit *et al.*, 2003) over a period of 4 h at 42°C using an Infinite 200 luminometer (Tecan). Miller units were calculated as described previously (Miller, 1972). With respect to the temperature optimum of LacS a conversion factor FT = 7 was determined and included in the final equation.

qRT-PCR and qualitative RT-PCR

Total RNA was isolated from 10 ml culture (OD_{600} = 0.4) grown in either full medium or tryptone minimal medium using the peqGOLD TriFast kit (Peqlab) following manufacturer's instructions. Residual chromosomal DNA was removed by treatment with RNase-free DNase I (Fermentas) for 2 h at 37°C. The absence of any DNA contamination was confirmed by ensuring the absence of PCR products before cDNA synthesis using the primer pairs RT-*fla*-fw/rev. Purified RNA served as template for cDNA synthesis using the First Strand cDNA Synthesis Kit (Fermentas) according to the manufacturer's instructions using either *fla* gene specific primers (*fla*_rev_cDNA) for transcript mapping or random hexamer primers for qRT-PCR analysis.

For transcript mapping, *fla* gene primer sets (*fla*-qRT-PCR-fw/rev) were used to check for abundance of *fla* transcripts. Control reactions were performed with DNase treated RNA as template. Each PCR reaction was performed in 25 μl including 2 μl cDNA and 1u Phusion High-Fidelity Polymerase (Thermo Fischer Scientific) in a Mastercycler gradient PCR machine (Eppendorf). An initial denaturation step of 5 min 94°C was followed by 30 cycles of 30 s 94°C, 15 s 60°C, 15 s 72°C and end step 5 min 72°C.

Quantitative PCR analysis was carried out using Maxima SYBR Green/ROX qPCR Master Mix (Fermentas) based on the Sybr green detection system (Real-Time 7300 PCR machine; Applied Biosystems). The C_q values (quantification

cycle) were determined automatically by Real-Time 7300 PCR software (Applied Biosystems) after 40 cycles. Samples were assayed at least in duplicates. Gene-specific primer sets (*fla*-qRT-PCR-fw/rev) were used to detect the amounts of transcripts of the *fla* gene cluster under both tested conditions. The efficiency of each primer pair was calculated from the average slope of a linear regression curve, which resulted from qRT-PCRs using a 10-fold dilution series (10 pg–10 ng) of *S. acidocaldarius* chromosomal DNA as template. Each obtained C_q value for the *fla* transcripts was standardized to the C_q value of the housekeeping gene Saci_0574 (*secY*) (van der Sluis *et al.*, 2006). Control reactions were performed with DNase treated RNA as template.

Transmission electron microscopy (TEM)

For TEM cells were fixed with 2.5% glutaraldehyde for 15 min at room temperature. Cell suspensions were placed directly on glow-discharged Formvar-coated 200-mesh copper grids (Plano, Wetzlar, Germany). The samples were washed twice with water, negatively stained for 20 s with 2% uranyl acetate, and air-dried. Grids were investigated either on a Philips CM 12 (FEI, Eindhoven, NL) equipped with a LaB_6 cathode operating at 120 kV or on a JEOL 2100 TEM, operated at 80 kV. Images were recorded with a 1 k × 1 k slow-scan CCD camera in case of the Philips CM12 or a fast-scan 2 k × 2 k camera F214 mounted to the JEOL 2100 TEM in combination with the EMMENU 4 software (TVIPS, Gauting Germany).

Thermomicroscopy

To determine swimming velocities of *S. acidocaldarius*, cells were analysed using a high temperature phase-contrast microscope (Olympus BX-50). The microscope, surrounded by circulating air heatings (max. 50°C), consists of a heatable microscope stage (max. 90°C) and an object lens heating (max. 80°C). All temperatures were regulated independently using a multi-channel control unit (Horn *et al.*, 1999).

Capillaries were inoculated at one end with the cell suspension and sealed at both ends. Motility video clips up to 15 s were obtained with a PCO Sensicam 1600 CCD camera. Movies were analysed with ImageJ 1.43u software (NIH) using the manual tracking plug-in. For each tested condition, the movement of 50 different cells of each tested strain – tracked on the centre of the cell – was followed for 25 consecutive images.

Swimming assays on semi-solid plates

Swimming motility on plates was analysed on semi-solid gelrite plates consisting of 0.15% gelrite supplemented with 0.001% tryptone. To inoculate the plates, 10^7 exponentially growing cells harvested from Brock medium without tryptone were dropped in the centre of the plate. Plates were incubated for 3 days in a humid chamber at 75°C. Swimming behaviour of the different *Sulfolobus* strains was analysed by measuring the swimming radius.

Preparation of membrane fractions of S. acidocaldarius

To prepare membrane fractions for immunoblotting, 50 ml of *S. acidocaldarius* cells were harvested at OD_{600} = 0.8 by centrifugation, washed with buffer (50 mM Tris-HCl pH 7.5, 10 mM $MgCl_2$, 0.25 mM PMSF), and were disrupted by sonication with a Sonoplus HD3100 (Bandelin Sonorex) at 4°C in lysis buffer [50 mM Tris-HCl pH 7.5, 10% (v/v) glycerol, 10 mM $MgCl_2$, 1 mM dithiothreitol, 0.5 mM PMSF, 1 mM EDTA]. After removal of intact cells and cell debris by low spin centrifugation using a Heraeus Multifuge 3SR (30 min, 4000 *g*, rotor 75006445), membrane fractions were collected by ultracentrifugation in a Optima MAX-XP ultracentrifuge (60 min, 240 000 *g*, rotor MLA-55, Beckman Coulter). The pellet contained the membrane fraction, whereas the supernatant represented the cytoplasmic fraction. Membrane fractions were washed once in TG buffer [50 mM Tris-HCl pH 7.5, 10% (v/v) glycerol] and finally suspended in 200 µl of the same buffer. Fractions were stored until use at –80°C.

Western blotting

Cultures for immunological analyses were grown in liquid medium to an OD_{600} = 0.5 and harvested by centrifugation at 3400 *g*. The pellet was resuspended in fresh medium lacking tryptone. After 4 h, 0.001% tryptone was added to each culture for growth over night. Cells were harvested and the pellet was resuspended in buffer containing 50 mM HEPES and 150 mM NaCl. Proteins were separated by SDS-PAGE according to the method of Laemmli (Laemmli, 1970) and transferred to a PVDF membrane (Roche) by semi-dry blotting as described previously (Towbin *et al.*, 1979).

Polyclonal peptide antibodies to *S. acidocaldarius* anti FlaB and anti FlaX were raised in rabbits (Eurogentec). Primary peptide antibodies were diluted to working concentrations of 1:1000 for anti FlaB and anti FlaX. As secondary antibody, the alkaline phosphatase goat anti-rabbit immunoglobulin G (Sigma), was used 1:7000 diluted. Blots were developed using the CDP-star chemiluminescent detection kit (Roche) according to the manufacturer's instructions. Chemifluorescence was measured in the LAS-4000 Luminescent image analyzer (Fujifilm, Düsseldorf, Germany). Intensities of obtained signals were analysed using ImageJ 1.43u software (NIH).

Acknowledgements

K.L. was supported by a fellowship from the IMPRS (International Max Planck Research School). S.V.A. received a VIDI grant from the Dutch Science Organization (NWO) and intramural funds from the Max Planck Society. We thank John Blackwood for his comments and critical reading of the manuscript.

References

Adams, G.M., Wright, R.L., and Jarvik, J.W. (1985) Defective temporal and spatial control of flagellar assembly in a mutant of *Chlamydomonas reinhardtii* with variable flagellar number. *J Cell Biol* **100:** 955–964.

Alam, M., and Oesterhelt, D. (1984) Morphology, function and isolation of halobacterial flagella. *J Mol Biol* **176:** 459–475.

Alam, M., Claviez, M., Oesterhelt, D., and Kessel, M. (1984) Flagella and motility behaviour of square bacteria. *EMBO J* **3:** 2899–2903.

Albers, S.V., and Pohlschröder, M. (2009) Diversity of

archaeal type IV pilin-like structures. *Extremophiles* **13:** 403–410.
Alm, R.A., and Mattick, J.S. (1997) Genes involved in the biogenesis and function of type-4 fimbriae in *Pseudomonas aeruginosa*. *Gene* **192:** 89–98.
Anderson, J.K., Smith, T.G., and Hoover, T.R. (2010) Sense and sensibility: flagellum-mediated gene regulation. *Trends Microbiol* **18:** 30–37.
Bardy, S.L., and Jarrell, K.F. (2002) FlaK of the archaeon *Methanococcus maripaludis* possesses preflagellin peptidase activity. *FEMS Microbiol Lett* **208:** 53–59.
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.
Berg, H.C., and Anderson, R.A. (1973) Bacteria swim by rotating their flagellar filaments. *Nature* **245:** 380–382.
Berkner, S., Wlodkowski, A., Albers, S.V., and Lipps, G. (2010) Inducible and constitutive promoters for genetic systems in *Sulfolobus acidocaldarius*. *Extremophiles* **14:** 249–259.
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 Mikrobiol* **84:** 54–68.
Chaban, B., Ng, S.Y., Kanbe, M., Saltzman, I., Nimmo, G., Aizawa, S., and Jarrell, K.F. (2007a) 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.
Chaban, B., Ng, S., Kanbe, M., Saltzman, I., Nimmo, G., Aizawa, S.-I., and Jarrell, K. (2007b) 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.
Craig, L., and Li, J. (2008) Type IV pili: paradoxes in form and function. *Curr Opin Struct Biol* **18:** 267–277.
Doyle, T.B., Hawkins, A.C., and McCarter, L.L. (2004) The complex flagellar torque generator of *Pseudomonas aeruginosa*. *J Bacteriol* **186:** 6341–6350.
Duan, X., and He, Z.G. (2011) Characterization of the specific interaction between archaeal FHA domain-containing protein and the promoter of a flagellar-like gene-cluster and its regulation by phosphorylation. *Biochem Biophys Res Commun* **407:** 242–247.
Faguy, D.M., Jarrell, K.F., Kuzio, J., and Kalmokoff, M.L. (1994) Molecular analysis of archael flagellins: similarity to the type IV pilin-transport superfamily widespread in bacteria. *Can J Microbiol* **40:** 67–71.
Ghosh, A., and Albers, S.V. (2011) Assembly and function of the archaeal flagellum. *Biochem Soc Trans* **39:** 64–69.
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.
Grogan, D.W. (1991) Evidence that beta-galactosidase of *Sulfolobus solfataricus* is only one of several activities of a thermostable beta-d-glycosidase. *Appl Environ Microbiol* **57:** 1644–1649.
Henche, A.L., Koerdt, A., Ghosh, A., and Albers, S.V. (2011) Influence on cell surface structures on crenarchaeal biofilm formation using a thermostable green fluorescent protein. *Environ Microbiol*. doi: 10.1111/j.1462-2920.2011.02638.x.
Heuner, K., Brand, B.C., and Hacker, J. (1999) The expression of the flagellum of *Legionella pneumophila* is modulated by different environmental factors. *FEMS Microbiol Lett* **175:** 69–77.
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.
Jarrell, K.F., and McBride, M.J. (2008) The surprisingly diverse ways that prokaryotes move. *Nat Rev Microbiol* **6:** 466–476.
Jarrell, K.F., Bayley, D.P., Florian, V., and Klein, A. (1996) Isolation and characterization of insertional mutations in flagellin genes in the archaeon *Methanococcus voltae*. *Mol Microbiol* **20:** 657–666.
Jarrell, K.F., Ng, S.Y., and Chaban, B. (2007) Flagellation and chemotaxis. In *Archaea: Molecular and Cellular Biology*. Cavicchioli, R. (ed.). Washington, DC: ASM Press, pp. 385–410.
Jones, C.J., and Aizawa, S. (1991) The bacterial flagellum and flagellar motor: structure, assembly and function. *Adv Microb Physiol* **32:** 109–172.
Jonuscheit, M., Martusewitsch, E., Stedman, K.M., and Schleper, C. (2003) A reporter gene system for the hyperthermophilic archaeon *Sulfolobus solfataricus* based on a selectable and integrative shuttle vector. *Mol Microbiol* **48:** 1241–1252.
Kearns, D.B. (2010) A field guide to bacterial swarming motility. *Nat Rev Microbiol* **8:** 634–644.
Koval, S.F., and Jarrell, K.F. (1989) Ultrastructure and biochemistry of the cell wall of *Methanococcus voltae*. *J Bacteriol* **169:** 1298–1306.
Laemmli, U.K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227:** 680–685.
Larsen, S.H., Adler, J., Gargus, J.J., and Hogg, R.W. (1974) Chemomechanical coupling without ATP: the source of energy for motility and chemotaxis in bacteria. *Proc Natl Acad Sci USA* **71:** 1239–1243.
Lewus, P., and Ford, R.M. (1999) Temperature-sensitive motility of *Sulfolobus acidocaldarius* influences population distribution in extreme environments. *J Bacteriol* **181:** 4020–4025.
Liu, J.Z., Dapice, M., and Khan, S. (1990) Ion selectivity of the *Vibrio alginolyticus* flagellar motor. *J Bacteriol* **172:** 5236–5244.
Macnab, R.M. (2003) How bacteria assemble flagella. *Annu Rev Microbiol* **57:** 77–100.
Macnab, R.M. (2004) Type III flagellar protein export and flagellar assembly. *Biochim Biophys Acta* **1694:** 207–217.
Manson, M.D., Tedesco, P., Berg, H.C., Harold, F.M., and Van der Drift, C. (1977) A protonmotive force drives bacterial flagella. *Proc Natl Acad Sci USA* **74:** 3060–3064.
Martins, L.O., and Santos, H. (1995) Accumulation of mannosylglycerate and di-myo-inositol-phosphate by *Pyrococcus furiosus* in response to salinity and temperature. *Appl Environ Microbiol* **61:** 3299–3303.
Maurer, L.M., Yohannes, E., Bondurant, S.S., Radmacher,

M., and Slonczewski, J.L. (2005) pH regulates genes for flagellar motility, catabolism, and oxidative stress in *Escherichia coli* K-12. *J Bacteriol* **187:** 304–319.
Meyer, B., 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*. doi: 10.1111/j.1365-2958.2011.07875.x.
Miller, J.H. (1972) *Experiments in Molecular Genetics*. p. 466 New York: Cold Spring Harbor Laboratory.
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.
Ng, S., Chaban, B., and Jarrell, K. (2006) Archaeal flagella, bacterial flagella and type IV pili: a comparison of genes and posttranslational modifications. *J Mol Microbiol Biotechnol* **11:** 167–191.
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.
Peabody, C.R., Chung, Y.J., Yen, M.R., Vidal-Ingigliardi, 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.
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 2010* pii: 754101.
Pisani, F.M., Rella, R., Raia, C.A., Rozzo, C., Nucci, R., Gambacorta, A., *et al.* (1990) Thermostable beta-galactosidase from the archaebacterium *Sulfolobus solfataricus*. Purification and properties. *Eur J Biochem* **187:** 321–328.
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.
Ryan, K.R., and Shapiro, L. (2003) Temporal and spatial regulation in prokaryotic cell cycle progression and development. *Annu Rev Biochem* **72:** 367–394.
Scharf, B. (2002) Real-time imaging of fluorescent flagellar filaments of *Rhizobium lupini* H13-3: flagellar rotation and pH-induced polymorphic transitions. *J Bacteriol* **184:** 5979–5986.
Schmitt, R. (2003) Helix rotation model of the flagellar rotary motor. *Biophys J* **85:** 843–852.
van der Sluis, E.O., Nouwen, N., Koch, J., de Keyzer, J., van der Does, C., Tampe, R., and Driessen, A.J. (2006) Identification of two interaction sites in SecY that are important for the functional interaction with SecA. *J Mol Biol* **361:** 839–849.
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.
Summers, K.E., and Gibbons, I.R. (1971) Adenosine triphosphate-induced sliding of tubules in trypsin-treated flagella of sea-urchin sperm. *Proc Natl Acad Sci USA* **68:** 3092–3096.
Szabó, Z., Sani, M., Groeneveld, M., Zolghadr, B., Schelert, J., Albers, S.-V., *et al.* (2007a) Flagellar motility and structure in the hyperthermoacidophilic archaeon *Sulfolobus solfataricus*. *J Bacteriol* **189:** 4305–4309.
Szabó, Z., Adriana, O., Albers, S., Kissinger, J., Driessen, A., and Pohlschröder, M. (2007b) Identification of diverse archaeal proteins with class III signal peptides cleaved by distinct archaeal prepilin peptidases. *J Bacteriol* **189:** 772–778.
Tarasov, V.Y., Pyatibratov, M.G., Tang, S.L., Dyall-Smith, M., and Fedorov, O.V. (2000) Role of flagellins from A and B loci in flagella formation of *Halobacterium salinarum*. *Mol Microbiol* **35:** 69–78.
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.
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.
Thomas, N.A., Mueller, S., Klein, A., and Jarrell, K.F. (2002) Mutants in *fla*I and *fla*J of the archaeon *Methanococcus voltae* are deficient in flagellum assembly. *Mol Microbiol* **46:** 879–887.
Towbin, H., Staehelin, T., and Gordon, J. (1979) Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. *Proc Natl Acad Sci USA* **76:** 4350–4354.
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.
Trachtenberg, S., Galkin, V.E., and Egelman, E.H. (2005) Refining the structure of the *Halobacterium salinarum* flagellar filament using the iterative helical real space reconstruction method: insights into polymorphism. *J Mol Biol* **346:** 665–676.
Turner, L., Ryu, W.S., and Berg, H.C. (2000) Real-time imaging of fluorescent flagellar filaments. *J Bacteriol* **182:** 2793–2801.
Weiss, R.L. (1973) Attachment of Bacteria to Sulphur in Extreme Environments. *J Gen Microbiol* **77:** 501–507.
Wurtzel, O., Sapra, R., Chen, F., Zhu, Y., Simmons, B.A., and Sorek, R. (2010) A single-base resolution map of an archaeal transcriptome. *Genome Res* **20:** 133–141.
Zähringer, 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.

Supporting information

Additional supporting information may be found in the online version of this article.

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

Supplementary material for

"Molecular analysis of the crenarchaeal flagellum" by

Lassak *et al.*

Table S1. Primers used in this study.

primer	sequence (5´- 3´)	purpose
	primers for operon mapping	
1464	GAA CTG GAG CCA TGA ATA AAG GTG TAT CA	operon-Saci_1179/*flaB*-fw
1465	GCT GTA TCC AAT CCT GCT AAA CCT CT	operon-Saci_1179/*flaB*-rev
1466	GCT GGA TCC GTA ATA GGT GTG ATG AT	operon-*flaB/flaX*-fw
1467	GGT ATA CCA ACT CCT ATC AAT ATT ATG AAT AAT G	operon-*flaB/flaX*-rev
1468	CCT GTT GGA GAT AAT GGA ATA TGA TGA AGA	operon-*flaX/flaG*-fw
1469	CGA CTG AAA AAA CGG TTC CCG CTA AAA C	operon-*flaX/flaG*-rev
1470	CGA TCT CTT ACC TTC AAA ATA GTG TAG TG	operon-*flaG/flaF*-fw
1471	GCA CTC AAC AAT ACA ATT CCA AAT GAA ACA C	operon-*flaG/flaF*-rev
1472	CCA TAT CCT CCG TAT CCA AAT ACT CAA G	operon-*flaF/flaH*-fw
1473	CTT AAT CAA GTA ATC TTT AGT AGT CTG CTC C	operon-*flaF/flaH*-rev
1474	GCG ACT ATA GGT GGA AGA AGA GTA AAG AT	operon-*flaH/flaI*-fw
1475	CCA TCC TCA CTT TTT ATA CTC TGA ATA TGA	operon-*flaH/flaI*-rev
1476	GCT TGT TAT ATG ATG AGT TGC AGA TGC	operon-*flaI/flaJ*-fw
1477	CCT AAA GAC CTT CAA ATA CGG ATG AAA AG	operon-*flaI/flaJ*-rev
1478	GCT CAC CGC GCA ACT GTT TCA A	operon-*flaJ*/Saci_1171-fw
1479	GGA AGT TAT AGA TAC GAG ATC GCC T	operon-*flaJ*/Saci_1171-rev
	primers for first strand cDNA synthesis	
1467	GGT ATA CCA ACT CCT ATC AAT ATT ATG AAT AAT G	operon-*flaB/flaX*-rev
2381	GCA GCA ATA GAA GCT AAC TC	*flaX*_rev_cDNA
3426	CGG AGA AGA CAA ATA GAT AAT GAT	*flaG*_rev_cDNA
1459	CCT CGA TAA GGA TAA TGG TAC AAC C	*flaH*_rev_cDNA
3427	CGC AAA ACT CAT GAC GAA AAG	*flaJ*_rev_cDNA
	internal primers to confirm cDNA integrity	
2378	ACT GCG TCT ACT GCG TTA TCT TTA TC	RT- *flaB*-fw
2379	TAT CAC TTC CGG CTA ATG GG	RT- *flaB*-rev
2380	TCT CAA GAA CAG CGA TTT CC	RT- *flaX*-fw
2381	GCA GCA ATA GAA GCT AAC TC	RT- *flaX*-rev

2382	TCA GTC GCC TCA TCT ATC TC	RT- *flaG*-fw
2383	TAC TGA CCC TGG CTG TAG TG	RT- *flaG*-rev
2384	TCC GTA TCA TTA ACG GGA AG	RT- *flaF*-fw
2385	AGA GCC TGG AGC TAA CAA AG	RT- *flaF*-rev
2386	AGA AGA CTA GGA GGA ATA CC	RT- *flaH*-fw
2387	CAG CCG AGA GCT TAA GAT AC	RT- *flaH*-rev
2388	GAG AAA CCG CAT CTG GAA AG	RT- *flaI*-fw
2389	TCG ACA ACC CTC CTG ATT AG	RT- *flaI*-rev
2390	GCA GAC ATG ACT GTC CTA TC	RT- *flaJ*-fw
2391	TCT CAC TGA GCA TAG TTA CG	RT- *flaJ*-rev
	primers for construction of deletion strains	
1408	CGG CTG CAG GCC TTA TTA GAT AGT CTA TAC CCT CC	Δ*flaB*-us-fw-*PstI*
1409	ACT GTC ACC CTT CTT TTC ATG TAT ATC ATG TTT CTA TC	Δ*flaB*-us-rev
1410	AAG AAG GGT GAC AGT TAT AGG ATA AAT GGT TTA AAA TG	Δ*flaB*-ds-fw
1411	GTG GGC CCG CTG AAT TAT TGT ATT AAA CGG TGA GG	Δ*flaB*-ds-rev-*ApaI*
2306	CTA ATT AGC CTA TGC TCA TCT TGT TTT CCT TCT TTT CAT GTA TAT	Δ*flaB*-us-rev-stop codons
2307	TGA GCA TAG GCT AAT TAG GGT TTA GCA GGA TTG GAT ACA G	Δ*flaB*-ds-fw-stop codons
1404	CCT CAC TAA CCA TTA CTC ATG CAC CGC CTT AGT G	Δ*flaX*-us-rev
1405	CGG TGC ATG AGT AAT GGT TAG TGA GGT TAT AAG TGA G	Δ*flaX*-ds-fw
1406	GTG CTG CAG GTT TAC AGC TGC TAG TAG AGG	Δ*flaX*-us-fw-*PstI*
1407	GTG GGG CCC GTC CAA TAA TAA GGG CTA GGG	Δ*flaX*-ds-rev-*ApaI*
1412	GCG CTG CAG AGG ATG TGT TAG GGA AGA TGG	Δ*flaG*-us-fw-*PstI*
1413	CGC CAA AGT TTG ACT AAC CAT TCA CAC CAC CAC C	Δ*flaG*-us-rev
1414	GTG AAT GGT TAG TCA AAC TTT GGC GTT TAT TCT AAT AAC	Δ*flaG*-ds-fw
1415	AGG GGC CCG ATG CTG GAT ATG GTA TTC CTC C	Δ*flaG*-ds-rev-*ApaI*
1416	GTG CTG CAG AGG GTG GTG GTG TGA ATG G	Δ*flaF*-us-fw-*PstI*
1417	GTA CTT ATA ATC ATT TTA GCT TAC CTT AAA CAT G	Δ*flaF*-us-rev
1418	GTA AGC TAA AAT GAT TAT AAG TAC CGG TAA TGA CG	Δ*flaF*-ds-fw
1419	GTG GGC CCT CTT CTT CCA CCT ATA GTC G	Δ*flaF*-ds-rev-*ApaI*
1420	GTG CTG CAG ATT GTT ACA CCC AAT GG	Δ*flaH*-us-fw-*PstI*
1421	CAC GAA GCT AAT CAT AGG CTT CCC CTC C	Δ*flaH*-us-rev
1422	CCT ATG ATT AGC TTC GTG GAA GAC TAT TTA AC	Δ*flaH*-ds-fw
1423	GTG GGC CCC AGA CTC GCA TCA ACT ACT GG	Δ*flaH*-ds-rev-*ApaI*
1438	GTG GGC CCG GAG CAG ACT ACT AAA GAT TA	Δ*flaI*-us-fw-*ApaI*
1439	TAT TAC TGA CAT GAC TTA TGC CCT CGA T	Δ*flaI*-us-rev

1440	ATA AGT CAT GTC AGT AAT ATT TGA CAG GAA TAA A	Δ*flaI*-ds-fw
1441	GTA GTC GAC CGT GGA GAG AAA GAA AAT CT	Δ*flaI*-ds-rev-*SalI*
638	GCG CTG CAG AAA CCG CAT CTG G	Δ*flaJ*-us-fw-*PstI*
600	GGT CCT TTC AAA ATA AGT ACC TTT GGT CAT ATA TTT TCA TCA AAT ATT ACT GAC ATA TTT TAT CGC CTC CTC C	Δ*flaJ*-us-rev
602	TCA AAT ATT ACT GAC ATA TTT TAT CGC CTC CTC CTG AAA ATA TAT GAC CAA AGG TAC TTA TTT TGA AAG GAC C	Δ*flaJ*-ds-fw
641	GCG GGA TCC GAG TGT TTG ACA TAC TTA GAG	Δ*flaJ*-ds-rev-*BamHI*
640	GCG CTG CAG TCA ACC TCT TCA TTC AAT AC	Δ*aapF*-us-fw-*PstI*
606	AAT TTA TAT ACT TTT ACT GTG TGA ATA TAC ACA ACT AGA TAA AGT TAA ATA TTT TTT ATA	Δ*aapF*-us-rev
607	TAT AAA AAA TAT TTA ACT TTA TCT AGT TGT GTA TAT TCA CAC AGT AAA AGT ATA TAA ATT	Δ*aapF*-ds-fw
642	GCG GGA TCC TCT ACC GGC AGG GAT AGA AG	Δ*aapF*-ds-rev-*BamHI*
	sequencing primers to confirm deletion strains	
1452	GCA CTG TTA ATT ATG TAG CTT TAG	Δ*flaB*-fw-seq
1453	CCA TGT TAA GTT TAC GTT ATT AG	Δ*flaB*-rev-seq
2326	CAAGATGAGCATAGGCTAATTAGGGT	Δ*flaB*-stop-seq
1454	CCA CCA ATA CAA AGA GTT ACT G	Δ*flaX*-fw-seq
1455	GCT TGA GTA TTT GGA TAC GGA G	Δ*flaX*-rev-seq
1456	GCA AAC CCA GCA AAG CCC TC	Δ*flaG*-fw-seq
1457	CGA ATT GAG CAG ATA ACA CAC TC	Δ*flaG*-rev-seq
1458	GCT GCA AAT ATA GCT GAG AGT TAC TC	Δ*flaF*-fw-seq
1459	CCT CGA TAA GGA TAA TGG TAC AAC C	Δ*flaF*-rev-seq
1460	GGT TCT TCC CCC TAT TGG AC	Δ*flaH*-fw-seq
1461	CCT CTT CTA CTT ATA TCA ACA CC	Δ*flaH*-rev-seq
1462	GGA GAT CAT GGC ACG GGT AAG	Δ*flaI*-fw-seq
1463	CAG TTA ATA AAA GTC TGT CTC TA	Δ*flaI*-rev-seq
581	GAG TCT GCC TGA CGG TTC T	Δ*flaJ*-fw-seq
582	GGA GAG TTA AGC TTT CGG CC	Δ*flaJ*-rev-seq
583	CTG CTA TTC TAT CTC CTG CAG G	Δ*aapF*-fw-seq
584	CAG TGT TGC TGG AGC TC	Δ*aapF*-rev-seq
939	CCG ATA CGC GTA CTG GAT CCT GAC CTA GGT TTG AGC AGT TCT AGT ACT TG	check-Sso-*pyrEF*-fw
928	ACG TCA TCT ACT AAC AAT ACT C	check-Sso-*pyrEF*-rev
	primers for complementations of deletion strains in trans	
3610	GGG GGG CCA TGG ATG GTA TAC ATG AAA AGA AGG	*flaB*-compl-fw-*NcoI*
3611	GGG GGG CGG CCG TTA TCC TAT AAC TGT CAC TTT TCC	*flaB*-compl-rev-*EagI*
3600	GGG GGG CCA TGG ATG GCT ATC CAA GAT CTA TTA CAA TC	*flaX*-compl-fw-*NcoI*
3601	GGG GGG CGG CCG CTA ACC ATT CAC ACC ACC AC	*flaX*-compl-rev-*EagI*

3602	GGG GGG CCA TGG ATG GTT AGT GAG GTT ATA AGT GAG	*flaG*-compl-fw-*NcoI*
3603	GGG GGG CGG CCG TTA GCT TAC CTT AAA CAT GTA AC	*flaG*-compl-rev-*EagI*
3604	GGG GGG CCA TGG ATG GGA GTG TCA CAA ACT TTG	*flaF*-compl-fw-*NcoI*
3605	GGG GGG CGG CCG TCA TAG GCT TCC CCT CCA TAT TAC	*flaF*-compl-rev-*EagI*
3606	GGG GGG CCA TGG ATG ATT ATA AGT ACC GGT AAT G	*flaH*-compl-fw-*NcoI*
3607	GGG GGG CGG CCG TTA TGC CCT CGA TAA GGA TAA	*flaH*-compl-rev-*EagI*
763	CCC CCC CAT GGT AAG CTT CGT GGA AGA C	*flaI*-compl-fw-*NcoI*
849	CCC CCC GGC CGT CAA ATA TTA CTG ACA TAT TTT ATC	*flaI*-compl-rev-*EagI*
3608	GGG GGG CCA TGG ATG TCA GTA ATA TTT GAC AGG	*flaJ*-compl-fw-*NcoI*
3609	GGG GGG CGG CCG TCA AGT TGT ATT CTG AAT GCC	*flaJ*-compl-rev-*EagI*
	quantitative RT-PCR primers to exclude polar effects	
1424	ACT GCG TCT ACT GCG TTA TCT TTA TC	*flaB*-qRT-PCR-fw
1425	GGA GAT AAG TCT ACA CTA GAT ACA CCA GAA	*flaB*-qRT-PCR-rev
1426	GCA GTT GAA GAG TTA GCC TTA TCT GTG	*flaX*-qRT-PCR-fw
1427	CCT ACT AAC TGA CTT ACG GTA CTA ATC T	*flaX*-qRT-PCR-rev
1428	CCT GGC TGT AGT GAA TTA GAT GTA ACT G	*flaG*-qRT-PCR-fw
1429	GTG TAG TGT ATT TCG GTC CAA ATG GTC A	*flaG*-qRT-PCR-rev
2310	GGA ATT GTA TTG TTG AGT GCA TAT TTT AAG AG	*flaF*-qRT-PCR-fw
2311	GGA GCT TAC AGT TAT CTT CGT TTC TAA CT	*flaF*-qRT-PCR-rev
1432	AGT TGA TGT GTA TCT TAA GCT CTC GG	*flaH*-qRT-PCR-fw
1433	CTG AAC CAG ATA TTC CTC CTG TAG TTT TTA	*flaH*-qRT-PCR-rev
1434	GGA GAA ACC GCA TCT GGA AAG ACA AC	*flaI*-qRT-PCR-fw
1435	GGA ACC GTC AAT TCT GGA GTG TCC T	*flaI*-qRT-PCR-rev
1436	CCA GAA AGG AGC AGA ACG GTA GG	*flaJ*-qRT-PCR-fw
1437	GCT ATT ACC GAA GCC AAT TCA CCA ATC	*flaJ*-qRT-PCR-rev
1480	CCT GCA ACA TCT ATC CAT AAC ATA CCG A	*secY*-housekeep-qRT-PCR-fw
1481	CCT CAT AGT GTA TAT GCT TTA GTA GTA G	*secY*-housekeep-qRT-PCR-rev
	primers for transcriptional reporter fusions (ONPG assays)	
1400	CAC CCG CGG GCC TTC TCA TCA GTA TCA TG	*flaB*-promoter-fw-*SacII*
1401	CCC CCC ATG GTT CTA TCA CTT TTT CAA AAT GAG ATT TCT AAG TTG GTC	*flaB*-promoter-rev-*NcoI*
1402	GCC CGC GGC GAG ACT GTG ACA ATT CAG ATA ACT CC	*flaX*-promoter-fw-*SacII*
1403	AGC CAT GGC TCA TGC ACC GCC TTA GTG GG	*flaX*-promoter-rev-*NcoI*

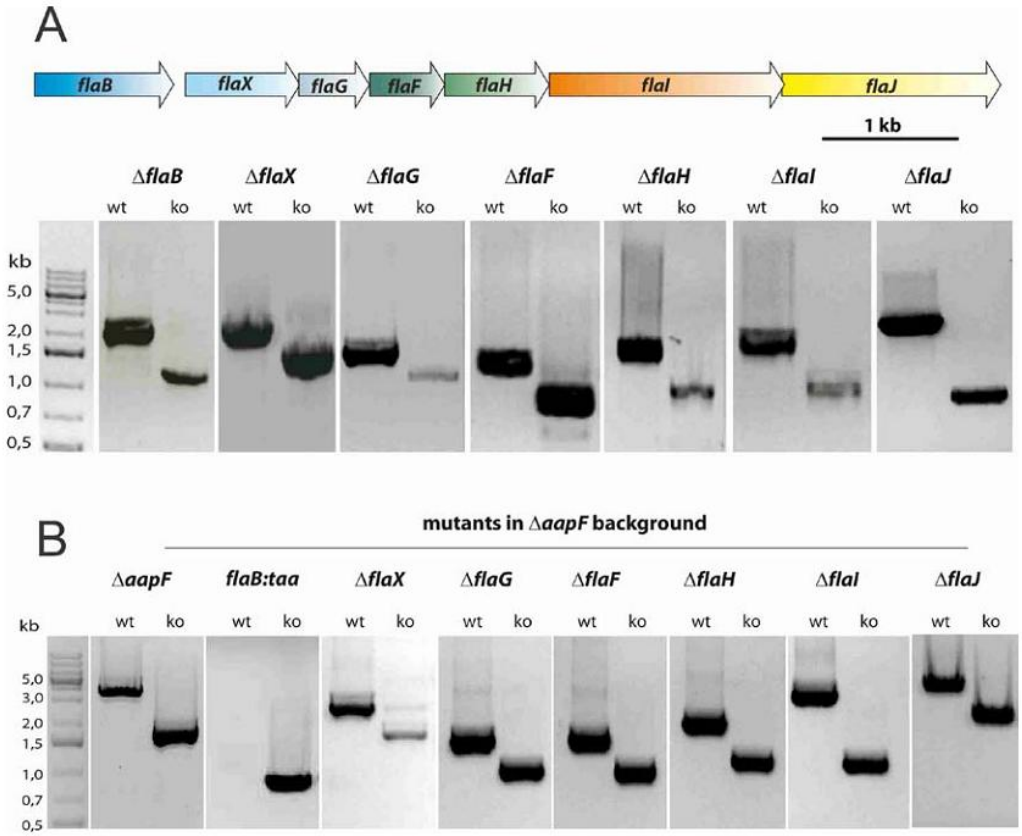


Fig. S1: PCR analysis of Δ*fla* deletion strains. Gene deletions were confirmed by PCR analysis using primers binding outside of the flanking regions of the corresponding *fla* genes that were used for constructing the deletion mutant plasmid. For each mutant, PCR amplifications were performed using genomic DNA of either the wild type MW001 (wt)/Δ*aapF* (wt), respectively, or the deletion strain (ko). (A) depicts the single mutants in the MW001 whereas (B) shows the PCR results on the mutants in the Δ*aapF* strain.

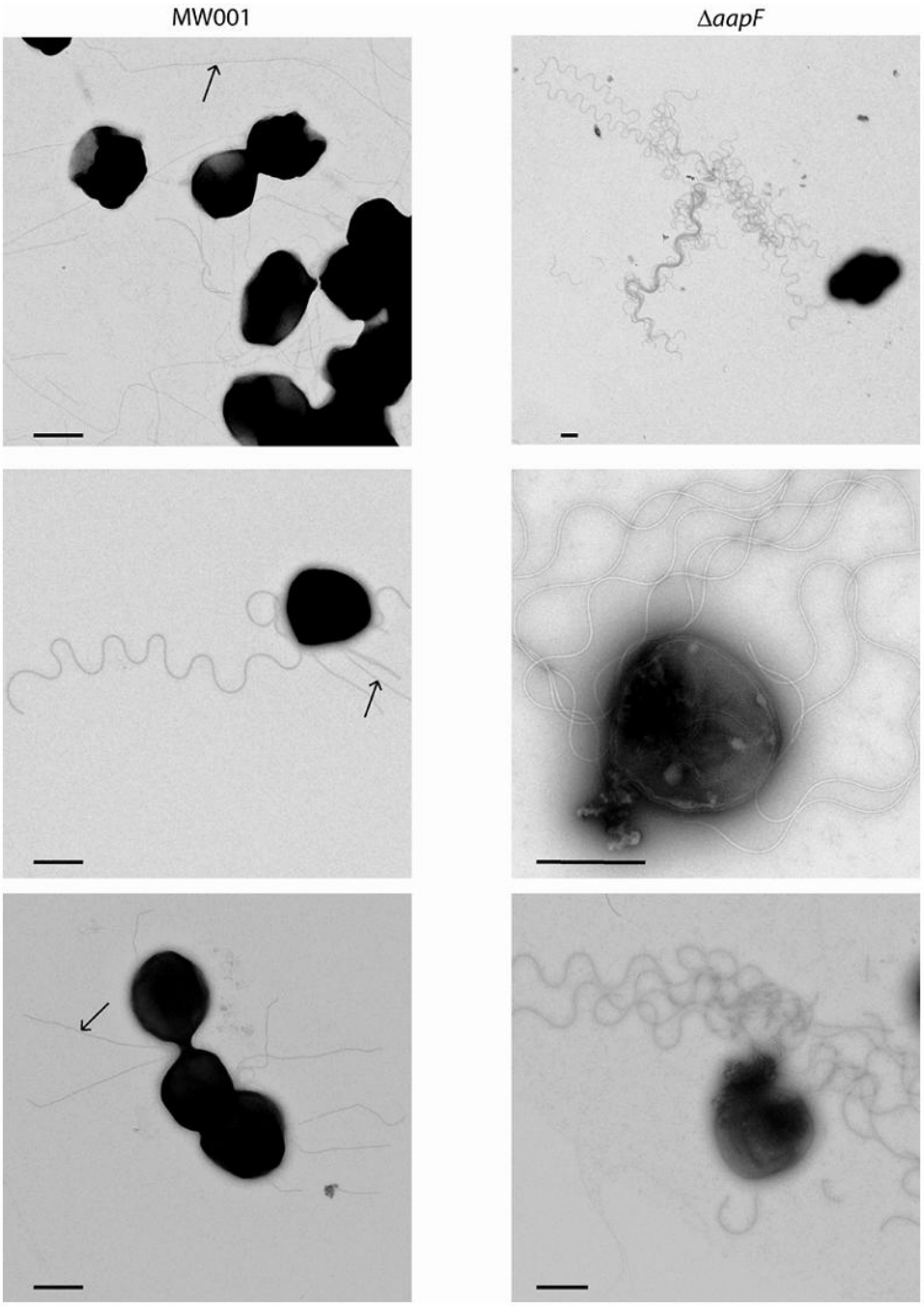


Fig. S2: Representative electron micrographs of negatively stained MW001 cells (left) and Δ*aapF* cells (right) reveal a different extent of flagella production in both *S. acidocaldarius* strains. Pili are highlighted with an arrow and were observed exclusively for MW001 cells. Size bars represent 600 nm.

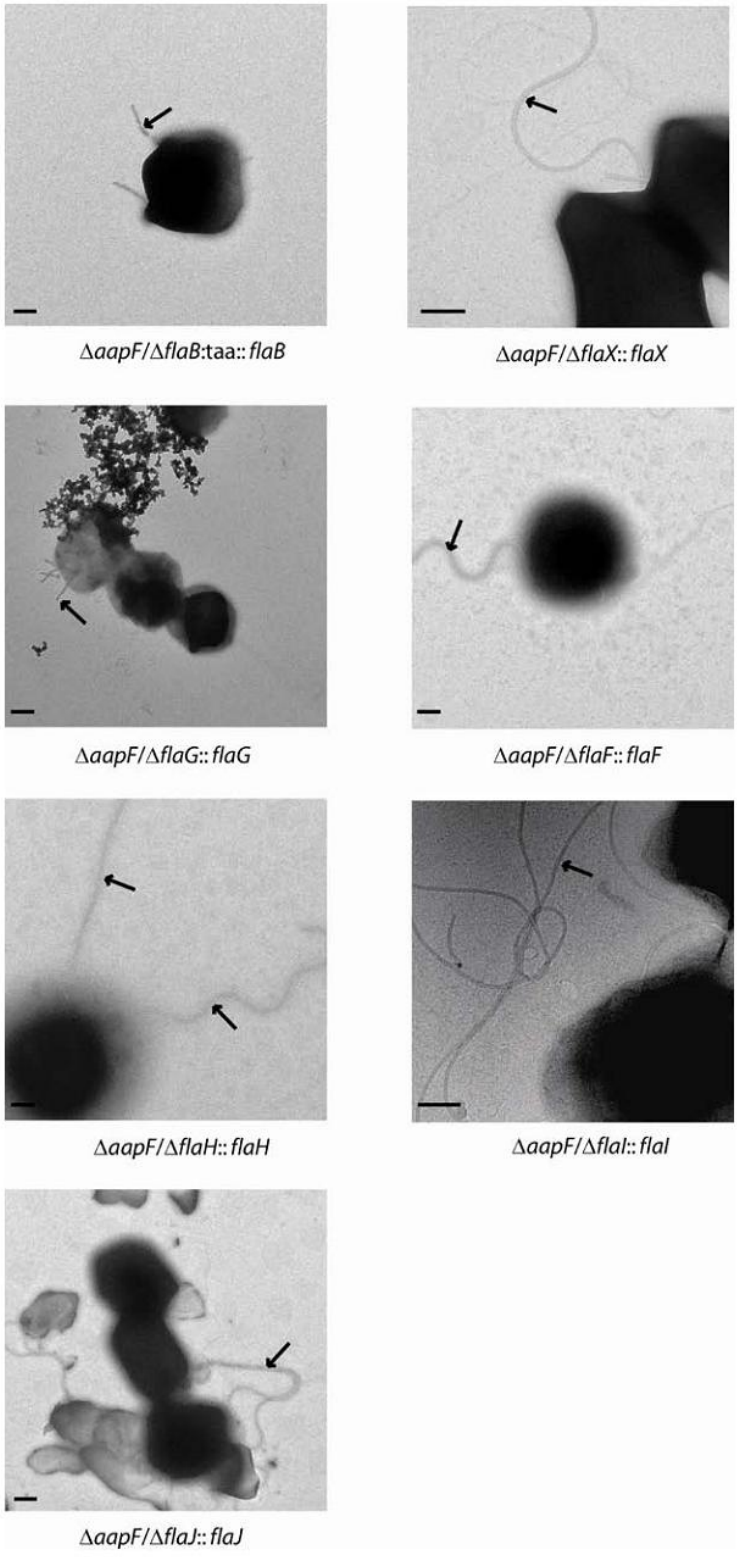


Fig. S3: TEM analysis of complemented *S. acidocaldarius* Δ*aapF* Δ*fla/ fla*$^{+}$ strains. Cells were negatively stained. Arrows indicate flagella. Size bars represent 250 nm.

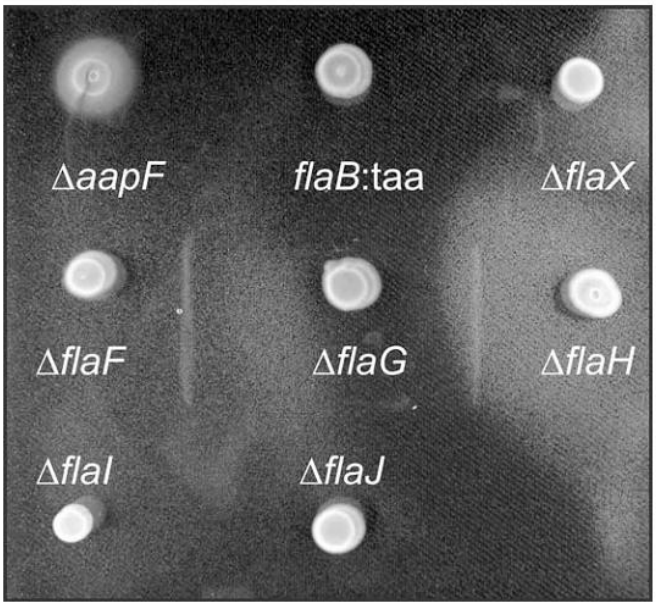


Fig. S4: Motility assay using semi-solid swarm plates containing 0.15% gelrite in Brock medium. Δ*aapF* is the positive control for motility. All the other deletion strains used are the double mutants Δ*aapF* Δ*fla*. For space reason only the *fla* mutation is indicated in the figure.

Movies M1-M3

M1: swimming motility of MW001 cells
M2: swimming motility of Δ*aapF* cells
M3: non-motile double deletion strain Δ*aapF flaB*:taa

3.2 Insights into FlaI functions in archaeal motor assembly and motility from structures, conformations, and genetics.

Summary:

In silico analysis predicted FlaI as the only ATPase in the archaellum system and classified this protein as a member of the "secretion ATPase superfamily" which groups ATPases found in bacterial type II/IV secretion and type IV pili systems. Due to the homology to T4P core components PilB/T and PilC respectively, FlaI and the polytopic membrane protein FlaJ are predicted to form analogously the assembly platform for the archaellum biosynthesis. FlaI from *S. acidocaldarius* was already confirmed as functional Mn^{2+}-dependent ATPase, which hydrolyses ATP in a co-operative manner and reaches optimum enzymatic activity at pH 6.5 at 75^{o}C (Ghosh et al., 2011a). It was also shown that FlaI functions as a hexamer and archaeal lipids have a stimulatory effect on its activity (Ghosh et al., 2011a). Here we present a detailed structural analysis of FlaI, which deepens our understanding of how FlaI operates within the archaellum complex. Using a full-length Walker B mutant, we were able to promote the oligomerisation of FlaI and solve the crystal structure of its hexameric form at 2.0 Å resolution. The FlaI hexamer forms a crown-like structure with an intrinsic 2-fold symmetry. The FlaI subunits within the hexamer, showed three different conformational states, assembled together in a rare cross-subunit interacting fashion. We have performed a structural comparison of FlaI to the type II secretion motor ATPase GspE, where we observed clear difference in the ATP-processing induced molecular dynamics. ATP hydrolysis causes in GspE motions directed perpendicular to the plane of the hexameric ring, whereas in FlaI the movement is directed parallel to the plane of the hexamer. This rotatory conformational change of FlaI explains the functional differences between FlaI and GspE and makes it potentially suitable to drive the archaella assembly and rotation.

It was known already, that FlaI is essential for the archaellum assembly (section 3.1), here we show however that not only the presence of FlaI is required, but also the hexamerisation and ATPase activity. Our study has also shown that the enzymatic activity and system specificity of FlaI are structurally separated. The ATPase activity is restricted to the C-terminal domain, whereas the N-terminal domain represents the functional part of FlaI, responsible also for the appropriate membrane location of this protein. In addition, we have proved that as the only functional ATPase in the archaellum, FlaI performs a dual role, generating energy for both, the archaellum assembly and its rotation. The deletion of first 29 amino acids of the N-terminus, led to uncoupling of these two functions, yielding an archaellated, but not motile *S. acidocaldarius* strain.

Contributions:

Sophia Reindl and Abhrajyoti Ghosh contributed equally. Both were involved in planning experiments and writing the manuscript, supervised by Sonja-Verena Albers and John A. Tainer. Abhrajyoti Ghosh constructed *in vivo* and *in vitro* mutants, purified FlaI and performed biochemical analysis. Sophia Reindl with help of Gareth J. Williams solved the crystal structure of FlaI, performed structural characterization and SAXS analysis. Kerstin Lassak, Tomasz Neiner and Anna-Lena Henche were involved in the biochemical and functional characterization of FlaI *in vitro* and *in vivo* and performed Electron Microscopy. All the authors revised the manuscript.

Molecular Cell
Article

Insights into FlaI Functions in Archaeal Motor Assembly and Motility from Structures, Conformations, and Genetics

Sophia Reindl,[1,2,4,5] Abhrajyoti Ghosh,[2,4] Gareth J. Williams,[1,4] Kerstin Lassak,[2] Tomasz Neiner,[2] Anna-Lena Henche,[2] Sonja-Verena Albers,[2,*] and John A. Tainer[1,3,*]
[1]Life Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA
[2]Molecular Biology of Archaea, Max Planck Institute for Terrestrial Microbiology, 35043 Marburg, Germany
[3]Department of Integrative Structural & Computational Biology, The Scripps Research Institute, La Jolla, CA 92037, USA
[4]These authors contributed equally to this work
[5]Present address: Department of Virology, Bernhard Nocht Institute for Tropical Medicine, 20359 Hamburg, Germany
*Correspondence: albers@mpi-marburg.mpg.de (S.-V.A.), jatainer@lbl.gov (J.A.T.)
http://dx.doi.org/10.1016/j.molcel.2013.01.014

SUMMARY

Superfamily ATPases in type IV pili, type 2 secretion, and archaella (formerly archaeal flagella) employ similar sequences for distinct biological processes. Here, we structurally and functionally characterize prototypical superfamily ATPase FlaI in *Sulfolobus acidocaldarius*, showing FlaI activities in archaeal swimming-organelle assembly and movement. X-ray scattering data of FlaI in solution and crystal structures with and without nucleotide reveal a hexameric crown assembly with key cross-subunit interactions. Rigid building blocks form between N-terminal domains (points) and neighboring subunit C-terminal domains (crown ring). Upon nucleotide binding, these six cross-subunit blocks move with respect to each other and distinctly from secretion and pilus ATPases. Crown interactions and conformations regulate assembly, motility, and force direction via a basic-clamp switching mechanism driving conformational changes between stable, backbone-interconnected moving blocks. Collective structural and mutational results identify in vivo functional components for assembly and motility, phosphate-triggered rearrangements by ATP hydrolysis, and molecular predictors for distinct ATPase superfamily functions.

INTRODUCTION

The assembly and movement of archaella (formerly termed archaeal flagella), bacterial type IV pili (T4P), and type 2 and type 4 secretion (T2S and T4S) systems are driven by superfamily ATPases. Both archaella and bacterial flagella function as rotating swimming organelles. However, archaella lack the hollow core of bacterial flagella and are assembled at the base (not the tip) by ATPase motors that are not homologous to bacterial flagella. Archaella are thus a unique motility structure with functional similarity to bacterial flagella but structural similarity to T4P (Ghosh and Albers, 2011; Pohlschroder et al., 2011; Trachtenberg and Cohen-Krausz, 2006). The archaellum (Jarrell and Albers, 2012) and T4P assembly machinery components (Bardy et al., 2004; Ng et al., 2006) are also homologous. Both bacterial prepilins and prearchaellins contain class III signal peptides processed by homologous T4P peptidases, PilD in bacteria and FlaK/PibD in archaea (Albers et al., 2003; Bardy et al., 2002). Moreover, secretion-system ATPase superfamily member FlaI and FlaJ membrane protein are homologous to PilT and PilB ATPases and the main transmembrane protein PilC in bacterial T4P and probably function as a core platform for assembly and rotation of the archaellum (Planet et al., 2001; Thomas et al., 2002). However, because of the limited biochemical and structural information for archaellum components, we lack a mechanistic understanding of archaellum assembly and motility.

Genetic manipulations in the genomes of archaellated archaea from the two major archaeal kingdoms, euyarchaeota and crenarchaeota, showed that all archaellum components are essential for assembly (Chaban et al., 2007; Lassak et al., 2012). As in haloarchaea, based on the conservation of archaella operons, archaellum movement is probably ATP dependent and not proton motive force driven (Streif et al., 2008).

Archaellum, bacterial T4P, T2S, and T4S assembly systems all use homologous ATPase motors (Craig and Li, 2008; Hansen and Forest, 2006) (Figure S1 available online). Those ATPases—PilE/PilF/TcpT/BfpD/CofH/LngH (T4P) (Chakraborty et al., 2008; Crowther et al., 2004; Gomez-Duarte et al., 2007; Tripathi and Taylor, 2007), GspE/XcpR/PulE/EpsE/ExeE/XpsE/OutE (T2S) (Filloux et al., 1998; Johnson et al., 2006; Russel, 1998; Shiue et al., 2006), VirB11 (T4S) (Savvides et al., 2003), or FlaI (archaellum) (Ghosh and Albers, 2011; Thomas et al., 2002)—are PilT/FtsK-secretory ATPases, a subgroup of RecA/Rad51-like motors (Iyer et al., 2004; Planet et al., 2001; Shin et al., 2003). In T4P systems, these ATPases act in assembly of the respective filaments (e.g., assembly ATPase PilB) and in disassembly or retraction of the pilus (e.g., retraction ATPase PilT) (Burrows, 2005). Structures of T4P filaments from *Pseudomonas*

Molecular Cell
Archaellum Assembly & Rotation by the ATPase FlaI

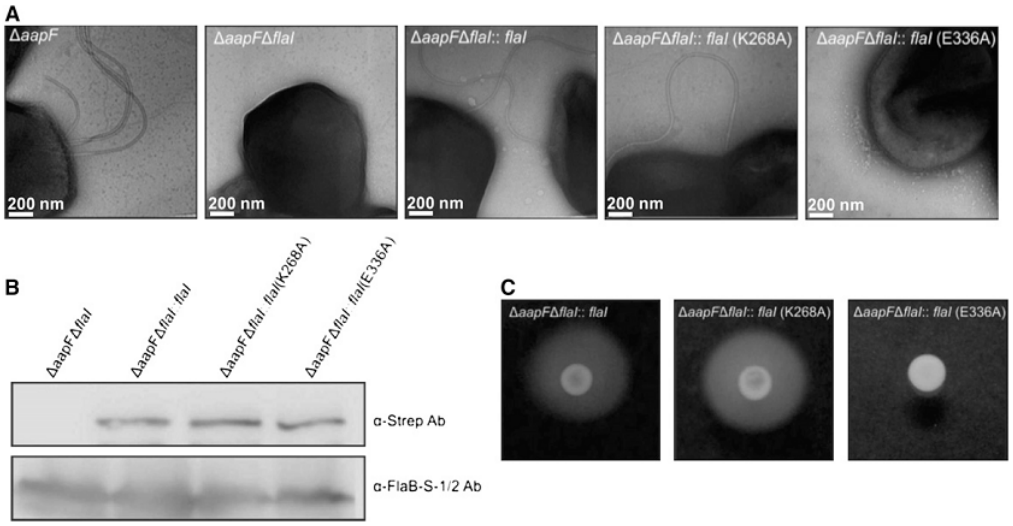


Figure 1. FlaI ATPase Activity Is Essential for Archaellum Formation

(A) EM images of the *S. acidocaldarius* cell surface (from left to right): The background is a hyperarchaellated strain (Δ*aapF*) that lacks pili. Deletion of the *flaI* gene results in no surface archaella. FlaI function is restored by introducing wild-type *flaI* or the Walker A mutant. The Walker B mutant strain is incapable of complementation, indicating the necessity of enzymatic ATPase activity for archaellum assembly.

(B) Wild-type FlaI, $FlaI^{K268A}$ Walker A, and $FlaI^{E336A}$ Walker B mutants are expressed in the Δ*aapF*Δ*flaI* mutant strain. Expression was detected by western blot, with the archaellin protein (FlaB) used as an endogenous control.

(C) Motility assays showing the ability of wild-type and FlaI variants to complement the motility phenotype in the Δ*aapF*Δ*flaI* strain. From left to right: Swimming behavior of the in trans complemented *ΔaapFΔflaI* with wild-type FlaI, FlaI Walker A^{K268A} mutant, and FlaI Walker B mutant. See also Figure S1 and Movies S1 and S2.

aeruginosa (Craig et al., 2003) and *Neissera gonorrhoeae* (Craig et al., 2006) suggest a model for T4P assembly and disassembly, which merits testing.

An added puzzle is how these sequence-similar T4P, T2S, and archaellum ATPases use ATP binding and hydrolysis to drive distinct processes: T4P for assembly and disassembly, T2S for the secretion of proteins, and archaella for assembly and filament motion. Informative crystal structures and models of T2S assembly ATPases (Abendroth et al., 2005; Chen et al., 2005; Robien et al., 2003; Yamagata and Tainer, 2007) and the T4P retraction ATPase (Misic et al., 2010; Satyshur et al., 2007) permitted few functional tests, because of the challenges of working with pathogens for genetic in vivo studies.

Fortunately, the crenarchaeon *Sulfolobus acidocaldarius* provides a prototypic system for investigating T4P, T2S, and archaellum ATPases, because genetic tools and nonpathogenicity aid genetic in vivo studies. Recent genetic tools for generating in-frame deletion mutants and functional in vivo experiments with proteins expressed in trans aid a more thorough and detailed analysis of the archaellum system (Albers et al., 2006; Wagner et al., 2009). Furthermore, FlaI, which shows ATP-dependent oligomerization and stimulation by archaeal tetraether lipids (Ghosh et al., 2011), provides an ATPase for examining both archaellum assembly and rotation.

To examine the molecular basis for FlaI ATPase dual activities and the functional diversity of this ATPase superfamily, we employed structural analyses by using crystallography and small-angle X-ray scattering (SAXS) in combination with in vitro and in vivo analyses. Collective results identified distinct residues and components essential for archaellum assembly and motility, a striking cross-subunit interface for the transmission of the ATPase force, intersubunit interactions to the active site, and an ADP-PO_4 release intermediate that together support a unified mechanistic understanding of filament assembly and motility functions.

RESULTS

FlaI Is the Assembly ATPase for *S. acidocaldarius* Archaella

The *S. acidocaldarius* Δ*aapF* strain lacks the membrane protein of the archaeal attachment-association pilus (aap) assembly system, resulting in a hyperarchaellated and hypermotile phenotype (Henche et al., 2012). Deletion of FlaI in this background (*ΔaapFΔflaI*) eliminates surface archaellation, implicating FlaI as the assembly ATPase (Lassak et al., 2012). To test and extend these observations and further test FlaI roles in archaellum assembly and filament movement, we complemented the Δ*aapF*Δ*flaI* strain in trans, using maltose-inducible expression vectors, with wild-type *flaI*, or the Walker A^{K268A} and Walker B^{E336A} *flaI* mutants (Figure 1). All strains were examined for archaella via electron microscopy (EM) (Figure 1A), for FlaI expression levels via western blot (Figure 1B), and for swimming in soft gelrite medium (Figure 1C) after inducing expression of complemented genes. Whereas the Δ*aapF*Δ*flaI* cell is nonmotile in swimming assays (Movies S1 and S2), EM revealed that

Molecular Cell
Archaellum Assembly & Rotation by the ATPase FlaI

Table 1. X-Ray Diffraction Data and Refinement Statistics

Data Set	FlaI^{E336A}-ADP (native)	FlaI^{E336A}-ADP (SeMet) peak/remote	FlaI-Apo
Diffraction Data Statistics			
Synchrotron beamline	ALS 8.3.1	ALS 8.3.1	ALS SIBYLS
Space group	C 2	C 2	P 6_3
Cell dimensions	a = 169.5Å; b = 148.1Å; c = 123.6Å α = γ = 90°; β = 131.6°	a = 169.6Å; b = 148.0Å; c = 124.4Å α = γ = 90°; β = 131.7°	a = b = 130.0Å; c = 311.8 Å α = β = 90°; γ = 120°
Wavelength (Å)	1.09	0.98/0.96	1.07
Resolution range (Å)[1]	46.0–2.0 (2.2–2.0)	57.9–2.5 (2.7–2.5)	49.9–3.6 (3.8–3.6)
Completeness (%)[1]	99.7 (95.6)	99.9 (97.6)	99.5 (98.3)
Observed reflections	5,453,245	2,088,354/2,170,286	1,135,996
Unique reflections	178,574	80,054	34,746
R_{sym} (%)[1,2]	7.1 (49.7)	5.4 (43.7)	9.8 (86.0)
I/σ[1]	25.9 (2.5)	36.4 (1.8)	17.7 (3.0)
Refinement Statistics			(refined as rigid body only)
Resolution range (Å)	46–2		50–3.6
R_{work}/R_{free} (%)[1,3]	18.5 (25)/21.4 (29)		29.6 (37)/31.8 (40)
No. protein atoms	12,565		16,448
No. ligand and ion atoms	98		–
No. water atoms	766		–
Average B factor	57.4		
Rmsd of bond lengths (Å)	0.008		
Rmsd of bond angles (°)	1.18		
PDB ID	4IHQ		4II7

[1]Values in parentheses are for the highest-resolution shell.
[2]R_{sym} is the unweighted R value on I between symmetry mates.
[3]R_{free} is the cross-validation R factor for 10% of reflections against which the model was not refined.

wild-type *flaI* and the *flaI* Walker A^{K268A} mutant could assemble archaea, and these complemented strains were motile in swimming assays (Figure 1). Importantly, complementation with the *flaI* Walker B^{E336A} mutation, which severely limits FlaI ATPase activity (Ghosh et al., 2011), does not support archaellum assembly or swimming. This essential role for FlaI ATPase activity in archaellum assembly motivated our detailed structural characterizations of this ATPase as the basis for archaellum assembly and swimming.

Domain Folds of the FlaI ATPase

For the examination of FlaI structures and conformations for insights into archaellum assembly, full-length Walker B mutant FlaI (FlaI^{E336A}) was purified and crystallized with ATP, which was found to be hydrolyzed to ADP in the structure. The crystal structure was solved in space group C2 via phasing with selenomethionine, allowing the building of contiguous polypeptide chains (minus the N-terminal methionines) for the three FlaI molecules in the asymmetric unit and refinement to a resolution of 2.0 Å (Table 1). Generation of the crystallographic hexamer using the 2-fold axis provided a detailed FlaI structure with bound nucleotide for comparisons to other superfamily ATPases (Figures 2A and 2B).

Each FlaI monomer is a bilobed structure consisting of a C-terminal ATPase domain (CTD) conserved among homologous ATPases and a much more variable N-terminal domain (NTD) (Figures 2 and S1). Both NTD and CTD form similarly sized globular structures with diameters of 45–55 Å, and a flexible linker connects the two domains (Figures 2A and 2B). The NTD forms two regions: NTD-1 (residues 1–126) and NTD-2 (residues 127–225). NTD-1 consists of a 29-residue α-β-α structure, followed by a three-stranded β sheet and a three-α-helix bundle. The electron density of NTD-1, and especially of the three-α-helix bundle, is less well defined, suggesting a high level of flexibility. NTD-2 has a well-defined electron density, consisting of two α helices and a large six-stranded β sheet.

The CTD also forms two subdomains. CTD-1 (residues 226–418) contains all the structural and functional features of a RecA-like ATPase, with Walker A and B and His box motifs. As in other RecA-like ATPases, it has a central β sheet core flanked on both sides by α helices. Two of the β strands are very long and extend toward the CTD of the neighboring subunit. CTD-2 (residues 419–C terminus) has two β strands and four α helices.

FlaI Hexameric Rings Are Built of Different Monomer Conformations

FlaI crystallized as a hexameric ring with intrinsic 2-fold symmetry resulting in three unique subunits. The overall hexamer shape resembles a crown: CTDs form the crown ring and NTDs form the crown points (Figure 2B). The FlaI crown has a diameter of 140 Å and a height from base to tip of 75 Å, with

Molecular Cell
Archaellum Assembly & Rotation by the ATPase FlaI

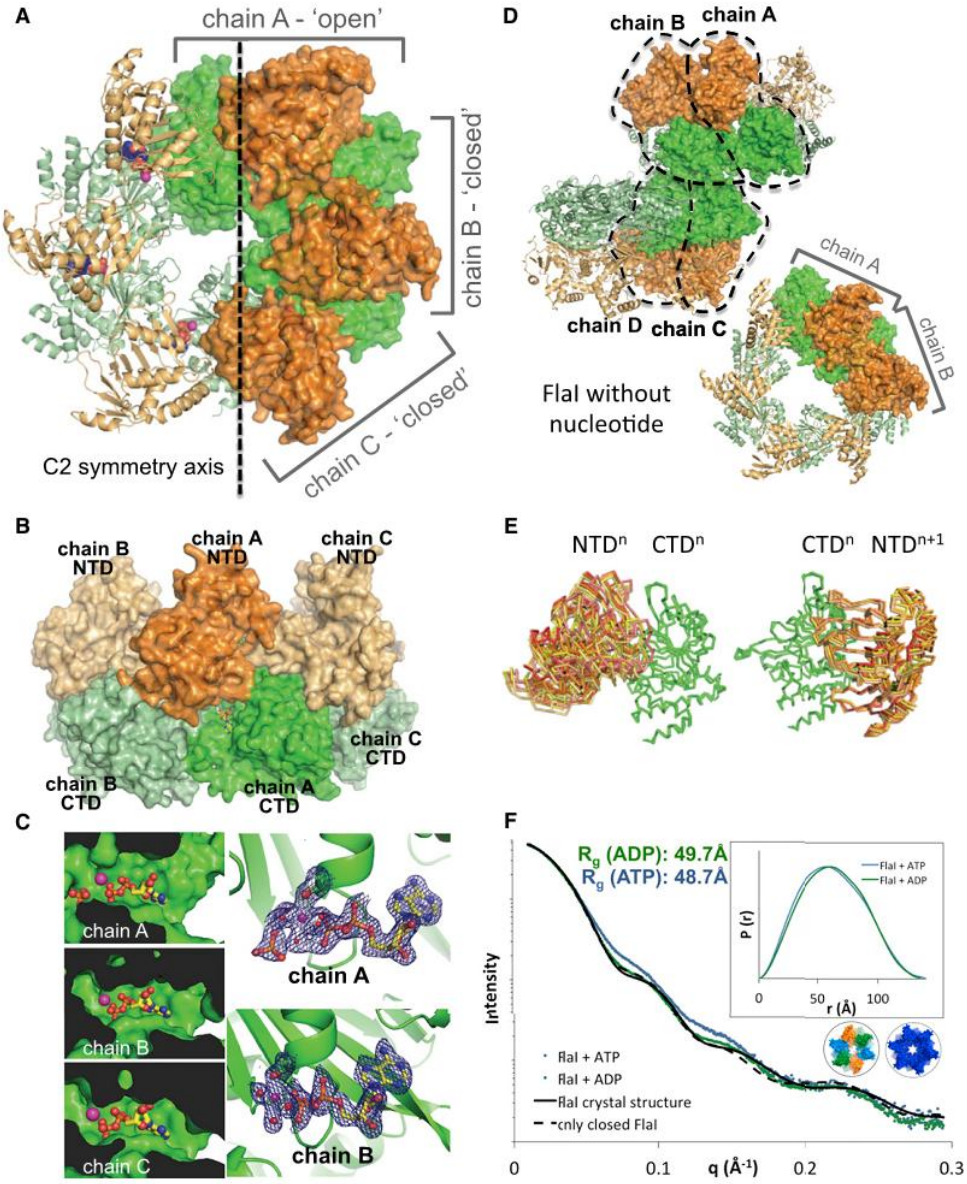


Figure 2. FlaI Crystal Structures Show Different Subunit Conformations

(A) Crystal structure of the Walker B mutant FlaI^{E336A} with NTDs (orange) and CTDs (green). A C2 symmetry axis divides the hexameric ring into two halves; three chains form the asymmetric unit (surface representation): chain A adopts a much more open conformation than chains B and C.

(B) The FlaI ring forms a crown-like structure. Side view of the FlaI^{E336A} hexameric ring shows the crown-like structure with the CTDs forming the base and the NTDs forming the points; nucleotide is shown as yellow sticks.

(C) The three active sites are not identical. The active site in chain A is more open and solvent accessible than in chains B and C. Electron density (2|Fo|-|Fc| map at 1.5 σ) for released phosphate is seen in chain A, but not for chains B and C, which have similar electron densities (chain B is shown as an example).

(D) Apo-FlaI without nucleotide also crystallized as a hexameric ring. Two subunits from two different hexameric rings form the asymmetric unit (surface representation), resulting in four unique subunit conformations. Colored as in (A).

(E) Comparison of the seven observed FlaI conformations reveals FlaI building blocks (three from ADP-FlaI and four from apo-FlaI). Left, alignment of the CTDs shows many different conformations of the corresponding NTDs. Right, building blocks of a CTD^{n} with NTD^{n+1} superimpose very well for all seven conformations.

(F) ATP-bound FlaI is more compact than ADP-FlaI. SAXS data of FlaI with ATP (blue) and ADP (green) shows a more compact structure for ATP-FlaI. Theoretical scattering curves for the crystal structure (black line) and an all closed model (dashed black line) are shown. Inset: P(r) distribution of the same data. See also Figure S2.

small and large openings formed by the CTD (9.5 Å) and NTD (75 Å) rings. Besides the linker, the CTD and NTD rings are connected within each FlaI monomer by a relatively small interface (buried surface area [BSA] of 365–400 $Å^2$). The FlaI monomers form the hexamer through much larger intersubunit interfaces, with the six-stranded β sheet of NTD-2 packing against the neighboring FlaI CTD, a CTD^n:NTD^{n+1} interaction, resulting in a BSA of ~1,240 $Å^2$.

Interestingly, the FlaI hexamer is asymmetric, and superimposition of the three full-length monomers shows that whereas chains B and C superimpose relatively well (root-mean-square deviation $[rmsd]_{C\alpha}$ = 0.68 Å), chain A adopts a dramatically different conformation ($rmsd_{C\alpha}$ = 2.4 Å to chain B and 2.8 Å to chain C). Superimposition of separated NTDs and CTDs shows that despite differences in full-length FlaI, the individual NTD and CTD structures are all similar, with only small changes in the NTD three-helix bundle at the crown points (average $rmsd_{C\alpha}$ of NTDs = 0.95 Å and CTDs = 0.75 Å). In contrast, the linker connecting the CTD and NTD is flexible, allowing a range of different conformations ($rmsd_{C\alpha}$ of individual subunits = 4.7 Å). Comparison shows that the FlaI hexamer bound to ADP contains two open active sites (from chain A) and four closed and buried active sites (from chains B and C) (Figures 2A–2C and Movie S3).

To compare nucleotide-bound and -free FlaI structures, we also crystallized wild-type FlaI without nucleotide, in space group $P6_3$, and solved the structure by molecular replacement using PHASER (McCoy, 2007). The CTD structure of ADP-bound $FlaI^{E336A}$ was used as a first search model, followed by a search with the NTD structure. The structure was refined to a resolution of 3.6 Å via rigid body refinement (Table 1). Although FlaI exists as a monomer in solution without nucleotide (Ghosh et al., 2011), it forms hexameric rings in our structure, probably due to the high protein concentration within the crystal. The asymmetric unit consists of four unique subunits as part of two different hexameric rings (Figure 2D). Similarly to the ADP-bound structure, the flexible linker connecting the NTD and CTD adopts four different conformations; however, the hexameric arrangement is different, with alternating open and closed active sites. Together with the three conformations seen in the ADP-bound structure, this gives a total of seven different FlaI conformations.

Comparing the seven structurally different FlaI subunits by superposition of their CTDs shows that the NTDs adopt multiple conformations (Figure 2E, left). Similarly, the intrasubunit interface between NTD and CTD is small compared to that between one NTD and the neighboring subunit CTD. Therefore, we compared building blocks consisting of a CTD plus the tightly bound NTD from the next subunit. Superposition of the CTD^ns of these building blocks shows that the NTD^{n+1}s all align with each other (Figure 2E, right).

An ADP-PO_4 Release Intermediate Is Visible in Two Active Sites of the Hexamer

Nucleotide binding occurs at the interface between two FlaI subunits (Figures 2A and S2). Given that $FlaI^{E336A}$ shows strongly reduced ATPase activity (Ghosh et al., 2011), we expected to trap an ATP-bound state when we crystallized it with ATP and $MgCl_2$. However, the electron density in the active site showed that ADP and Mg^{2+} were bound in all subunits (Figure 2C), indicating ATP hydrolysis during crystallization. Interestingly, the major difference between FlaI subunits is the rotation of one molecule (chain A), resulting in an open active site compared to a closed conformation for chains B and C. Closer examination of the chain A conformation revealed extra electron density, consistent with a phosphate moiety (Figures 2C and S3). The phosphate is 4 Å from the active site Mg^{2+} and 4.2 Å from the ADP β-phosphate, suggesting a release intermediate that has not previously been seen in this superfamily. ADP and Mg^{2+} binding residues are similar within the CTDs of all three subunits, with the only major rearrangement being a dramatic conformational switch of Arg338, which moves to coordinate the released phosphate in chain A (Figure 3). Importantly, the most dramatic difference between chain A and chains B/C is the rotation of the NTD. The NTD of chain A rotates ~20°, and due to their extensive shared interface this conformational rearrangement shifts the chain B CTD toward the chain A active site (Figure 3A). The net result is the fitting of a chain B α helix into a groove near the chain A active site, positioning two basic residues into hydrogen-bonding distance of the released phosphate: Arg326 is 3.6 Å away and Lys322 is 3.4 Å away (Figure 3B). These residues are conserved throughout most of this ATPase superfamily (Figure S1).

Solution Analysis Reveals ATP Binding Induces a More Compact Form of FlaI

Experiments aiming to soak ATP into the crystals for obtaining ATP-bound FlaI failed because the crystals dissolved, suggesting conformational differences between ATP- and ADP-bound states. To test the differences between ATP- and ADP-bound states in solution, we performed SAXS experiments with $FlaI^{E336A}$ using either nucleotide (Figure 2F). The Walker B mutation significantly slows the rate of ATP hydrolysis, allowing the ATP-bound state to be observed during the short time frame of the SAXS experiment. The experimental scattering curves for $FlaI^{E336A}$-ADP and $FlaI^{E336A}$-ATP are clearly different. The calculated theoretical scattering curve for the crystal structure of $FlaI^{E336A}$-ADP (Figure 2F, solid black line) overlaps well with the equivalent experimental scattering curve, with an improved fit obtained by generating an all-closed 6-fold symmetric FlaI ring (dotted black line) with SymmDock, a geometry-based docking program (Schneidman-Duhovny et al., 2005). The FlaI-ATP pair distribution (P(r)) function, which provides experimental electron-pair distances to accurately compare structures by SAXS (Putnam et al., 2007), is shifted to smaller distances compared to FlaI-ADP, particularly between 20 and 90 Å, indicating a more compact structure. Collectively, the solution and crystal structures suggest that ATP binding locks the FlaI hexamer into a more symmetrical and less dynamic conformation, which may promote hexamer assembly by stabilizing interactions across adjacent subunits.

Archaellum Assembly Requires FlaI Hexameric Assembly

FlaI undergoes hexameric assembly in solution upon ATP binding (Ghosh et al., 2011), as do other secretion ATPase superfamily members. However, a key question is whether this

Molecular Cell
Archaellum Assembly & Rotation by the ATPase FlaI

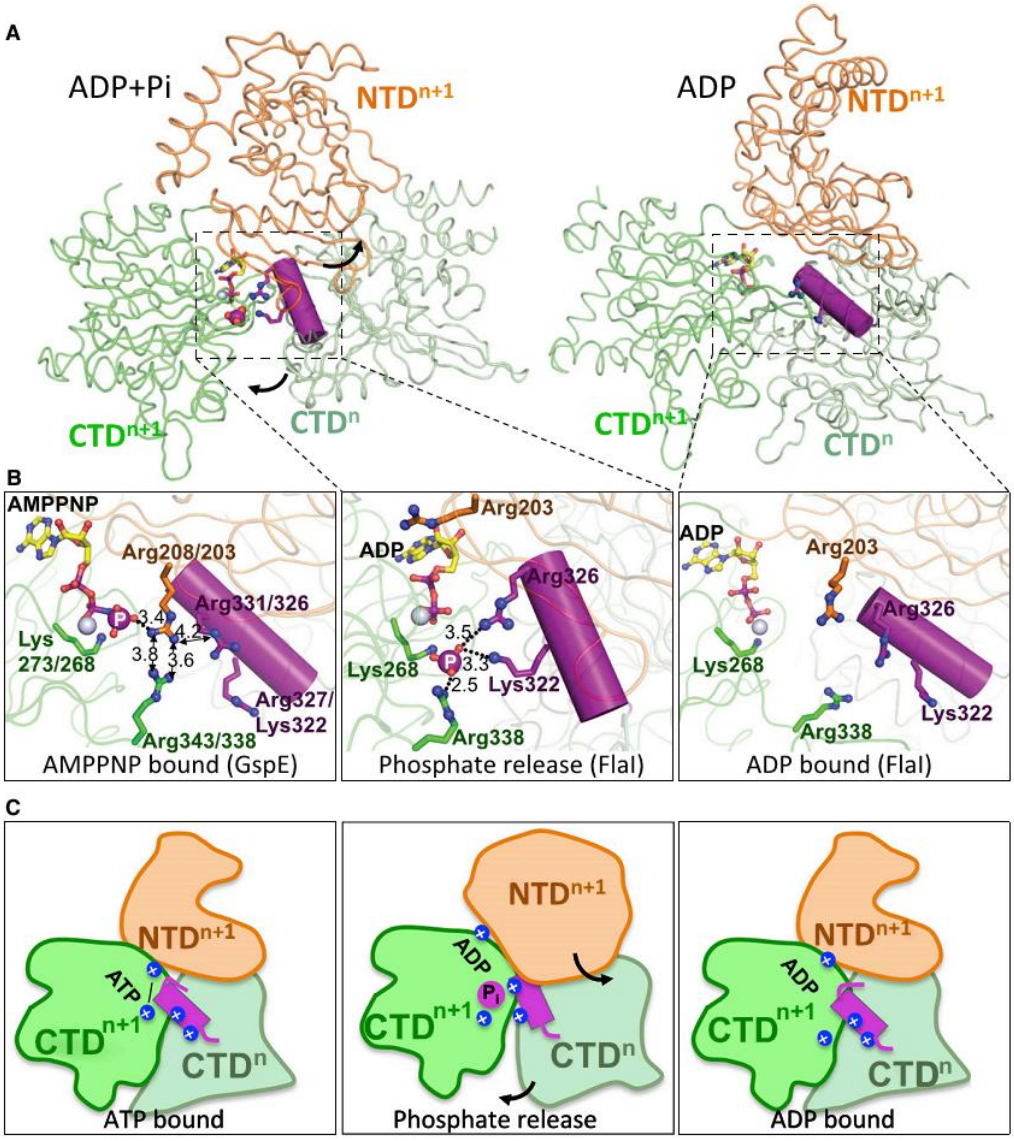


Figure 3. Basic-Clamp Switching Drives Secretion Family ATPase Conformational Changes

(A) Overview of conformational changes between FlaI ADP plus inorganic phosphate (ADP+Pi) and ADP states. A CTD^{n} helix containing a basic clamp (purple) enters the neighboring active site to coordinate the released Pi in the ADP+Pi state (left). Phosphate (purple sphere) dissociation results in conformational rotation of CTD^{n} toward CTD^{n+1} and NTD^{n+1} away from CTD^{n+1} (arrows), with the basic-clamp helix moving away from the active site (right). Boxed regions are shown in (B).

(B) Close up of ATP hydrolysis and Pi release-driven conformational changes. The GspE structure bound to AMPPNP (left) shows close proximity (distances are in Å) of intramolecular and intermolecular Arg clamps, probably stabilized by Arg203 hydrogen bonding to the γ-phosphate. In this state the intermolecular basic clamp (purple) on CTD^{n} is outside the active site. Residues are numbered for GspE and FlaI. Upon ATP hydrolysis (middle), the basic-clamp helix rotates into the active site, with FlaI Arg326 and Lys322 of CTD^{n} displacing Arg203 of NTD^{n+1} to make intersubunit interactions to the released phosphate (large purple sphere). Residues are numbered for FlaI. Upon phosphate release (right), the basic clamps reset, with the basic-clamp helix moving out of the active site and NTD^{n+1} Arg203 moving back into the active site, positions similar to those seen in the AMPPNP state, ready for subsequent rounds of ATP binding and hydrolysis. Residues are numbered for FlaI.

(C) Schematic showing the conformational changes observed between ATP-bound (left), phosphate-release intermediate (middle), and ADP-bound (right) states for FlaI and superfamily members. Colors are as in (A), and basic-clamp residues are highlighted with blue spheres. See also Figure S3.

Molecular Cell
Archaellum Assembly & Rotation by the ATPase FlaI

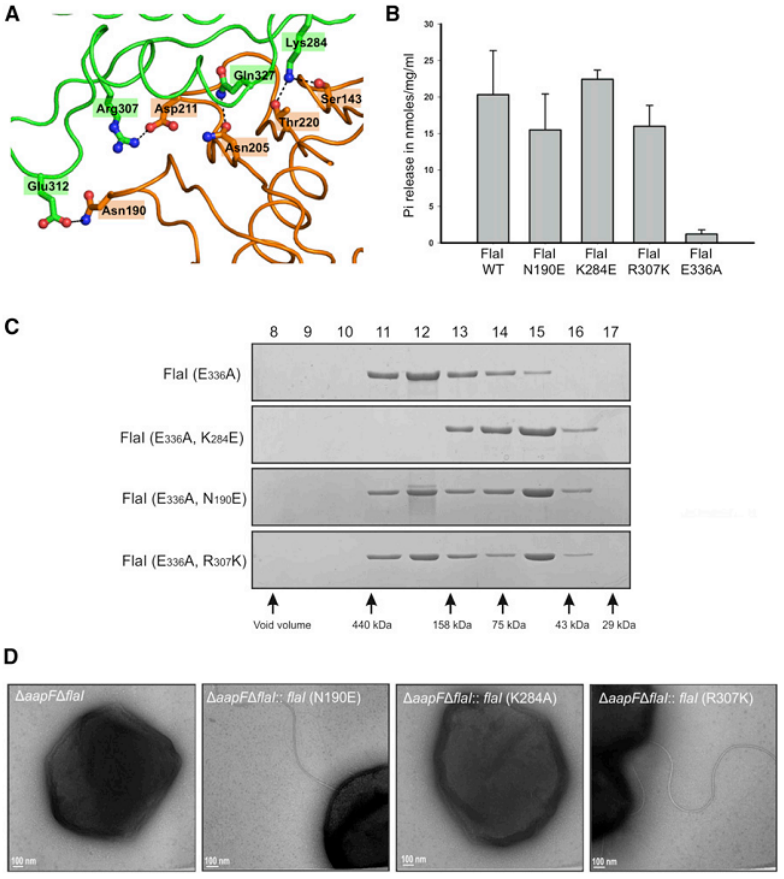


Figure 4. FlaI Hexameric Conformation Is Necessary for Function

(A) Residues involved in salt bridges between FlaI subunits were mutated.

(B) Interface mutants are active ATPases. ATPase activity was analyzed by measuring released Pi upon ATP hydrolysis at 70°C (mean of three independent experiments ± SD). WT, wild-type.

(C) Analytical gel filtration shows the effect of interface mutation on hexamerization. Analysis of interface mutations (FlaI$^{K284E/E336A}$, FlaI$^{N190E/E336A}$, and FlaI$^{R307K/E336A}$) shows that these interface mutants have a reduced ability to form hexamers upon ATP binding compared to FlaI^{E336A}.

(D) Complementation analysis of the Δ*aapF*Δ*flaI* mutant strain with FlaI interface mutants. Expression of FlaI interface mutants in trans in a Δ*aapF*Δ*flaI* strain revealed partial complementation phenotypes for FlaI^{N190E} and FlaI^{R307K}. However, FlaI^{K284E} was unable to complement the mutant phenotype (Δ*aapF*Δ*flaI*), indicating that FlaI needs to be in its hexameric form to be functional.

hexameric assembly is indispensible for the association and functionality of their respective macromolecular machines. To test whether ATP-dependent oligomerization is important for archaellum assembly in *S. acidocaldarius*, we made three FlaI mutations, designed to disrupt subunit interactions based on our structures. Residues in salt bridges at the haxameric interface of two subunits were mutated: Asn190 to Glu, Lys284 to Glu, and Arg307 to Lys (Figure 4A), and the FlaI variants were characterized in vitro and in vivo.

We made these mutations in wild-type *flaI*, to analyze the effect on ATPase activity, and also in the Walker B^{E336A} *flaI* mutant, to test the impact on oligomerization. Purified interface mutants in the wild-type background were analyzed for ATP hydrolysis. All three point mutants were active ATPases (Figure 4B), showing that the ATP binding and hydrolysis motifs conserved within each FlaI monomer are sufficient for ATPase activity. The interface mutants in the WalkerB^{E336A} background were tested for oligomerization by analytical gel-filtration analysis. For oligomerization experiments, the WalkerB^{E336A} background is necessary for disabling the ATPase activity and trapping the structure in the hexameric state. Self-assembly into hexameric rings was decreased for all mutants compared to FlaI^{E336A} (Figure 4C). The most striking effect was evident for FlaI$^{K284E/E336A}$, which was unable to form stable hexamers even with ATP.

To test the functional impact of disrupting FlaI oligomerization, we performed in trans complementation of the Δ*aapF*Δ*flaI S. acidocaldarius* strain with each point mutant. The ΔaapF mutant is a hyperarchaellated strain that therefore gives stronger motility phenotypes (Henche et al., 2012). EM analysis of complemented strains revealed that FlaI^{N190E} and FlaI^{R307K} restored the wild-type phenotype, whereas FlaI^{K284E} did not, with only 6% of cells overexpressing FlaI^{K284E} assembling archaella (Figure 4D). These collective analyses indicate that FlaI must be able to assemble into its hexameric state to function as the archaellum-assembly ATPase.

Covalently Interlocked Blocks Move within the Ring to Make FlaI a Stable Motor

How GspE and FlaI achieve their different physiological functions in protein secretion and movement has been a puzzle. Comparison of FlaI and GspE conformations (Figure 5) shows that the directionality of movement induced by ATP binding, hydrolysis, and release differs. FlaI moves with a rotating translation parallel and across the plane of the hexameric ring (Movie S3), a movement suitable to drive both assembly and rotation of archaella. In contrast, GspE moves up and down perpendicular to the plane of the ring, with little rotary movement (Movie S4). The building block seen for FlaI and GspE is also conserved across other superfamily ATPases (Figures 5C and S4).

The FlaI NTD Is Essential for Archaellum Formation, but Not ATPase Activity

The most variable region of the T2S/T4S/T4P ATPase superfamily is the NTD. In this domain the T4P retraction ATPase

Molecular Cell
Archaellum Assembly & Rotation by the ATPase FlaI

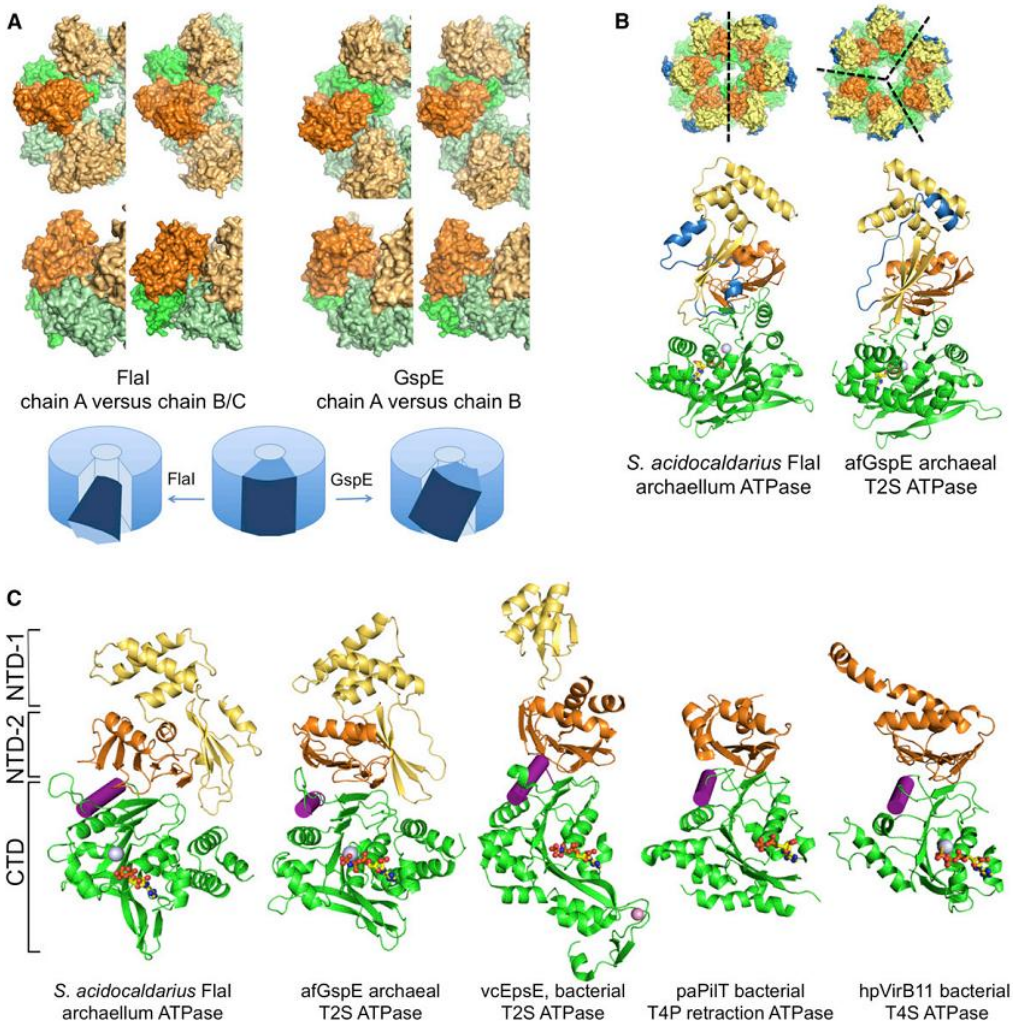


Figure 5. FlaI Compared to Other Secretion-Family ATPases

(A) Movements induced by ATP hydrolysis differ between FlaI and GspE. The two building blocks for comparison are shown in the same orientation within the hexameric ring. For FlaI, NTD^{chainA} + CTD^{chainB} is compared to NTD^{chainB} + CTD^{chainC}; for GspE, the only two building blocks present in the structure are shown. The different building-block conformations in the FlaI and GspE rings allow us to propose domain movements during ATPase hydrolysis. The directionality of movement differs between FlaI and GspE (top, top view; bottom, side view).

(B) The largest structural difference between FlaI and GspE is in residues 1–29 (blue). This region forms the outer side surface of the crown-like structure (top). Crystal-related symmetry axes are indicated as black dashed lines.

(C) Like FlaI, other secretion-system ATPases form building blocks. CTDs (green) are shown with NTD-2 (orange) and, when present, the NTD-1 (yellow) of the neighboring subunit with nucleotides (sticks), Mg^{2+} (silver), and Zn^{2+} ions (pink spheres). The α helix mediating the block movement described in Figure 3 is highlighted (purple). See also Figure S4 and Movies S3 and S4.

PilT (Misic et al., 2010) and the T4S ATPase VirB11 (Savvides et al., 2003) lack the NTD-1 that is present in the T2S ATPases GspE (Yamagata and Tainer, 2007), EpsE (Abendroth et al., 2005; Robien et al., 2003), and FlaI (Figure 5C). Although the overall fold of FlaI NTD-1 resembles GspE NTD-1, it is unrelated to the fold of the EpsE NTD-1. Full-length EpsE has yet to be crystallized, but EspE1 NTD-1 has been crystallized in complex with the inner-membrane protein EpsL cytoplasmic domain (Abendroth et al., 2005; Robien et al., 2003). Collectively, these results suggest that the superfamily NTD domains link the conserved ATPase domain to the distinct biological systems.

To investigate the structurally implicated biological roles of the FlaI NTD, we generated a Δ*aapF*ΔN*flaI* strain, which lacks FlaI amino acids 1–224 (yellow and orange in Figure 5C). EM analysis of the Δ*aapF*ΔN*flaI* deletion mutant revealed that no archaellum was assembled (Figure 6A) and swimming motility was abolished

Molecular Cell
Archaellum Assembly & Rotation by the ATPase FlaI

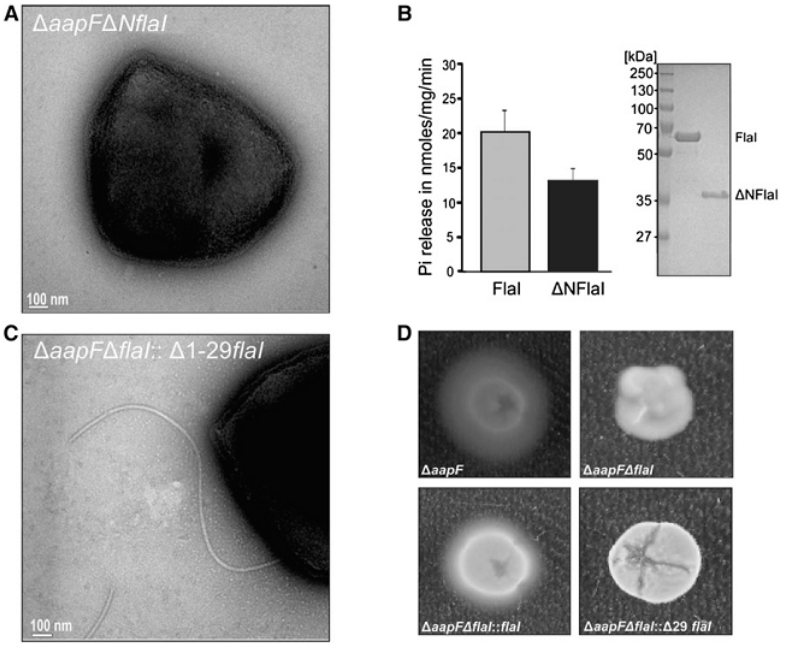


Figure 6. The FlaI NTD Is Necessary for Archaellum Assembly, but Not ATPase Activity

(A) Deletion of FlaI residues 1–224 (Δ*aapF*Δ*NflaI*) results in a phenotype similar to that of *flaI* deletion (Δ*aapF*Δ*flaI*).

(B) The CTD alone is an active ATPase. ATPase assays comparing purified full-length FlaI are compared to the CTD alone (ΔNFlaI). ΔNFlaI showed 75% activity compared to wild-type, indicating that the Δ*aapF*Δ*NflaI* phenotype cannot be explained by a loss of enzymatic activity (mean of three independent experiments ± SD).

(C) FlaI residues 1–29 are not essential for archaellum assembly. Complementation analysis of Δ*aapF*Δ*NflaI* mutant strain with $FlaI^{\Delta 1-29}$ shows archaellum assembly in the complemented strain (10%–20% of cells showed flagella assembly).

(D) Motility assays showing the inability of $FlaI^{\Delta 1-29}$ to restore the motility phenotype in the Δ*aapF*Δ*NflaI* strain. From left to right: Swimming of the hyperarchaellated background strain (Δ*aapF*) was observed in Gelrite plates, whereas a deletion of FlaI resulted in nonmotile cells (Δ*aapF*Δ*NflaI*). Complementation of the strain with wild-type FlaI (*flaI*::Δ*aapF*Δ*NflaI*) restored the wild-type phenotype, whereas strains complemented with $FlaI^{\Delta 1-29}$ had no motility. See also Movie S5.

(Movie S5), indicating a central role of the NTD in archaellum assembly. To test whether this functional effect was due to a loss of FlaI ATPase activity from NTD deletion, we expressed and purified ΔNFlaI from *Escherichia coli*. The ΔNFlaI protein still exhibits 75% of ATPase activity compared to the full-length FlaI, indicating that assembly failure in the Δ*aapF*Δ*NflaI* deletion mutant was not due to the loss of ATPase activity (Figure 6B): the Δ*aapF*Δ*NflaI* deletion provides a separation-of-function mutation for ATPase versus assembly and motility.

The First 29 Residues of FlaI Are Essential for Motility, but Not Assembly

According to our structures and analyses, a striking difference between the FlaI and GspE crystal structures is in amino acids 1–29 (Figure 5B). In FlaI these residues form a surface-exposed α-β-α structure (Figures 5B and S1), located on the outside of the hexamer farthest away from the center. We hypothesized that this subdomain is a specific adaptor that links the ATPase to archaellum assembly or motility. To test this, we generated a construct of FlaI in which amino acids 1–29 were deleted ($FlaI^{\Delta 1-29}$), expressed this variant in a *S. acidocaldarius* Δ*aapF*Δ*flaI* background, and assayed for archaellum assembly and motility. EM analysis of the strain overexpressing $FlaI^{\Delta 1-29}$ revealed that 10%–20% of the cell population can assemble archaella, whereas complementation with wild-type FlaI leads to 40%–50% of cells having archaella (Figure 6C), indicating a nonessential role of the amino acids 1–29 in assembly. However, unlike hyperarchaellated Δ*aapF* and *flaI*-complemented Δ*aapF*Δ*flaI* strains, which show motility under standard assay conditions, the $FlaI^{\Delta 1-29}$-expressing Δ*aapF*Δ*flaI* strain showed no motility, similar to the nonmotile Δ*aapF*Δ*flaI* strain (Figure 6D). Together, these analyses indicate a role of FlaI amino acids 1–29 in motility rather than assembly of the *S. acidocaldarius* archaellum: $FlaI^{\Delta 1-29}$ provides a separation-of-function mutation for assembly versus motility.

The FlaI Three-Helix Bundle Is Required for Membrane Localization and Archaellum Assembly

FlaI and the membrane protein FlaJ form the key platform on which the archaellum assembly takes place. It was suggested that *Methanocaldococcus* sp. FlaI interacts with the membrane and membrane-anchored archaellum assembly components (Thomas and Jarrell, 2001). Together with our finding that FlaI NTD is not responsible for enzymatic activity and the structure of the T2S ATPase EpsE-NTD in complex with membrane protein EpsL (Abendroth et al., 2005), we hypothesized that the FlaI ring-top surface, formed by NTD-1, may constitute the membrane and/or membrane protein interaction surface. Mapping crystallographic temperature factors (B values) onto the FlaI structure, as proven useful for observing differences in mobility with functional implications (Tainer et al., 1984), identifies a three-helix bundle (amino acids 61–125) that has higher B factors than the rest of the protein (Figure 7A), suggesting increased local motion. This three-helix bundle forms the tip of the FlaI crown and is placed so as to interact with components of the membrane.

We examined the effect of removing this three-helix bundle from FlaI on archaellum assembly by expressing the $FlaI^{\Delta 61-125}$ internal deletion in trans in the *ΔaapFΔflaI* strain. EM analysis revealed that loss of the three-helix bundle abolishes archaella assembly (Figure 7A). Furthermore, cellular fractionation confirmed that the $FlaI^{\Delta 61-125}$ protein was unable to localize to the membrane, indicating a role of the three-helix bundle in mediating protein-protein and/or protein-membrane interactions (Figure 7A).

Molecular Cell
Archaellum Assembly & Rotation by the ATPase FlaI

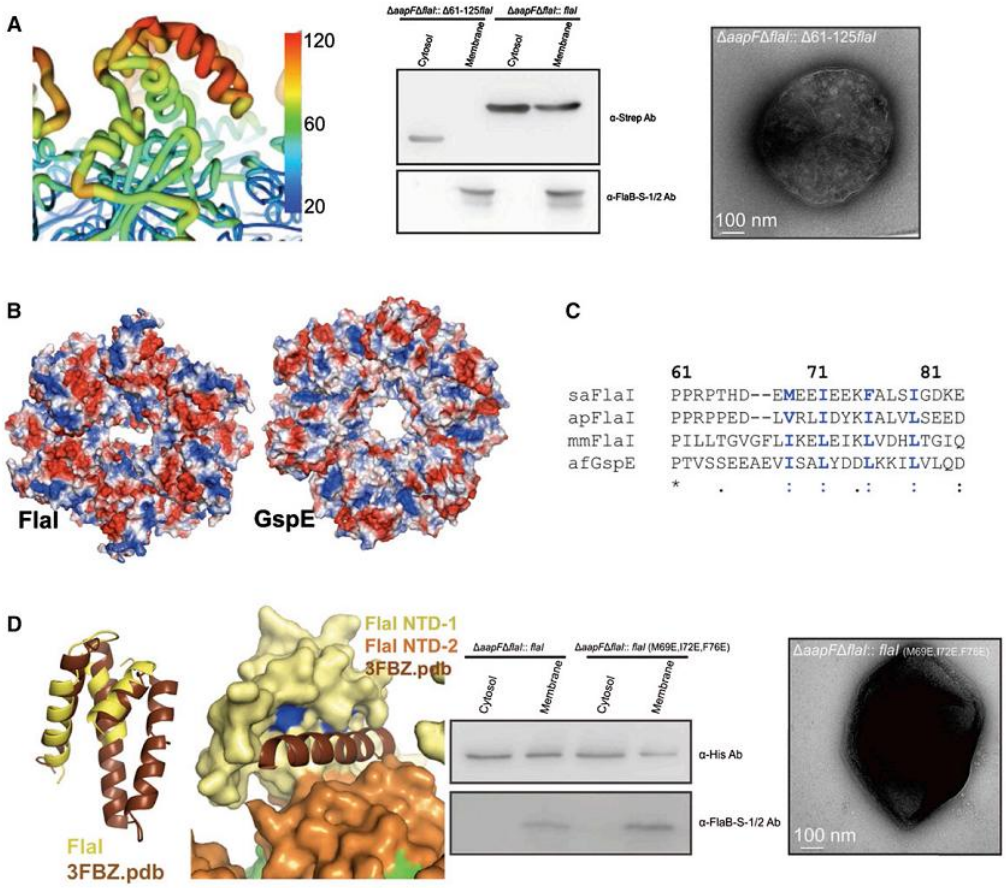


Figure 7. N-Terminal Helix Bundle Domain Interaction Site

(A) A three-helix bundle is essential for archaellum formation and FlaI localization. Left: a strongly charged patch is the three-helix bundle with relatively high B factors (red with thicker ribbon), indicative of intrinsic flexibility. Middle: FlaI missing the helix bundle domain (FlaI$^{\Delta 61-125}$) does not localize to the membrane. Protein presence in membrane and cytoplasmic fractions is shown by western blot. Right: the deletion of the helix bundle domain in vivo results in nonarchaellated cells.

(B) Electrostatic surface (blue, positive; red, negative) shows differences between FlaI and GspE.

(C) Conserved, surface-exposed hydrophobic residues in the helix bundle domain (blue) are good candidates for a protein-protein interaction surface.

(D) Dali search for the three-helix bundle identifies several structural domains that consist of an additional fourth helix. Left: superposition of the Dali result (PDB ID code: 3FBZ) positions the fourth helix directly in the groove between NTD-1 and NTD-2 and close to the conserved hydrophobic residues (blue). Mutation of these residues (Met79, Ile72, and Phe76) to glutamates results in both decreased FlaI association with the membrane, as compared to wild-type FlaI in our membrane-association assays (middle), and a nonarchaellated phenotype (right). See also Movie S6.

To further examine interaction surfaces in the FlaI$^{61-125}$ region, we did more detailed structural analysis. Comparison of the FlaI and GspE NTDs shows that the charge distribution differs greatly on the top surface where this three-helix bundle is located (Figure 7B). In FlaI this surface is positively charged, and we expect these regions will most likely interface with the membrane-lipid negatively charged phosphates. In contrast, stable protein-protein interaction sites commonly involve surface-exposed, conserved hydrophobic residues. Examination of our FlaI NTD structures identifies three hydrophobic surface-exposed residues (Met69, Ile72, and Phe76) that map to the three-helix bundle and are conserved among all the FlaIs in archaea (Figure 7C). Furthermore, a Dali protein-structure database search for structural homologs (Holm and Rosenström, 2010) using only the three-helix bundle identified predominantly four-helix bundles; the best hit is shown in Figure 7D (Protein Data Bank [PDB] ID code: 3FBZ). Interestingly, superimposition of this four-helix bundle with FlaI positions the fourth helix precisely into a groove formed within the three-helix bundle that is lined with the conserved hydrophobic residues.

To test the structurally implicated FlaI-FlaJ interface, we constructed a FlaI$^{M69E/I72E/F76E}$ triple-point mutant and expressed this in trans in the Δ*aapF*Δ*flaI* deletion strain. EM analysis revealed no archaellum assembly and swimming assays showed

no motility (Movie S5) in the complemented strain overexpressing the triple-point mutant. Furthermore, membrane localization assays showed that FlaI$^{M69E/I72E/F76E}$ localization to the membrane was significantly reduced compared to that of wild-type (Figure 7D). Collectively, these results support an important role of these residues in archaellum assembly, most likely through mediating protein-protein interactions.

DISCUSSION

Combined structural and functional analyses of the prototypic archaeal motor ATPase FlaI suggest the basis for its dual function in archaellum assembly and movement, which further informs superfamily functions. We found that *S. acidocaldarius* FlaI ATPase activity and hexameric assembly are essential for archaellum assembly and identified N-terminal FlaI regions important for archaellum motility, assembly, and FlaI localization. Moreover, our crystal structure of FlaI bound to nucleotide traps a previously unobserved intermediate that informs the mechanistic basis for conformational changes characteristic of this superfamily.

Arg-Clamp Switches Drive Building-Block Movements

The seven FlaI conformations in our crystal structures expand the number of conformations observed for this ATPase superfamily. GspE (Yamagata and Tainer, 2007), PilT (Misic et al., 2010; Satyshur et al., 2007), and VirB11 (Savvides et al., 2003) hexameric structures have two or more conformations, and symmetric hexamers of VirB11 and PilT have also been seen. Dynamic hexamer formation is therefore characteristic of this superfamily, and the propensity to form symmetric versus asymmetric states may reflect differences in synchronization of ATP binding and hydrolysis between subunits. Analyses of VirB11 and GspE led to suggestions they act as motors by linking a rigid basal ring formed by the ATP binding CTD with moving NTDs on top (Savvides et al., 2003; Yamagata and Tainer, 2007). Yet we find that in all superfamily members with known structures, neighboring NTDs and CTDs form a tight interface (Figure 2E) and act as a moving building block, supporting and extending a mechanism proposed for PilT (Satyshur et al., 2007). We propose that this "flexible crown" structure is key for FlaI motor functions that drive archaellum assembly in an ATP-hydrolysis-dependent manner. Furthermore, differences in building-block movements driven by ATP binding and hydrolysis may explain why similar proteins can provide the energy for distinct biological processes such as protein secretion versus archaellum assembly and movement.

Interestingly, in our crystal structure of FlaI bound to nucleotide, one subunit is trapped in a phosphate-release state. In this conformation, basic residues from the neighboring CTD subunit protrude into the active site cleft and coordinate with the released phosphate. Previous low-resolution PilT observations implicated analogous interactions, involving PilT Arg203 and Arg207, and suggested that this could be similar to the F_1F_0ATPase Arg finger (Satyshur et al., 2007). Our results thus test and extend these observations, providing direct evidence of intersubunit communication to the nucleotide in the active site, and suggest that phosphate coordination immediately following ATP hydrolysis is a switch that triggers building-block rearrangements. Closed conformations of GspE and PilT (Misic et al., 2010; Satyshur et al., 2007; Yamagata and Tainer, 2007) have identified an intrasubunit Arg clamp that binds the ATP γ-phosphate. In our ADP-bound FlaI structures these clamp residues, Arg338 in the CTD and Arg203 in the NTD, are not in an active conformation. However, consistent with our comparison to GspE, wherein the Arg-clamp residues must move closer together to bind the ATP γ-phosphate, our SAXS data supports a model wherein the FlaI subunits compact further to adopt a closed conformation when bound to ATP. Surprisingly, in GspE these Arg residues are in close proximity (<4 Å) and would repel each other without stabilization from the γ-phosphate (Figure 3B). Furthermore, Arg203 is close (4.2 Å) to Arg326 of the neighboring CTD. Collectively, this leads us to hypothesize a switch model wherein hydrolysis of ATP for releasing the γ-phosphate would move the Arg-clamp residues into closer proximity, leading to their repulsion (Figure 3). Because the NTD is no longer anchored to the nucleotide, this repulsion would drive the conformational rotation of the NTD and the concerted movement of the neighboring CTD to which it is interfaced, positioning the second conserved basic clamp (Arg326 and Lys322 in FlaI) to intermolecularly bind to the released phosphate. Interestingly, this movement opens up the back face of the FlaI hexamer suitable for ADP dissociation (presumably, a high nucleotide concentration allows ADP to remain in our crystal structure). Due to the high concentration of positive charge around the phosphate, this intermediate may be metastable (temperature trapped for the thermophile), and repulsion from these residues would drive further subunit movements and subsequent phosphate release so that the cycle of ATP binding, hydrolysis, and product release can continue efficiently for archaellum assembly and motility. Intriguingly, another insight from our structure is an intramolecular hydrophobic interface formed between the basic-clamp helix of the nucleotide binding subunit with secondary structure elements surrounding Arg338, which is in close contact with the leaving phosphate. This interface provides another potential mechanism by which ATP hydrolysis between subunits is coordinated: the basic-clamp helix may sense the phosphate-release state through its connection to Arg338.

The NTD Links Building-Block Movements to Archaellum Assembly and Motility

The FlaI NTDs, which form the crown points, are strategically located on the outside edge to readily form functional interactions (Figures 5B and 7A): striking differences between FlaI and GspE NTDs indicate this is an area that interacts with the distinct membrane components in each system. Interestingly, we found that in FlaI the first 29 residues, on the crown tip sides, are required for driving archaellum motility, whereas the three-helix bundle, at the crown point, localizes FlaI to the membrane and is required for archaellum assembly and motility. Furthermore, the three-helix bundle is a probable interaction surface for FlaJ, given that at physiological pH (*S. acidocaldarius* intracellular pH is ~6.5) the three-helix bundle is negatively charged (isoelectric point [pI] of 5.75) and the conserved cytoplasmic domains of FlaJ are positively charged (pI of 9.75 and 10.3).

Molecular Cell
Archaellum Assembly & Rotation by the ATPase FlaI

FlaI and FlaJ interactions may also be driven by hydrophobic regions between a groove in the FlaI three-helix bundle with an α helix of FlaJ, with protruding helices for this interaction existing in the cytoplasmic domains of membrane proteins in the Eps (Abendroth et al., 2005) and Bfp systems (Yamagata et al., 2012).

Combining FlaI biochemical data with our structural and genetic data suggests three key steps for FlaI functions in both archaellum assembly and movement, which supports and extends ideas proposed from sequence homologous T2S and T4P systems (Figure S5): (1) Monomeric FlaI protein in the cytoplasm self-assembles into a hexameric ring upon ATP binding. (2) Specific interactions between the FlaI NTD and both membrane lipids and its binding partner (most likely FlaJ) move the FlaI hexamer to the membrane. (3) Basic-clamp switching during successive rounds of ATP hydrolysis, product release, and ATP binding results in conformational translation with rotation of the stable FlaI intersubunit blocks, which drives both archaellum assembly and rotation. Interestingly, the top surface of FlaI moves up (and down) 8–10 Å when ATP is hydrolyzed (Movie S3). This shift range matches the spacing between pilin proteins in T4P filaments (Craig et al., 2003; Craig et al., 2006; Hartung et al., 2011), with reconstructed pili from various organisms all showing a distance of 5–10 Å between pilin neighbors. Thus, the conformational change in FlaI may be transferred through the probably tight interaction with FlaJ to push the archaellum filament out of the membrane, allowing a new archaellin subunit to bind and grow the filament.

Our results elucidate how FlaI structures and conformations play key roles in archaellum assembly and motility functions with implications for superfamily evolution and functions: crown geometry determines sites of interaction, and tight intersubunit moving blocks provide power for filament assembly and rotation, whereas loop connections provide interface elasticity by allowing functional movements without disrupting assembly. Overall, collective results reveal detailed and global FlaI activities in transducing nucleotide binding and hydrolysis into translational and rotatory motions suitable for assembly and motility without destabilizing the integrity of the hexameric crown assembly.

EXPERIMENTAL PROCEDURES

Strains and Growth Conditions

S. acidocaldarius strain DSM639 was used for all in vivo experiments, and deletion experiments were done in a Δ*aapF* background; see Supplemental Information.

ATPase Assay and Analytical Size-Exclusion Chromatography

Cloning, expression, and purification of recombinant FlaI and mutants are described in Supplemental Information. ATPase assays used the malachite green assay (Ghosh et al., 2011). For analytical size-exclusion chromatography, the protein was incubated at 50°C for 10 min and then 4°C for 6 hr with either 1 mM AMPPNP (adenylyl imidodiphosphate) or ADP plus 10 mM $MgCl_2$. The samples were applied to a Superdex 200 10/300 column, and fractions were analyzed with SDS-PAGE.

SAXS

SAXS data were collected and processed (see Supplemental Information) at the Advanced Light Source (ALS) beamline (BL) 12.3.1 (SIBYLS) Lawrence Berkeley National Laboratory (Classen et al., 2010; Hura et al., 2009). SAXS data have been deposited in BIOISIS.net under accession codes FLIADP and FLIATP for FlaI with ADP and ATP, respectively (ATP data published in Ghosh et al., 2011).

Crystallization and Structure Determination of FlaI

FlaI crystals were grown and data were collected, processed, and refined as described in the Supplemental Information. X-ray data were collected on beamlines 8.3.1 and 12.3.1 at the ALS. Refined models of $FlaI^{E336A}$-ADP at 2.0 Å and FlaI-Apo at 3.6 Å have good statistics (Table 1).

In-Frame Gene Deletion and Gene Complementation

Deletion mutant strains and recombinant plasmids were constructed as described in the Supplemental Information. All bacterial strains, primers, and plasmids are listed in Table S1.

Cell Fractionation and Localization Analysis

To test roles of highly conserved and specific residues in FlaI, we used cellular fractionation of the complemented deletion mutant (Δ*aapF*Δ*flaI*) expressing different FlaI variants (see Supplemental Information).

Electron Microscopy

For transmission EM (TEM), cells were fixed with 2.5% glutaraldehyde for 15 min at room temperature. Cell suspensions were placed directly on glow-discharged Formvar-coated 200-mesh copper grids (Plano, Wetzlar, Germany). Samples were negatively stained for 20 s with 2% uranyl acetate, washed twice with water, and air dried. TEM was performed using a JEOL JEM 3010 electron microscope (300 kV; LaB6 cathode).

Swimming Assays on Semisolid Plates and Thermomicroscopy

Analyses of motility and swimming behavior of *S. acidocaldarius* are described in Supplemental Information.

ACCESSION NUMBERS

Coordinates and structure factors have been deposited in the Protein Data Bank under accession numbers 4IHQ and 4II7. SAXS data for FlaI with ADP have been deposited in BIOISIS.net under accession code FLIADP.

SUPPLEMENTAL INFORMATION

Supplemental Information includes five figures, one table, six movies, and Supplemental Experimental Procedures and can be found with this article online at http://dx.doi.org/10.1016/j.molcel.2013.01.014.

ACKNOWLEDGMENTS

We thank Albers and Tainer laboratory members for helpful comments, especially Susan Tsutakawa, Michal Hammel, and ALS BL12.3.1 and BL8.3.1 staff. This study was supported by National Institute of Health (NIH) grant AI022160 to J.A.T. The SIBYLS beamline (BL12.3.1) is supported by the United States Department of Energy program IDAT and by NIH grant GM105404. A.G. received a Max Planck postdoctoral fellowship, and S.V.A. was supported by a VIDI grant from the Dutch Science Organization (NWO) and Max Planck Society intramural funds.

Received: May 8, 2012
Revised: October 23, 2012
Accepted: January 10, 2013
Published: February 14, 2013

REFERENCES

Abendroth, J., Murphy, P., Sandkvist, M., Bagdasarian, M., and Hol, W.G. (2005). The X-ray structure of the type II secretion system complex formed by the N-terminal domain of EpsE and the cytoplasmic domain of EpsL of Vibrio cholerae. J. Mol. Biol. *348*, 845–855.

Albers, S.V., Szabó, Z., and Driessen, A.J. (2003). Archaeal homolog of bacterial type IV prepilin signal peptidases with broad substrate specificity. J. Bacteriol. *185*, 3918–3925.

Albers, S.V., Jonuscheit, M., Dinkelaker, S., Urich, T., Kletzin, A., Tampé, R., Driessen, A.J., and Schleper, C. (2006). Production of recombinant and tagged proteins in the hyperthermophilic archaeon Sulfolobus solfataricus. Appl. Environ. Microbiol. *72*, 102–111.

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.

Bardy, S.L., Ng, S.Y., and Jarrell, K.F. (2004). Recent advances in the structure and assembly of the archaeal flagellum. J. Mol. Microbiol. Biotechnol. *7*, 41–51.

Burrows, L.L. (2005). Weapons of mass retraction. Mol. Microbiol. *57*, 878–888.

Chaban, B., Ng, S.Y., Kanbe, M., Saltzman, I., Nimmo, G., Aizawa, S., and Jarrell, K.F. (2007). Systematic deletion analyses of the fla genes in the flagella operon identify several genes essential for proper assembly and function of flagella in the archaeon, Methanococcus maripaludis. Mol. Microbiol. *66*, 596–609.

Chakraborty, S., Monfett, M., Maier, T.M., Benach, J.L., Frank, D.W., and Thanassi, D.G. (2008). Type IV pili in Francisella tularensis: roles of pilF and pilT in fiber assembly, host cell adherence, and virulence. Infect. Immun. *76*, 2852–2861.

Chen, Y., Shiue, S.J., Huang, C.W., Chang, J.L., Chien, Y.L., Hu, N.T., and Chan, N.L. (2005). Structure and function of the XpsE N-terminal domain, an essential component of the Xanthomonas campestris type II secretion system. J. Biol. Chem. *280*, 42356–42363.

Classen, S., Rodic, I., Holton, J., Hura, G.L., Hammel, M., and Tainer, J.A. (2010). Software for the high-throughput collection of SAXS data using an enhanced Blu-Ice/DCS control system. J. Synchrotron Radiat. *17*, 774–781.

Craig, L., and Li, J. (2008). Type IV pili: paradoxes in form and function. Curr. Opin. Struct. Biol. *18*, 267–277.

Craig, L., Taylor, R.K., Pique, M.E., Adair, B.D., Arvai, A.S., Singh, M., Lloyd, S.J., Shin, D.S., Getzoff, E.D., Yeager, M., et al. (2003). Type IV pilin structure and assembly: X-ray and EM analyses of Vibrio cholerae toxin-coregulated pilus and Pseudomonas aeruginosa PAK pilin. Mol. Cell *11*, 1139–1150.

Craig, L., Volkmann, N., Arvai, A.S., Pique, M.E., Yeager, M., Egelman, E.H., and Tainer, J.A. (2006). Type IV pilus structure by cryo-electron microscopy and crystallography: implications for pilus assembly and functions. Mol. Cell *23*, 651–662.

Crowther, L.J., Anantha, R.P., and Donnenberg, M.S. (2004). The inner membrane subassembly of the enteropathogenic Escherichia coli bundle-forming pilus machine. Mol. Microbiol. *52*, 67–79.

Filloux, A., Michel, G., and Bally, M. (1998). GSP-dependent protein secretion in gram-negative bacteria: the Xcp system of Pseudomonas aeruginosa. FEMS Microbiol. Rev. *22*, 177–198.

Ghosh, A., and Albers, S.V. (2011). Assembly and function of the archaeal flagellum. Biochem. Soc. Trans. *39*, 64–69.

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.

Gomez-Duarte, O.G., Chattopadhyay, S., Weissman, S.J., Giron, J.A., Kaper, J.B., and Sokurenko, E.V. (2007). Genetic diversity of the gene cluster encoding longus, a type IV pilus of enterotoxigenic Escherichia coli. J. Bacteriol. *189*, 9145–9149.

Hansen, J.K., and Forest, K.T. (2006). Type IV pilin structures: insights on shared architecture, fiber assembly, receptor binding and type II secretion. J. Mol. Microbiol. Biotechnol. *11*, 192–207.

Hartung, S., Arvai, A.S., Wood, T., Kolappan, S., Shin, D.S., Craig, L., and Tainer, J.A. (2011). Ultrahigh resolution and full-length pilin structures with insights for filament assembly, pathogenic functions, and vaccine potential. J. Biol. Chem. *286*, 44254–44265.

Henche, A.L., Koerdt, A., Ghosh, A., and Albers, S.V. (2012). Influence of cell surface structures on crenarchaeal biofilm formation using a thermostable green fluorescent protein. Environ. Microbiol. *14*, 779–793.

Holm, L., and Rosenström, P. (2010). Dali server: conservation mapping in 3D. Nucleic Acids Res. *38*(Web Server issue), W545–W549.

Hura, G.L., Menon, A.L., Hammel, M., Rambo, R.P., Poole, F.L., 2nd, Tsutakawa, S.E., Jenney, F.E., Jr., Classen, S., Frankel, K.A., Hopkins, R.C., et al. (2009). Robust, high-throughput solution structural analyses by small angle X-ray scattering (SAXS). Nat. Methods *6*, 606–612.

Iyer, L.M., Leipe, D.D., Koonin, E.V., and Aravind, L. (2004). Evolutionary history and higher order classification of AAA+ ATPases. J. Struct. Biol. *146*, 11–31.

Jarrell, K.F., and Albers, S.V. (2012). The archaellum: an old motility structure with a new name. Trends Microbiol. *20*, 307–312.

Johnson, T.L., Abendroth, J., Hol, W.G., and Sandkvist, M. (2006). Type II secretion: from structure to function. FEMS Microbiol. Lett. *255*, 175–186.

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.

McCoy, A.J. (2007). Solving structures of protein complexes by molecular replacement with Phaser. Acta Crystallogr. D Biol. Crystallogr. *63*, 32–41.

Misic, A.M., Satyshur, K.A., and Forest, K.T. (2010). P. aeruginosa PilT structures with and without nucleotide reveal a dynamic type IV pilus retraction motor. J. Mol. Biol. *400*, 1011–1021.

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.

Planet, P.J., Kachlany, S.C., DeSalle, R., and Figurski, D.H. (2001). Phylogeny of genes for secretion NTPases: identification of the widespread tadA subfamily and development of a diagnostic key for gene classification. Proc. Natl. Acad. Sci. USA *98*, 2503–2508.

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.

Putnam, C.D., Hammel, M., Hura, G.L., and Tainer, J.A. (2007). X-ray solution scattering (SAXS) combined with crystallography and computation: defining accurate macromolecular structures, conformations and assemblies in solution. Q. Rev. Biophys. *40*, 191–285.

Robien, M.A., Krumm, B.E., Sandkvist, M., and Hol, W.G. (2003). Crystal structure of the extracellular protein secretion NTPase EpsE of Vibrio cholerae. J. Mol. Biol. *333*, 657–674.

Russel, M. (1998). Macromolecular assembly and secretion across the bacterial cell envelope: type II protein secretion systems. J. Mol. Biol. *279*, 485–499.

Satyshur, K.A., Worzalla, G.A., Meyer, L.S., Heiniger, E.K., Aukema, K.G., Misic, A.M., and Forest, K.T. (2007). Crystal structures of the pilus retraction motor PilT suggest large domain movements and subunit cooperation drive motility. Structure *15*, 363–376.

Savvides, S.N., Yeo, H.J., Beck, M.R., Blaesing, F., Lurz, R., Lanka, E., Buhrdorf, R., Fischer, W., Haas, R., and Waksman, G. (2003). VirB11 ATPases are dynamic hexameric assemblies: new insights into bacterial type IV secretion. EMBO J. *22*, 1969–1980.

Schneidman-Duhovny, D., Inbar, Y., Nussinov, R., and Wolfson, H.J. (2005). PatchDock and SymmDock: servers for rigid and symmetric docking. Nucleic Acids Res. *33*(Web Server issue), W363–W367.

Shin, D.S., Pellegrini, L., Daniels, D.S., Yelent, B., Craig, L., Bates, D., Yu, D.S., Shivji, M.K., Hitomi, C., Arvai, A.S., et al. (2003). Full-length archaeal Rad51 structure and mutants: mechanisms for RAD51 assembly and control by BRCA2. EMBO J. *22*, 4566–4576.

Shiue, S.J., Kao, K.M., Leu, W.M., Chen, L.Y., Chan, N.L., and Hu, N.T. (2006). XpsE oligomerization triggered by ATP binding, not hydrolysis, leads to its association with XpsL. EMBO J. *25*, 1426–1435.

Molecular Cell
Archaellum Assembly & Rotation by the ATPase FlaI

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.

Tainer, J.A., Getzoff, E.D., Alexander, H., Houghten, R.A., Olson, A.J., Lerner, R.A., and Hendrickson, W.A. (1984). The reactivity of anti-peptide antibodies is a function of the atomic mobility of sites in a protein. Nature *312*, 127–134.

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.

Thomas, N.A., Mueller, S., Klein, A., and Jarrell, K.F. (2002). Mutants in flaI and flaJ of the archaeon Methanococcus voltae are deficient in flagellum assembly. Mol. Microbiol. *46*, 879–887.

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.

Tripathi, S.A., and Taylor, R.K. (2007). Membrane association and multimerization of TcpT, the cognate ATPase ortholog of the Vibrio cholerae toxin-coregulated-pilus biogenesis apparatus. J. Bacteriol. *189*, 4401–4409.

Wagner, M., Berkner, S., Ajon, M., Driessen, A.J., Lipps, G., and Albers, S.V. (2009). Expanding and understanding the genetic toolbox of the hyperthermophilic genus Sulfolobus. Biochem. Soc. Trans. *37*, 97–101.

Yamagata, A., and Tainer, J.A. (2007). Hexameric structures of the archaeal secretion ATPase GspE and implications for a universal secretion mechanism. EMBO J. *26*, 878–890.

Yamagata, A., Milgotina, E., Scanlon, K., Craig, L., Tainer, J.A., and Donnenberg, M.S. (2012). Structure of an essential type IV pilus biogenesis protein provides insights into pilus and type II secretion systems. J. Mol. Biol. *419*, 110–124.

Molecular Cell, Volume *49*

Supplemental Information

Insights into FlaI Functions in Archaeal Motor Assembly and Motility from Structures, Conformations, and Genetics

Sophia Reindl, Abhrajyoti Ghosh, Gareth J. Williams, Kerstin Lassak, Tomasz Neiner, Anna-Lena Henche, Sonja-Verena Albers, and John A. Tainer

Supplemental Information Inventory

Figure S1A

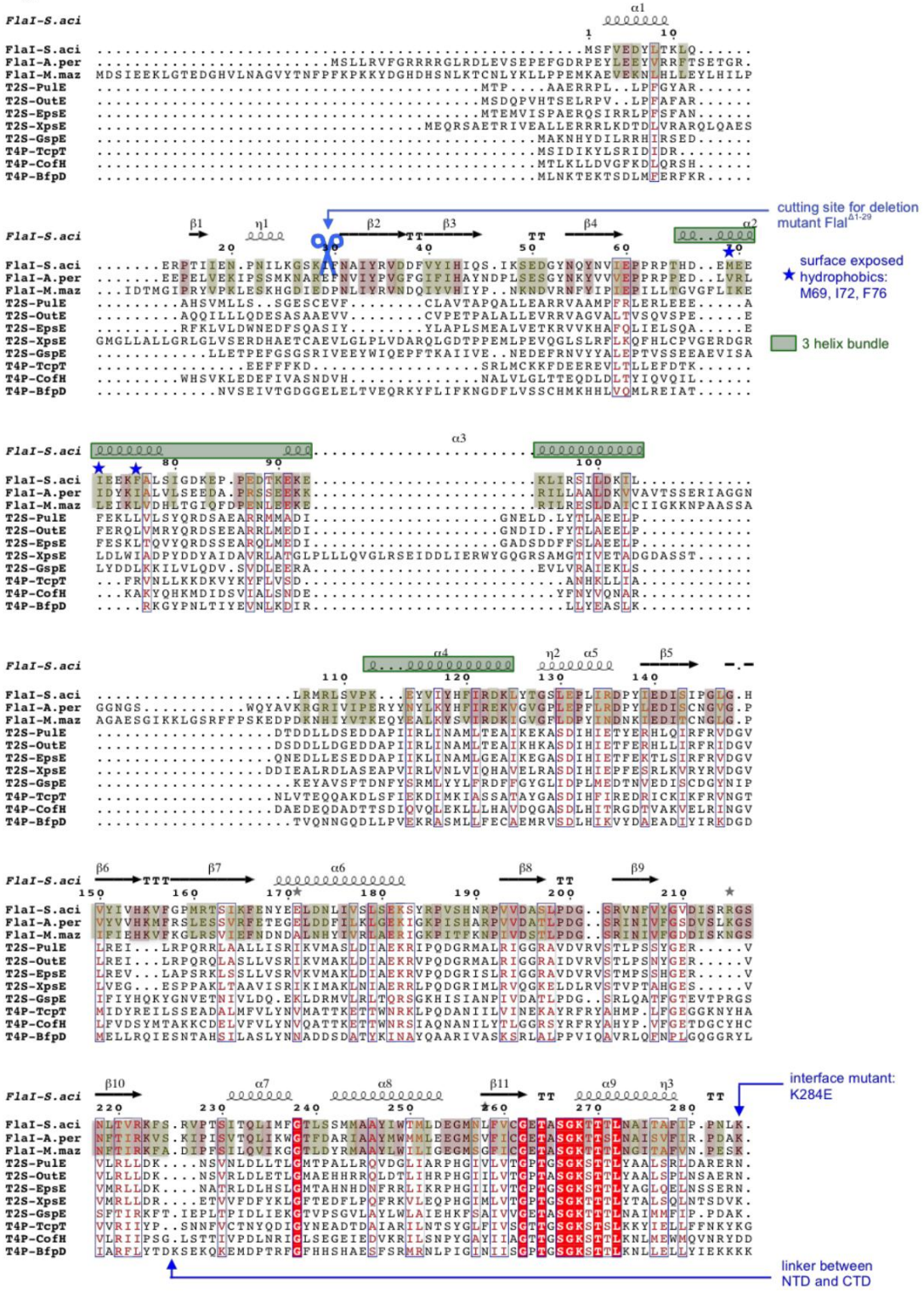

cutting site for deletion mutant FlaI Δ1-29
surface exposed hydrophobics: M69, I72, F76
3 helix bundle
interface mutant: K284E
linker between NTD and CTD

Figure S1A - continued

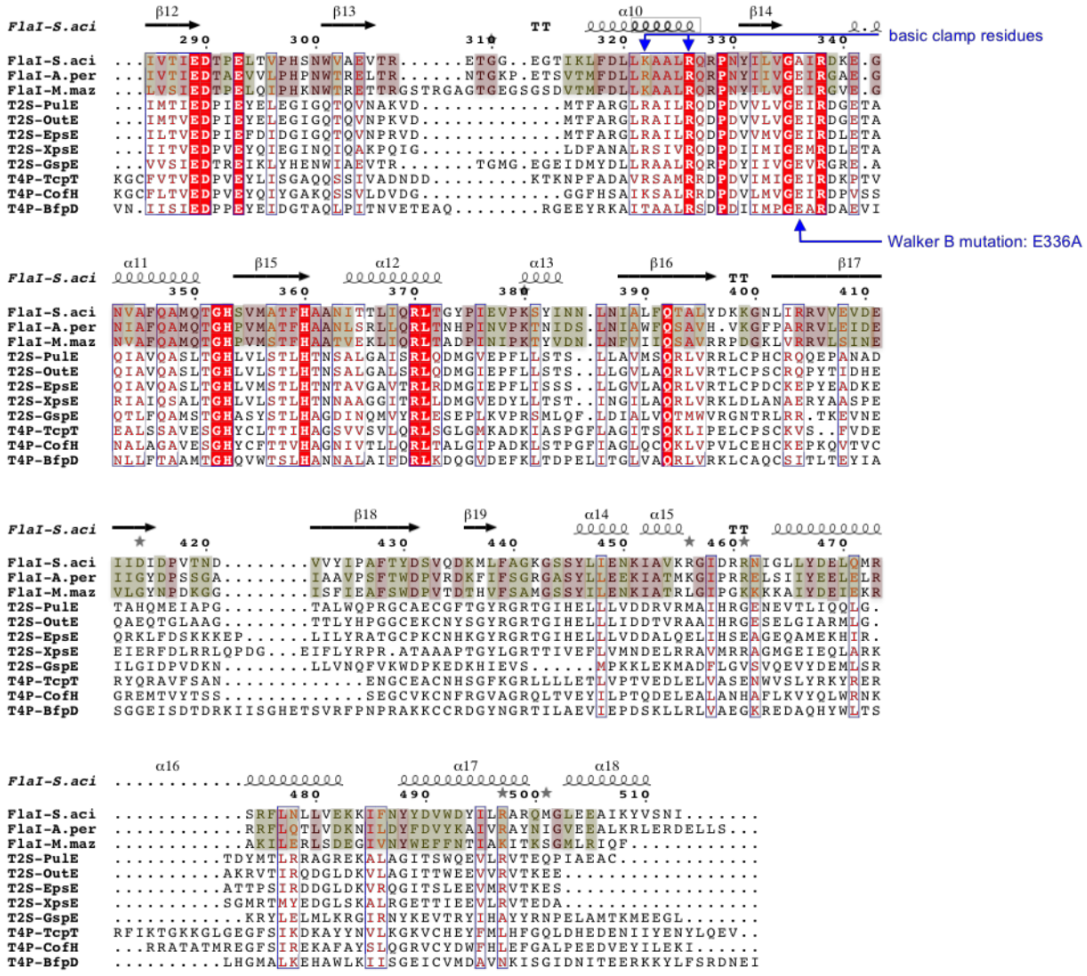


Figure S1B

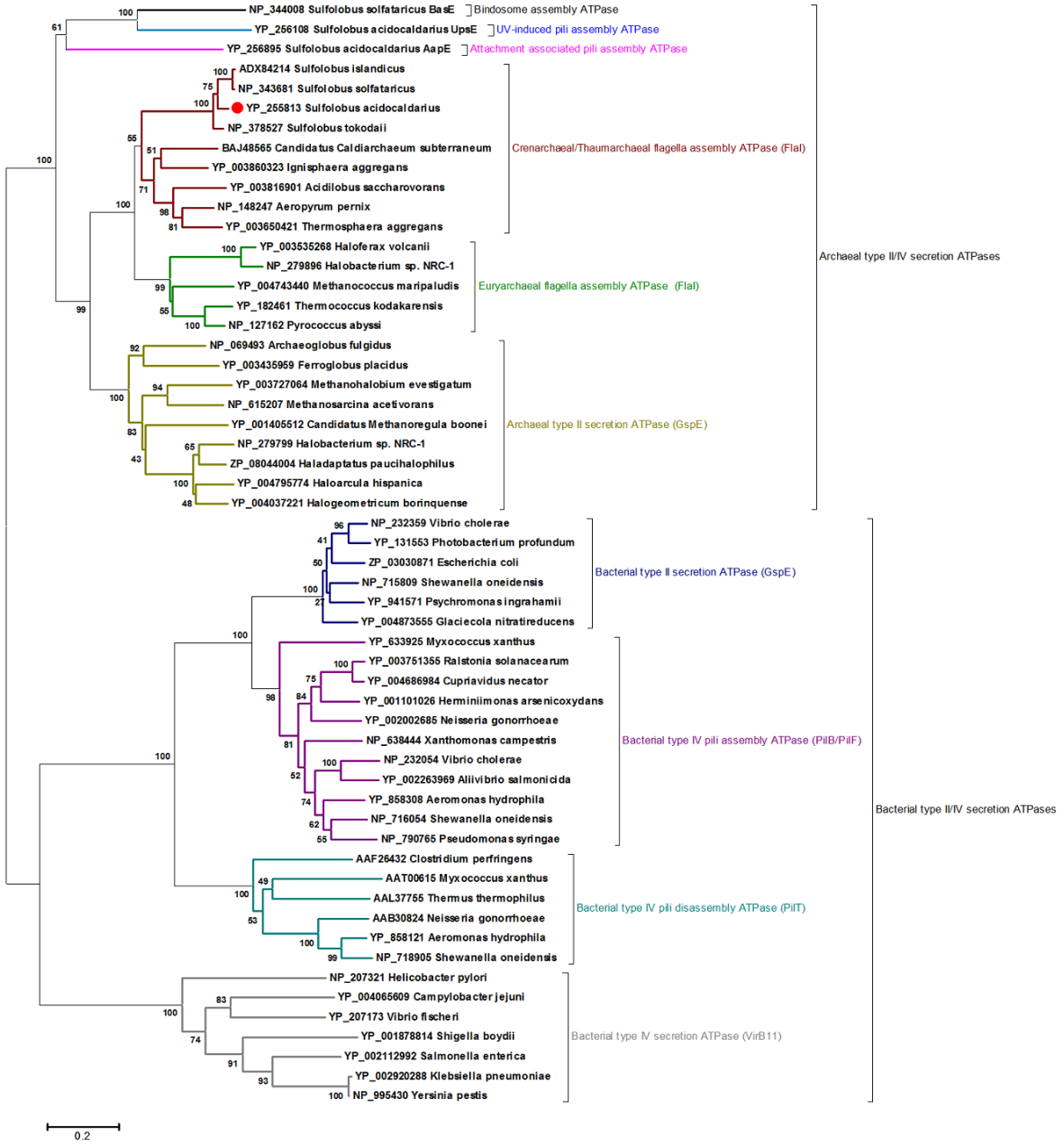


Figure S1. Similarity between Assembly ATPases for T4P and T2S/T4S, Related to Figure 1
(A) Sequence alignment between *S. acidocaldarius* FlaI, other FlaI proteins, and homologs from T2S and T4P systems**.** Identical residues are highlighted in red, similar residues in red letters. Homology between FlaI proteins is shown in light red (identical residues) and yellow (similar residues). FlaI secondary structure elements are shown above the alignment for β-sheets (arrows), α-helices (coils) and β-turns (T). Regions and residues relevant for this study are highlighted and marked. The sequence alignment was prepared with ClustalW (Larkin et al, 2007) and ESPript (Gouet et al, 1999). Sequence identity between FlaI and other crystallized homologs are as follows: GspE 34%; VirB11 16%; PilT 12% and EpsE: 7%. **(B)** Phylogenetic tree showing the relationship between the different assembly ATPases for T4P and T2S/T4S in bacteria and archaea and the archaeal flagellum.

Movie S1. Swimming Motility of Hyperflagellated Δ*aapF* Strain of *S. acidocaldarius* Analyzed by Thermomicroscopy, Related to Figure 1

Movie S2. Nonmotile Δ*aapF*Δ*flaI* Strain of *S. acidocaldarius* Analyzed by Thermomicroscopy, Related to Figure 1

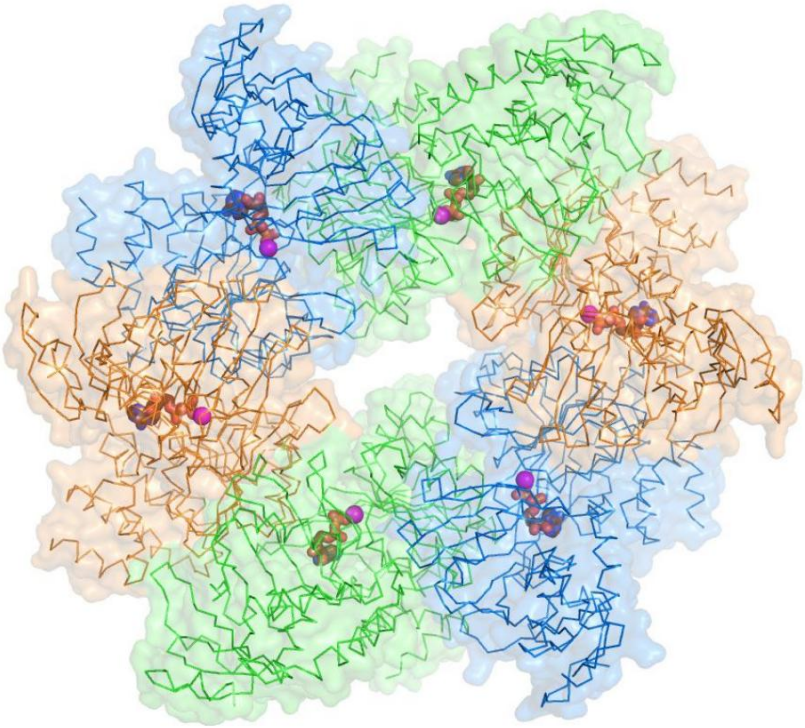


Figure S2. Hexameric Structure of FlaI with Bound Nucleotide, Related to Figure 2

FlaI is shown as ribbons and ADP in spheres with Mg^{2+} in magenta. The nucleotide is located between two CTDs and one NTD. Chain A is shown in green, chain B in orange and chain C in blue.

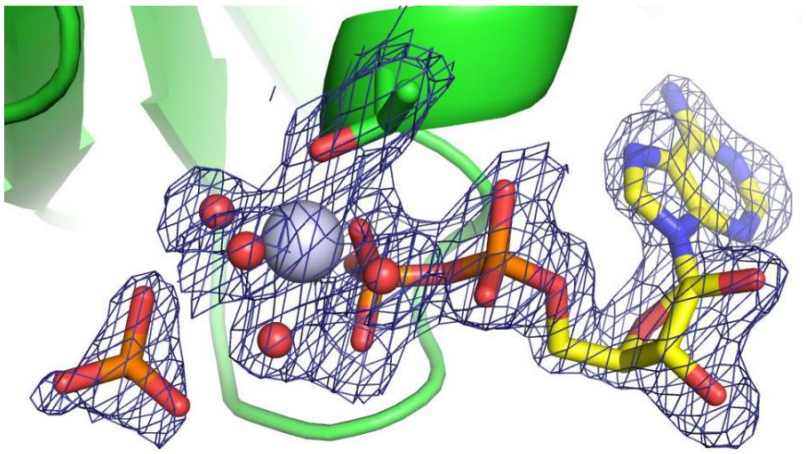


Figure S3. Simulated Annealing Composite Omit Map of the Phosphate-Release State in Chain A of the FlaI^{E336A}-ADP Crystals Is Shown at 1 σ for Verification of the Existence of a Phosphate, Related to Figure 3

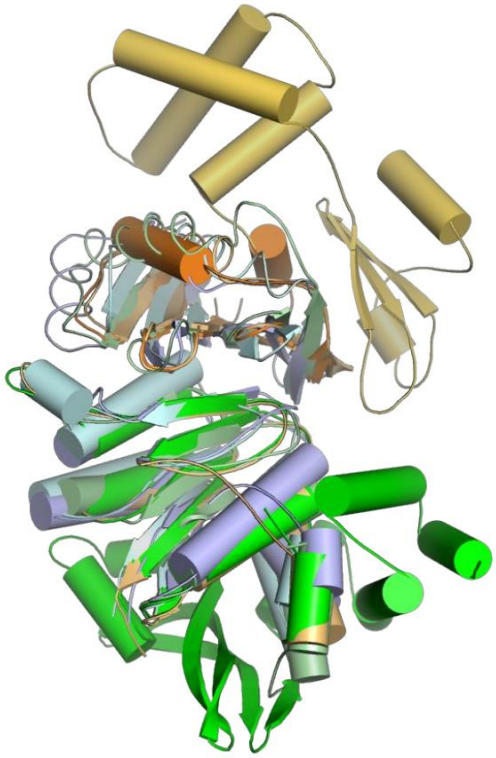


Figure S4. Superposition of Building Blocks of Different T2S/T4S Secretion ATPases, Related to Figure 5

FlaI is shown with secondary structure elements as cartoon and colored as in Figure 5 (NTD-1 in yellow, NTD-2 in orange and CTD in green). For GspE (dark orange), EpsE (cyan), VirB11 (dark green) and PilT

(blue) only the regions that overlap with the FlaI building block are shown, as with the secondary structure elements as cartoons.

Movie S3. Domain Movement within the FlaI Hexameric Ring, Related to Figure 5
From the different conformations present in the crystal structures we observe the following movement induced by ATP hydrolysis. The hexameric ring is shown in side view with two building blocks facing the reader. A second copy of the hexamer is rotated by 60°, so the right building blocks superimpose. The movie shows a morph between the left building block of one hexamer (NTD^{chainA} + CTD^{chainB}) and the left building block of the second hexamer (NTD^{chainB} + CTD^{chainC}). One FlaI subunit is shown in green, the CTD of the subunit to the left in blue and the NTD of the subunit to the right in orange. The base of the hexameric ring formed by the CTDs is shown in gray. The observed movement clearly demonstrates that a rigid building block on one CTD (blue) and the NTD of the neighboring molecule (green) is moving, not the NTD is moving with respect to the ring of CTDs.

Movie S4. Domain Movement within the GspE Hexameric Ring, Related to Figure 5
Prepared in analogy to Movie S3. As seen for FlaI, building blocks of CTD^{n}:NTD^{n+1} move with respect to each other. The observed movement for the secretion ATPase GspE differs in its directionality from the one seen for the archaellum ATPase FlaI. For easier comparison the subunits are colored as in Movie S1.

Movie S5. Nonmotile *ΔaapFΔNflaI* Strain of *S. acidocaldarius* Analyzed by Thermomicroscopy, Related to Figure 6

Movie S6. Nonmotile *ΔaapFΔflaI* Strain of *S. acidocaldarius* Expressing $FlaI^{M69E,I72E,F76E}$ In Trans from a Multicopy Plasmid pSVA1450 Analyzed by Thermomicroscopy, Related to Figure 7

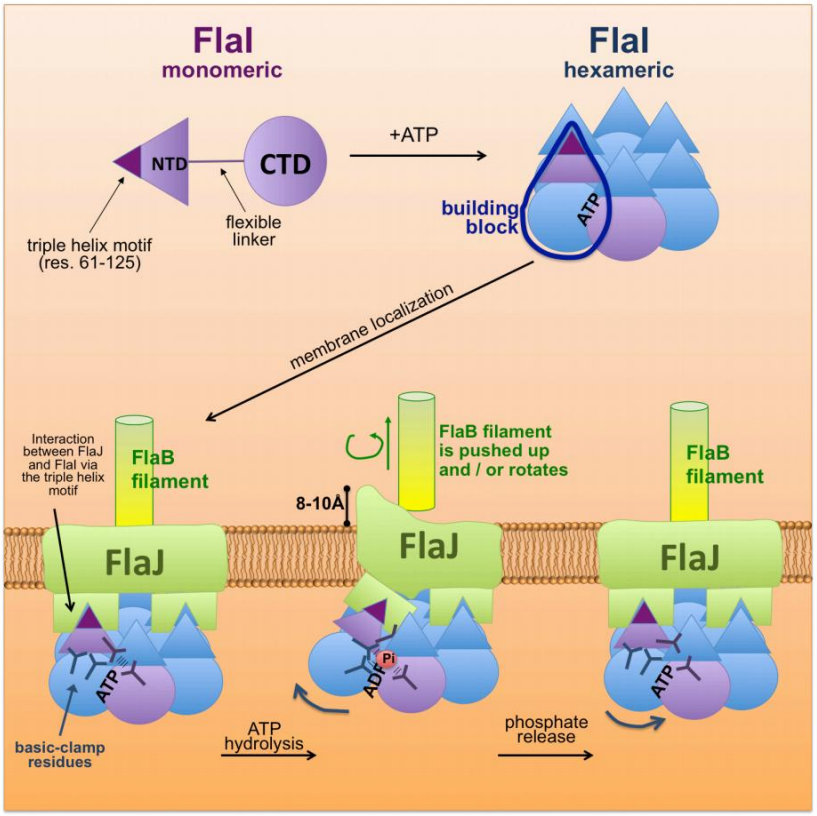


Figure S5. Cartoon of FlaI Functions in Archaeal Motor Assembly, related to Figure 1-7

Table S1. Bacterial Strains, Primers, and Plasmids

Bacterial strains, primers and plasmids	Relevant characteristics	Source or reference
Strains		
E.coli		
DH5α	l2 f80d/*lacZDM15* D(*lacZYA-argF*)*U169 recA1 endA1 hsdR17* (rK2 mK1) *supE44 thi-1 gyrA relA1*	Invitrogen
Walker cells- RIL (C43)	F-*ompT gal hsdSB*(rB-mB-) *dcmlon*_DE3 *endA*Hte [*argUileYleuW*Camr]	Miroux& Walker
K12 (ER1821)	F$^-$ *glnV44 e14*$^-$*(McrA*$^-$*) rfbD1? relA1? endA1 spoT1? thi-1 Δ(mcrC-mrr)114::IS10* (containing pM.EsaBC4I plasmid)	NEB
Sulfolobus acidocaldarius		
DSM639	Hyperthermo-acidophilic crenarchaeon which grows optimally at 80˚C and pH 2-3	DSMZ
MW001	Wild-type, *pyrE*$^-$ (a 322 bp deletion in the *pyrE* gene)	Wagner *et.al.*
Δ*flaI*Δ*aapF*(MW456)	MW001, ΔSaci_1173/ΔSaci_2318, *pyrE*$^-$	Lassak *et.al.*
Δ*flaI*Δ*aapF* (MW501)	MW001, ΔNSaci_1173 (deletion of 672 nucleotides at the 5' end of saci_1173), *pyrE*$^-$	This study
Primers		
FlaI713	5'-GGGGGCATATGAGCTTCGTGGAAGAC-3'; Forward primer for FlaI containing NdeI restriction site	Ghosh *et.al.*
FlaI714	5'-GGGGGCTCGAGTCAAATATTACTGACATATTTTATC-3'; Reverse primer for FlaI containing XhoI restriction site	Ghosh *et.al.*
FlaI728	5'- CCCCCGAATTCGATGAGCTTCGTGGAAGAC-3'; Forward primer for FlaI containing EcoRI restriction site	Ghosh *et.al.*
FlaI729	5'- CCCCCAAGCTTTCAAATATTACTGACATATTTTATC-3'; Reverse primer for FlaI containing HindIII restriction site	Ghosh *et.al.*
FlaI742	5'- ATCTGGAGCGACAACTAC-3'; Overlapping internal forward primer for K_{268}A mutation in FlaI	Ghosh *et.al.*
FlaI743	5'- AATGTAGTTGTCGCTCCAG-3'; Overlapping internal reverse primer for K_{268}A mutation in FlaI	Ghosh *et.al.*
FlaI763	5'-CCCCCCCATGGTAAGCTTCGTGGAAGAC-3'; Forward primer for FlaI containing NcoI restriction site	This study
FlaI764	5'-CCCCCGGATCCCCAAATATTACTGACATATTTTATCG-3'; Reverse primer for FlaI containing BamHI restriction site	This study
FlaI784	5'-ACTTGTGGGAGCAATTAGAG-3'; Overlapping internal forward primer for E_{336}A mutation in FlaI	Ghosh *et.al.*
FlaI785	5'- CTTTGTCTCTAATTGCTCCCACAAG-3'; Overlapping internal reverse primer for E_{336}A mutation in FlaI	Ghosh *et.al.*
ΔNFlaI$_{ko}$799	5'-CCCCCGGGCCCAGGGATATGTGATTAC-3'; ΔN*fla*-us-fw: Upstream forward primer with ApaI restriction site	This study
ΔNFlaI$_{ko}$800	5'-ACTAAACATGACTTATGCCCTCGATAAG-3'; ΔN*fla*-us-rev: Upstream reverse primer	This study
ΔNFlaI$_{ko}$801	5'-GTCATGTTTAGTAGAGTTCCAACCAGTATAAC-3'; ΔN*fla*-ds-fw: Downstream forward primer	This study
ΔNFlaI$_{ko}$802	5'-CCCCCCTGCAGAACCCTCCTGATTAGATTCCC-3'; ΔN*fla*-ds-rev: Downstream reverse primer with PstI restriction site	This study
ΔNFlaI$_{Exp}$815	5'-CCCCCGAATTCGTTTAGTAGAGTTCCAACC-3'; Forward primer for FlaI containing EcoRI restriction site	This study
FlaI816	5'-TTTCTCATGAAAGACCAGTAG-3'; Overlapping internal forward primer for N_{268}E mutation in FlaI	This study
FlaI817	5'-TACTGGTCTTTCATGAGAAAC-3'; Overlapping internal reverse primer for N_{268}E mutation in FlaI	This study
FlaI818	5'-CAAATCTAGAGATAGTTACTATTGAG-3'; Overlapping internal forward primer for K_{284}E mutation in FlaI	This study
FlaI819	5'-CCTCAATAGTAACTATCTCTAGATTTGG-3'; Overlapping internal reverse primer for K_{284}E mutation in FlaI	This study
FlaI820	5'-TCACTAAAGAAACTGGTGGTGAAG-3'; Overlapping internal forward primer for R_{307}K mutation in FlaI	This study
FlaI821	5'-ACCACCAGTTTCTTTAGTGACCTC-3'; Overlapping internal reverse primer for R_{307}K mutation in FlaI	This study
Δ1-29FlaI830	5'-CCCCCCCATGGTATTTAATGCAATCTATAGAGTAGACG-3'; Forward primer for Δ1-29FlaI containing NcoI restriction site	This study
Δ61-126FlaI834	5'-CAATAATAGAAAACCATATTCAGAGTATAAAAAGTGAGG-3'; Overlapping internal forward primer for Δ61-126FlaI	This study
Δ61-126FlaI835	5'-CACTGAATATGGTTTTCTATTATTGTAGGTCTTTCCTG-3'; Overlapping internal reverse primer for Δ61-126FlaI	This study

Table S1. Bacterial Strains, Primers, and Plasmids - continued

Bacterial strains, primers and plasmids	Relevant characteristics	Source or reference
Primers		
FlaI838	5'-CATGATGAAGAGGAGGAAGAAGAGGAGAAGGAAGCCTTATCTATCGG-3'; Overlapping internal forward primer for $M_{69}E$, $I_{72}E$, $F_{76}E$ mutation in FlaI	This study
FlaI839	5'-GATAAGGCTTCCTTCTCCTCTTCTTCCTCCTCTTCATCATGAGTAGGTC-3'; Overlapping internal reverse primer for $M_{69}E$, $I_{72}E$, $F_{76}E$ mutation in FlaI	This study
FlaI849	5'-CCCCCCGGCCGTCAAATATTACTGACATATTTTATC-3'; Reverse primer for FlaI containing EagI restriction site	This study
Plasmids		
pETDuet-1	Amp^r, Cam^r, expression plasmid containing replicon ColE1 (pBR322) and two MCS (MCS1 and MCS2)	
pSVA251	Amp^r, Cam^r, pETDuet-1 carrying untagged FlaI in MCS2 using restriction sites NdeI-XhoI	Ghosh *et.al.*
pSVA263	Amp^r, Cam^r, pETDuet-1 carrying N-terminal His-tagged FlaI in MCS1 using restriction sites EcoRI-HindIII	Ghosh *et.al.*
pSVA265	Amp^r, Cam^r, pETDuet-1 carrying N-terminal His-tagged FlaI($K_{268}A$) in MCS1 using restriction sites EcoRI-HindIII	Ghosh *et.al.*
pSVA307	Amp^r, Cam^r, pETDuet-1 carrying N-terminal His-tagged FlaI($D_{290}A$) in MCS1 using restriction sites EcoRI-HindIII	Ghosh *et.al.*
pSVA308	Amp^r, Cam^r, pETDuet-1 carrying N-terminal His-tagged FlaI($E_{336}A$) in MCS1 using restriction sites EcoRI-HindIII	Ghosh *et.al.*
pCMalLacS	pRN1-based shuttle vector with *S.solfataricus pyrEF* gene and *lacS* reporter gene β-glycosidase from *Sulfolobus solfataricus* containing Mal-promoter for inducible expression in *S.acidocaldarius*	Berkner *et.al.*
pSVA406	pGEMT-easy based vector containing the *S. solfataricuspyrEF*gene	Wagner *et.al.*
pSVA1423	pCMalLacS vector containing *saci_1161* (MRP) for higher inducible expr. of target gene	Wagner *et.al.*
pSVA1450	pSVA1423 vector containing deletion of 2 bp in the *mal*-promoter for higher inducible expression of the target gene	Wagner & Albers pers. Comm.
pMZ1	Amp^r, cloning vector containing replicon ColE1 (pBR322) and a 10X His-Strep tag sequence at the 3' end of the MCS	Zolghardt *et.al.*
pSVA1481	pGEMT-easy based vector containing the cloning cassette from pMZ1	Wagner & Albers pers. Comm.
pSVA296	Amp^r, pMZ1 carrying FlaI in MCS using restriction sites NcoI-BamHI	This study
pSVA298	Amp^r, pMZ1 carrying FlaI($K_{268}A$) in MCS using restriction sites NcoI-BamHI	This study
pSVA326	Amp^r, Cam^r, pETDuet-1 carrying untagged FlaI($E_{336}A$) in MCS2 using restriction sites NdeI-XhoI	This study
pSVA329	$\Delta NFlaI_{ko}$ in pSVA406	This study
pSVA334	Amp^r, pMZ1 carrying FlaI($E_{336}A$) in MCS using restriction sites NcoI-BamHI	This study
pSVA335	Amp^r, pSVA1423 carrying FlaI($E_{336}A$) in MCS using restriction sites NcoI-EagI with C-terminal 10X His-Strep tag sequence	This study
pSVA336	Amp^r, pSVA1423 carrying FlaI in MCS using restriction sites NcoI-EagI with C-terminal 10X His-Strep tag sequence	This study
pSVA337	Amp^r, pSVA1423 carrying FlaI($K_{268}A$) in MCS using restriction sites NcoI-EagI with C-terminal 10X His-Strep tag sequence	This study
pSVA338	Amp^r, Cam^r, pETDuet-1 carrying N-terminal His-tagged ΔNFlaI(Δ1-224) in MCS1 using restriction sites EcoRI-HindIII	This study
pSVA343	Amp^r, Cam^r, pETDuet-1 carrying N-terminal His-tagged FlaI($N_{190}E$) in MCS1 using restriction sites EcoRI-HindIII	This study
pSVA344	Amp^r, Cam^r, pETDuet-1 carrying N-terminal His-tagged FlaI($K_{284}E$) in MCS1 using restriction sites EcoRI-HindIII	This study
pSVA345	Amp^r, Cam^r, pETDuet-1 carrying N-terminal His-tagged FlaI($R_{307}K$) in MCS1 using restriction sites EcoRI-HindIII	This study
pSVA348	Amp^r, Cam^r, pETDuet-1 carrying N-terminal His-tagged FlaI($K_{284}E$, $E_{336}A$) in MCS1 using restriction sites EcoRI-HindIII	This study
pSVA349	Amp^r, Cam^r, pETDuet-1 carrying N-terminal His-tagged FlaI($N_{190}E$, $E_{336}A$) in MCS1 using restriction sites EcoRI-HindIII	This study
pSVA350	Amp^r, Cam^r, pETDuet-1 carrying N-terminal His-tagged FlaI($R_{307}K$, $E_{336}A$) in MCS1 using restriction sites EcoRI-HindIII	This study
pSVA351	Amp^r, pSVA1450 carrying FlaI in MCS using restriction sites NcoI-EagI	This study
pSVA352	Amp^r, pSVA1450 carrying FlaI($K_{268}A$) in MCS using restriction sites NcoI-EagI	This study
pSVA353	Amp^r, pSVA1450 carrying FlaI($E_{336}A$) in MCS using restriction sites NcoI-EagI	This study
pSVA354	Amp^r, pSVA1481 carrying Δ1-29FlaI in MCS using restriction sites NcoI-BamHI	This study
pSVA355	Amp^r, pSVA1450 carrying Δ1-29FlaI in MCS using restriction sites NcoI-EagI with C-terminal 10X His-Strep tag sequence	This study

Table S1. Bacterial Strains, Primers, and Plasmids - continued

Bacterial strains, primers and plasmids	Relevant characteristics	Source or reference
Plasmids		
pSVA359	Amp^r, pSVA1481 carrying Δ61-126FlaI in MCS using restriction sites NcoI-BamHI	This study
pSVA366	Amp^r, pSVA1481 carrying FlaI(M_{69}E, I_{72}E, F_{76}E) in MCS using restriction sites NcoI-BamHI	This study
pSVA368	Amp^r, pSVA1450 carrying FlaI in MCS using restriction sites NcoI-EagI with C-terminal 10X His-Strep tag sequence	This study
pSVA370	Amp^r, pSVA1450 carrying Δ61-126FlaI in MCS using restriction sites NcoI-EagI with C-terminal 10X His-Strep tag sequence	This study
pSVA371	Amp^r, pSVA1481 carrying FlaI(N_{190}E) in MCS using restriction sites NcoI-BamHI	This study
pSVA372	Amp^r, pSVA1481 carrying FlaI(K_{284}E) in MCS using restriction sites NcoI-BamHI	This study
pSVA373	Amp^r, pSVA1481 carrying FlaI(R_{307}K) in MCS using restriction sites NcoI-BamHI	This study
pSVA374	Amp^r, pSVA1450 carrying FlaI(N_{190}E) in MCS using restriction sites NcoI-EagI with C-terminal 10X His-Strep tag sequence	This study
pSVA375	Amp^r, pSVA1450 carrying FlaI(K_{284}E) in MCS using restriction sites NcoI-EagI with C-terminal 10X His-Strep tag sequence	This study
pSVA376	Amp^r, pSVA1450 carrying FlaI(R_{307}K) in MCS using restriction sites NcoI-EagI with C-terminal 10X His-Strep tag sequence	This study
pSVA380	Amp^r, pSVA1450 carrying FlaI(M_{69}E, I_{72}E, F_{76}E) in MCS using restriction sites NcoI-EagI with C-terminal 10X His-Strep tag sequence	This study

Supplemental Experimental Procedures

Figure Preparation

Figures were prepared using PyMOL (Schrodinger, 2010) and Chimera (Pettersen et al., 2004).

Strains and Growth Conditions

S. acidocaldarius DSM639 was grown aerobically at 75°C in Brock's basal salts medium (Brock et al, 1972) adjusted to pH 3.5 with sulphuric acid and supplemented with 0.1% (w/v) tryptone (Roth) or NZAmine AS (Sigma) and 0.2% (w/v) dextrin. For the uracil auxotrophic strains; *S. acidocaldarius* MW001 (containing a 322 bp deletion in the *pyrE* gene) (Wagner and Albers, unpublished), *S. acidocaldarius flaI* mutant strain (MW456; *ΔflaI* in *S. acidocaldarius* Δ*aapF* background) (Lassak et al., 2012) and *S. acidocaldarius* Δ*flaI* mutant strain (MW501; Δ*flaI* in *S .acidocaldarius* Δ*aapF* background); the growth medium was supplemented with 10 µg ml^{-1} uracil. To prepare plates, Brock medium was solidified by adding a final concentration of 0.6% (w/v) Gelrite (Sigma) and $MgCl_2$ and $CaCl_2$ to 10 mM and 3 mM, respectively. Plates were then incubated for five days at 75°C. For the propagation of plasmids *E. coli* strain DH5α was used. For the methylation of the plasmids *E.coli* ER1821^{kan} was used. The *E.coli* BL21(DE3)-C43 containing RIL plasmid was used for heterologous expression analysis.

Cloning and Expression

Plasmids pSVA251, pSVA263, pSVA265, pSVA308, pSVA326 containing the *flaI* gene (*saci_1173*) either without tag or N-terminal His tag, K268A (walker A mutant) *flaI* mutant with N-terminal His tag, and E336A (walker B mutant) *flaI* mutant either without tag or N-terminal His tag, respectively were constructed using expression vector pETDuet-1 (Novagen). Additionally, the fragments of *flaI* gene containing either Asn190Glu, Lys284Glu, Ard307Lys mutations in the interface region of the hexameric FlaI were constructed by site-directed mutagenesis and were cloned as EcoRI/HindIII fragments into pETDuet-1 expression vector yielding recombinant expression vectors pSVA343 (FlaI^{N190E}), pSVA344 (FlaI^{K284E}), pSVA345 (FlaI^{R307K}), respectively. The same interface point mutations were also introduced into pSVA308 (FlaI^{E336A}, *flaI* walker B mutant) using similar techniques and generated expression vectors pSVA348 (FlaI^{K284E,E336A}), pSVA349 (FlaI^{N190E,E336A}), pSVA350 (FlaI^{R307K,E336A}), respectively. An N-terminally truncated *flaI* (Δ*NFlaI*, N-terminal 224 amino acids were removed from full length FlaI) was cloned as EcoRI/HindII fragment into pETDuet-1 resulting in the pSVA338 (Δ*flaI*) expression vector. All the *flaI* fragments with or without mutations, as well as the Δ*flaI* construct were

overexpressed from the respective constructs in *E.coli* Bl21 (DE3)-C43 cells contacting RIL^{Cm} plasmid and induced by the addition of IPTG as described previously. After expression the cells were collected by centrifugation, resuspended in lysis buffer [50 mM Hepes-NaOH (pH 7.5), 150 mM NaCl and 10% glycerol] containing the complete EDTA-free protease inhibitor cocktail (1 tablet/50 ml of lysate; Roche), frozen in liquid nitrogen and stored at −80°C.

For *in vivo* studies, substitution mutations in the full-length *flaI*; $flaI^{K268A}$, $flaI^{E336A}$, $flaI^{M69E,I72E,F76E}$ were created by site-directed mutagenesis and cloned as NcoI/BamHI fragments into the entry vector pMZ1/pSVA1481 to introduce C-terminal tandem tag (Strep-10xHis-tag) and then sub-cloned as NcoI/EagI fragments into the expression vectors pSVA1423/pSVA1450 (derived from pCmalLacS vector). The resultant expression constructs pSVA337 ($FlaI^{K268A}$), pSVA335 ($FlaI^{E336A}$), pSVA370 ($FlaI^{M69E,I72E,F76E}$) were methylated via a passage through *E.coli* ER1821 strain and then transformed into *S.acidocaldarius flaI* mutant strain (Δ*aapFΔflaI S.acidocaldarius*) for expression analysis. Additionally, *flaI* constructs containing deletions in either 1-29 amino acids ($flaI^{\Delta 1-29}$) or 61-126 amino acids ($flaI^{\Delta 61-126}$) were constructed and cloned into pSVA1450 in a similar method mentioned earlier.

Methylated expression plasmids were electroporated in electro-competent *S. acidocaldarius* cells (wild type cells MW001,Δ*aapFΔflaI*, Δ*aapE*) using a Gene PulserXcell (BioRad) with a constant time protocol with input parameters 1.5 kV, 25 µF, 600 Ω in 1 mm cuvettes. Before plating on uracil lacking and NZAmine containing plates, cells were regenerated for 30 min at 75°C in two-fold recovery solution (1% sucrose, 20 mM β-alanine, 1.5 mM malate buffer, pH 4.5, 10 mM $MgSO_4$).

Protein Purification

Before purification, the cells were thawed in the lysis buffer (50 mM HEPES-NaOH, pH 7.5, 150 mMNaCl, 10% glycerol) containing the complete EDTA-free protease inhibitor cocktail (1 tablet/ 50 ml of lysate; Roche, Mannheim, Germany). Cells were lysed by sonication with a Sonoplus HD3100 sonicator (BandelinSonorex Biotechnique, Germany) using probe HD3100. Cell debris was removed by centrifugation at 6,000 x g for 15 min followed by centrifugation at 150,000 x g for 30 min at 4°C in an Optima™ MAX-XP ultracentrifuge (rotor MLA 55) (Beckman Coulter, USA) to remove the membranes. The supernatant was incubated at 70°C for 20 min in a water bath. The sample was cooled on ice and precipitated proteins were removed by centrifugation at 150,000 x g for 30 min at 4°C in an Optima™ MAX-XP ultracentrifuge (rotor MLA 55) (Beckman Coulter, USA). For purification of His-tagged proteins, the supernatant was applied to a Ni^{2+}- affinity column (Native IMAC) on the Profinia™ protein purification system (BioRad Laboratories, Inc, USA). Bound protein was eluted in lysis buffer containing 500 mM imidazole. The eluted fraction was dialyzed overnight at 4°C against 50 mM HEPES-NaOH pH 7.5, 150 mM NaCl. The dialyzed fraction was loaded on a Superdex 200 10/300 gel filtration column on an AKTA-Purifier system (GE Health Care, Germany), equilibrated with 50 mM HEPES-NaOH pH 7.5, 150 mM NaCl (flow rate 0.5 mL/min). The untagged FlaI (or $FlaI^{E336A}$ mutant) was expressed, sonicated and heat treated, as described above. The protein was then precipitated at room temperature with 20% ammonium sulfate (w/v) for 30 min. The precipitated protein was re-suspended in buffer Blue A (50mM Tris-HCl pH 8.0, 200 mM NaCl, 2 mM EDTA) and loaded on a 5 mL HiTrap BLUE column in the AKTA-Purifier system (GE Health Care). The column was washed with buffer Blue A and the bound protein was eluted with buffer Blue B (50 mM Tris-HCl pH 8.0, 1 M NaCl, 2 mM EDTA). The eluted protein was further purified on a 5 mL HiTrap PHENYL column. The protein was eluted using a gradient from buffer Blue B to buffer PHENYL (50 mM Tris-HCl pH 8.0, 2 mM EDTA). The eluted protein was then concentrated (<2 mL) using an Amicon concentrator system (Milipore, USA), and loaded on a Superose 6 10/300 size-exclusion column equilibrated with PHENYL buffer and eluted with 50 mM HEPES-NaOH pH 7.5 with 150 mM NaCl.

Protein purity was monitored by reducing sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE). Gels were stained with Biosafe (BioRad) protein stain. Protein concentration was determined with the Bio Rad protein assay dye-dependent reagent (Bio Rad) and absorption at 280 nm.

Crystallization and Structure Determination of FlaI

FlaI^{E336A} (8 mg/mL) was crystallized in the presence of 2.5 mM ATP and 5 mM $MgCl_2$ in 7.5% PEG 335 MME, 32.5% ethylene glycol, 82 mM imidazole, 18 mM malate (pH ~7.6), by sitting drop vapor diffusion. Crystals were flash cooled in liquid nitrogen and diffraction data was collected to 2 Å at beamline 8.3.1 at the ALS. Wild-type FlaI (15 mg/mL) was crystallized in 100mM HEPES pH 7.0, 22% Polyvinylpyrrolidone K 15, 10 mM $CoCl_2$ and 15% glycerol using the hanging drop vapor diffusion. Crystals were flash frozen in liquid nitrogen and diffraction data was collected to 3.6 Å at the SIBYLS beamline at the ALS. Datasets were processed with ELVES (Holton & Alber, 2004) and HKL2000 (Otwinowski & Minor, 1997). For phase determination FlaI^{E336A} was produced in its Selenomethionine-substituted form (SeMet-FlaI), which crystallized in the same space group as the native protein. The structure was solved with the AutoSol package implemented in PHENIX (Adams et al, 2010) and refined by iterative cycles of manual model building in COOT and computational refinement with PHENIX. The wild-type structure was solved by molecular replacement using PHASER (McCoy, 2007) with the CTD of FlaI^{E336A} (residues 228-513) as first search model followed by a search for the NTD (residues 2-224). Due to the low-resolution this structure was only refined by the rigid body refinement settings in PHENIX. Data collection and refinement statistics are shown in Supplementary Table S1.

Small Angle X-ray Scattering (SAXS*)*

SAXS data were collected at the ALS beamline 12.3.1 (SIBYLS) Lawrence Berkeley National Laboratory, Berkeley, CA. Using a wavelength of $\lambda = 1.0$ Å and a sample-to-detector distance of 1.5 m, scattering vectors (q) ranging from 0.01 to 0.32 $Å^{-1}$ were collected. The scattering vector is defined as $q = 4\pi \sin\theta/\lambda$, where 2θ is the scattering angle. All experiments were performed at 20°C, and data were processed as described (Putnam et al, 2007). The buffer used in SAXS experiments was 50 mM HEPES-NaOH pH 7.5, 200 mM NaCl and 5% glycerol, and a buffer blank was collected prior to FlaI samples for buffer subtraction. SAXS data of FlaI was collected using short (0.5 s) and long (2 s) exposure times at three different concentrations, with the lowest being the peak fraction of the gel filtration and the medium and highest being concentrated to half and a quarter volume, respectively. For experiments in the presence of nucleotides, FlaI was mixed at a 1:1 ratio with buffer containing 5 mM ATP or ADP. Pair distribution functions (P(r)) for the experimental SAXS data for different protein concentrations were calculated by the program GNOM (Svergun, 1992). Different possible conformations of the FlaI hexameric ring were created using SymmDock (Schneidman-Duhovny et al, 2005) with chain A or chain C as an input molecule, the conformation that fitted the SAXS data best was determined using the FoXS and MES web server (Schneidman-Duhovny et al, 2010).

In-frame Gene Deletion

In order to generate an N-terminal in frame deletion of the first 672 nt in *flaI* (corresponding to amino acids 1-224) ~800 bp of the up- and downstream flanking regions of N-terminal gene fragment (1-224 amino acids) of *flaI* were amplified by PCR using the corresponding primer pairs (ΔN*fla*-us-fw/ ΔN*fla*-us-rev and ΔN*fla*-ds-fw/ ΔN*fla*-ds-rev) listed in Table 2. By overlap extension PCR the flanking regions were fused to the in-frame deletion product. This PCR product was then cloned into the suicide vector pSVA406 yielding pSVA329 (Δ*FlaI*: Δ*saci_1173*). The final deletion construct was confirmed by sequencing analysis. The deletion plasmid pSVA329 was methylated prior to the electroporation into the uracil auxotrophic strain of *S. acidocaldarius* Δ*aapF* following the protocol described earlier. After five days of growth on plates without uracil, insertion of the plasmids into the genome was verified by colony PCR by amplification of the *pyrEF* cassette using the primers 928 and 936. Finally, to generate the deletion strain, clones harboring the gene fragment targeting plasmid were streaked on counter selection medium. The second recombination step was selected in the presence of 5-FOA and uracil. Uracil auxotrophic colonies could be obtained on plates after 5 days incubation at 75°C. The in-frame deletion strain containing a deletion of 672 nt at the N-terminus of *flaI* was confirmed by PCR and DNA sequencing using primer pairs *A* and *B*.

Gene Complementation

The deletion mutant strains and recombinant plasmids were constructed using standard molecular biology techniques described under the 'cloning and expression' and 'Construction of plasmids for gene deletion' sections and are listed in table 2. To complement deletion mutants (Δ*aapF*Δ*flaI*) different recombinant constructs were electroporated into the respective deletion mutant strains. To complement Δ*aapF*Δ*flaI*, wild type *flaI* constructs pSVA336/pSVA351/pSVA368, Walker A mutant (K268A) of *flaI* pSVA337/pSVA352, Walker B mutant (E336A) of *flaI* pSVA335/pSVA353, triple point (M69E,I72E,F76E) mutant of *flaI* pSVA370, Δ1-29 *flaI* pSVA355 and Δ61-126 *flaI* pSVA370 were electroporated in electro-competent Δ*aapF*Δ*flaI* cells. After electroporation cells were grown in Brock medium containing 0.1% N-Z-amine and 0.1-0.4% maltose. In pCMalLacS derived expression plasmids, pSVA1423/pSVA1450, the target gene is cloned under the control of the maltose-inducible promoter malA. The expression of the target genes was verified by SDS-PAGE followed by Western blotting using either α-Strep Ab or α-His Ab. The expression of *flaB* (the flagellin) was tested using α-FlaB-S1/S2 Ab and used as a marker for internal control as well as a control for checking purity of the membrane fractions tested in the present study.

Cell Fractionation and Localization Analysis

To understand to role of highly conserved and specific residues in FlaI protein, we attempted cellular fractionation of the complemented deletion mutant (Δ*aapF*Δ*flaI*) expressing different variations of FlaI protein [wild type *flaI* constructs pSVA336/pSVA351/pSVA368, walker A mutant (K268A) of *flaI* pSVA337/pSVA352, Walker B mutant (E336A) of *flaI* pSVA335/pSVA353, triple point (M69E,I72E,F76E) mutant of *flaI* pSVA370, Δ1-29 *flaI* pSVA355 and Δ61-126 *flaI* pSVA370].

To fractionate membrane and cytoplasm, 400 mL of logarithmically growing *S. acidocaldarius* culture (OD_{800}=0.8) was harvested by centrifugation (4,500 x g, 30 min). The cell pellet was resuspended in 50 mM HEPES-NaOH pH 7.5, 150 mM NaCl, 10% glycerol containing the complete EDTA-free protease inhibitor cocktail (1 tablet/50 mL of lysate; Roche, Mannheim, Germany). Cells were lysed using sonication with a Sonoplus HD3100 sonicator (Bandelin Sonorex Biotechnique, Germany) using probe HD3100. Cell debris was removed by centrifugation at 6,000 x g for 20 min. The supernatant was then centrifuged at 150,000 x g for 30 min at 4°C in an Optima™ MAX-XP ultracentrifuge (rotor MLA 55) (Beckman Coulter, USA) to pellet the membrane. The supernatant was treated as the cytosolic fraction in the following analysis. On the other hand, the resulting membrane pellet was resuspended in 50 mM HEPES-NaOH pH 7.5 containing 2% DDM and the complete EDTA-free protease inhibitor cocktail (1 tablet/50 mL of lysate; Roche, Mannheim, Germany). The resuspended membrane was kept on ice for 30 min for better solubilization and then flash-frozen in liquid nitrogen. Both the membrane and cytoplasmic fractions were estimated for protein content using the BioRad protein estimation kit. The purity of the membrane and cytosolic fractions was confirmed using western blotting using the anti-FlaB-S1/S2 antibodies.

Swimming Assays on Semi-Solid Plates

Swimming motility on semi-solid surface was analyzed on 0.15% gelrite plates supplemented with 0.001% tryptone. Approximately 10^7 exponentially growing cells were harvested from Brock medium without tryptone and used as inoculum for the assay. The inoculated plates were incubated for 3 days in a humid chamber at 75°C. Swimming behavior of different *S. acidocaldarius* strains was assessed by measuring the swimming radius.

Thermomicroscopy

To determine swimming velocities of different *S. acidocaldarius* strains, the cells were analyzed using a high temperature phase-contrast microscope (Olympus BX-50). The microscope consisted of a heatable sample stage (max. 90°C) and an object lens heating system (max. 80°C) in combination to a circulating air heating chamber (max. 50°C) surrounding the microscope. All the temperatures were regulated independently using a

multi-channel control unit. The sample capillaries were fed with the cell suspension and sealed at both the ends. Motility video clips up to 15 s were obtained with a PCO Sensicam 1600 CCD camera. Movies were analyzed with ImageJ 1.43u software (NIH) using the manual tracking plug-in.

Supplemental References

Berkner, S., Wlodkowski, A., Albers, S.V., and Lipps, G. (2010). Inducible and constitutive promoters for genetic systems in Sulfolobus acidocaldarius. Extremophiles *14*, 249-259

Gouet, P., Courcelle, E., Stuart, D.I. and Metoz, F. (1999). ESPript: multiple sequence alignments in PostScript. Bioinformatics. *15*, 305-8

Henche, A.L., Koerdt, A., Ghosh, A., and Albers, S.V. (2012). Influence of cell surface structures on crenarchaeal biofilm formation using a thermostable green fluorescent protein. Environmental microbiology *14*, 779-793

Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R, Thompson JD, Gibson TJ and Higgins DG (2007). ClustalW and ClustalX version 2. Bioinformatics *23*, 2947-2948

Miroux, B. and Walker, J. E. (1996). Over-production of proteins in Escherichia coli: mutant hosts that allow synthesis of some membrane proteins and globular proteins at high levels. Journal of molecular biology *260*, 289-298

Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C., and Ferrin, T.E. (2004). UCSF Chimera--a visualization system for exploratory research and analysis. J Comput Chem *25*, 1605-1612.

Schrodinger, LLC (2010). The PyMOL Molecular Graphics System, Version 1.3r1.

Wagner, M., van Wolferen, M., Wagner, A., Lassak, K., Meyer B., Reimann, J., and Albers, S.V. (2012). Versatile genetic tool box for the crenarchaeote Sulfolobus acidocaldarius, Frontiers in Microbiology, *3*, 214

Zolghadr, B., Weber, S., Szabo, Z., Driessen, 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. Molecular Microbiology *64*, 795-806

3.3 Insights into subunit interactions in the *Sulfolobus acidocaldarius* archaellum cytoplasmic complex

Summary:

The previously described motor protein FlaI and the Walker ATPase-like protein FlaH are the only two cytoplasmic components of the crenarchaeal archaellum (Jarrell & Albers, 2012). All the remaining five Fla proteins of *S. acidocaldarius* are predicted to carry transmembrane domains. Furthermore FlaH, FlaI and the polytopic membrane protein FlaJ are highly conserved among all archaellated archaea and therefore believed to be involved in the formation of the core structure of the archaellum (Ghosh & Albers, 2011a). FlaI and FlaJ show moreover homology to T4P core components PilB/T and PilC, believed to be the central part of the pillus assembly machinery (Jarrell & Albers, 2012). Additionally, recently published data have shown that FlaX can form ring like structures with 30 nm diameter, which interacts with FlaI (Banerjee et al., 2012a). Taken together we proposed that the archaellar basal body is composed of FlaX, FlaH, FlaI and FlaJ. To confirm this model we have performed detailed interaction studies between these purified Fla proteins *in vitro*.

Using Micro Scale Thermophoresis (MST) we have confirmed that FlaX, FlaH and FlaI interact with each other. All the protein pairs have interacted with high affinities with K_D values in the nM range. We have expressed the two N and C-terminal domains of FlaI separately and tested both of them for interaction with FlaXc. In pull-down assays the N-terminal and the C-terminal domain of FlaI can interact with FlaX individually, what led us to hypothesize that the FlaI hexamer may function inside the FlaX rings. The diameter of the FlaX ring is 30 nm and would provide enough space for the FlaI hexamer which is 14 nm in diameter.

We also isolated a stable FlaXc-FlaH-FlaI ternary complex *in vitro*. After mixing purified FlaX and FlaI, both strep tagged, with His tagged FlaH, we were able to pull-down all combined components using Ni-resin. This result confirmed that all three proteins, FlaXc-FlaH-FlaI form a tight complex *in vitro*. The obtained results led us to propose that the three proteins FlaX, FlaH and FlaI form the basal body of the crenarchaeal archaellum, in which FlaX act as a membrane bound cytoplasmic scaffold for its assembly.

Contributions:

Ankan Banerjee purified FlaX and FlaI. Tomasz Neiner purified FlaH. Under supervision of Sonja-Verena Albers, both designed experimental procedures, performed interaction studies and wrote the manuscript. Patrick Tripp was involved in performing interaction studies, including MST analysis and revised the manuscript.

the FEBS Journal

Insights into subunit interactions in the *Sulfolobus acidocaldarius* archaellum cytoplasmic complex

Ankan Banerjee, Tomasz Neiner, Patrick Tripp and Sonja-Verena Albers

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

Keywords
archaea; archaellum; basal body; flagella; type IV pili

Correspondence
S. V. Albers, Molecular Biology of Archaea, Max Planck Institute for Terrestrial Microbiology, Karl von Frisch Straße 10, 35043 Marburg, Germany
Fax: +49 6421178429
Tel: +49 6421178426
E-mail: albers@mpi-marburg.mpg.de

(Received 24 June 2013, revised 12 August 2013, accepted 9 September 2013)

doi:10.1111/febs.12534

Archaella are the archaeal motility structure that is the functional pendant of the bacterial flagellum but is assembled by a mechanism similar to that for type IV pili. Recently, it was shown by Banerjee *et al.* that FlaX, a crenarchaeal archaellum subunit from *Sulfolobus acidocaldarius*, forms a ring-like oligomer, and it was proposed that this ring may act as a static platform for torque generation in archaellum rotation [Banerjee A *et al.* (2012) *J Biol Chem* **287**, 43322–43330]. Moreover, the hexameric crystal structure of FlaI was solved, and its dual function in the assembly and the rotation of the archaellum was demonstrated [Reindl S *et al.* (2013) *Mol Cell* **49**, 1069–1082]. In this study, we show by biochemical and biophysical techniques that FlaX from *S. acidocaldarius* acts as a cytoplasmic scaffold in archaellum assembly, as it interacts with FlaI as well as with the recA family protein FlaH, the only cytoplasmic components of the archaellum. Interaction studies using various truncated versions of FlaI demonstrated that its N- and C-termini interact with FlaX. Moreover, using microscale thermophoresis, we show that FlaI, FlaX and FlaH interact with high affinities in the nanomolar range. Therefore, we propose that these three proteins form the cytoplasmic motor complex of the archaellum.

Structured digital abstract

- FlaH and FlaI bind by mst (View interaction)
- FlaXc physically interacts with FlaI by pull down (1, 2)
- FlaI and FlaXc bind by mst (1, 2)
- FlaI binds to FlaXc by pull down (View interaction)

Introduction

Sulfolobus acidocaldarius is a hyper-thermoacidophilic model organism that has been isolated from solfataric hotsprings and is one of the best characterized crenarchaea. In response to changes in the environment, *S. acidocaldarius* expresses various surface appendages; for example, archaeal adhesive pili, which are type IV pili (T4P), are used for attachment during favorable conditions [1,2]. In addition, archaella (archaeal flagella) are produced to evade unfavorable conditions by swimming (motility) [3,4]. The archaellum is the best-characterized surface structure in archaea [4].

The archaellum is structurally similar to T4P of Gram-negative bacteria, but functionally resembles bacterial flagella as it is used for swimming (motility). The homology between the polytopic membrane protein FlaJ in the archaella assembly system and PilC in the T4P assembly systems, the similarity of the P-loop ATPases FlaI and PilT/PilB, the proximal assembly of the subunits FlaB and PilA in the archaellum and T4P

Abbreviations
CTD, C-terminal domain; MST, microscale thermophoresis; NTD, N-terminal domain; T2SS, type II secretion system; T4P, type IV pilus.

filaments, respectively, the narrow diameter of the filament and the lack of a central lumen led to the hypothesis that archaella and T4P are assembled by similar processes [5–7].

Genetic and physiological experiments have shown that deletion of any of the archaella-related proteins leads to non-archaellated cells in all studied phyla [8–11]. In *S. acidocaldarius*, the archaellum is 12–14 nm in diameter (Fig. 1A) and comprises seven proteins that are encoded by a single gene locus (Fig. 1C). The archaellin FlaB is the filament-forming component of the archaellum cluster. The function of the monotopic membrane proteins FlaF and FlaG is unknown, but they are present in all archaellated archaea. The genetic order of *flaHIJ* genes is well conserved in all archaella operons [7]. FlaJ is a polytopic membrane protein that forms the membrane core of the archaellum machinery, while FlaH is a RecA family protein containing a classical Walker A motif and an incomplete Walker B motif [7]. Therefore, FlaH may be involved in nucleotide binding and might be involved in regulation of archaellum assembly or rotation. So far, FlaH has only been found in its monomeric form (T.N. and S.-V.A., unpublished results). On the other hand, it has been shown that *S. acidocaldarius* FlaI forms hexameric species in an ATP-dependent manner [12]. The crystal structure of FlaI was recently solved [13], showing that FlaI has two separate globular domains (the N- and C-terminal domains; NTD and CTD, respectively). The NTD and CTD together form a crown-like structure, in which the CTD is the base and the NTD forms the tip of the crown [13]. The first 29 amino acids on the NTD of FlaI were unstructured in the crystal structure, and deletion of these 29 amino acids of FlaI led to non-motile cells; however, archaella were assembled in this mutant and therefore it was proposed that FlaI is involved in archaellum assembly as well as archaellum rotation [13]. On the other hand, it was shown that FlaX, a *Sulfolobales*-specific archaellum subunit, forms a ring-shaped oligomeric structure of 30 nm diameter *in vitro*. The three C-terminal helices were important for ring closure, and it was also demonstrated that the C-terminal half of the protein interacts with the ATPase FlaI [14]. A current model of the archaellum is shown in Fig. 1B.

In the present study, we characterized the subunit interactions in the *S. acidocaldarius* motor complex forming the basal body of *S. acidocaldarius* archaella. We demonstrate that FlaX may be a platform for

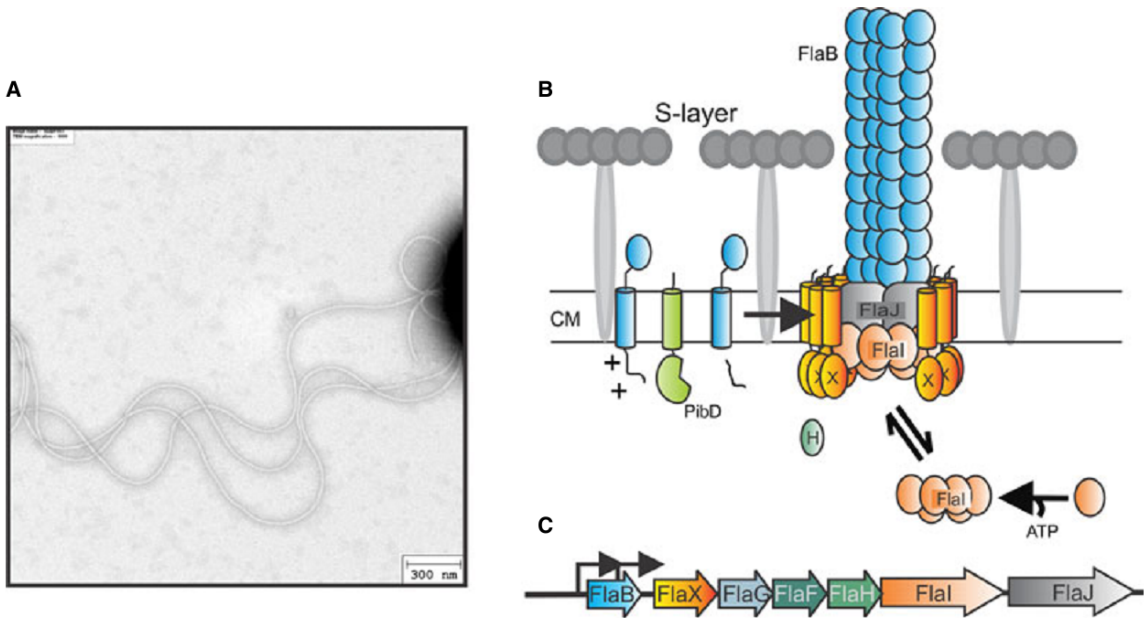


Fig. 1. (A) Electron micrograph image of *Sulfolobus acidocaldarius* showing several archaealla. (B) Present model of the archaellum of *S. acidocaldarius*, in which FlaX, a monotopic membrane protein, forms a ring complex [14]. CM, cell membrane. The S-layer is a proteinaceous layer that enfolds the cell membrane [23]. FlaI forms hexameric species in an ATP-dependent manner, and interacts with the lipids and with FlaX rings [12,14]. PibD is the membrane-bound class III signal peptidase [24]. (C) *S. acidocaldarius* archaella gene cluster. The major transcript is *flaB*, and the second transcript is expressed as *flaXGFHIJ*, as indicated by the black arrows [8].

the motor complex. Using a fluorescence-based biophysical technique, microscale thermophoresis (MST), the interaction of FlaXc (an N terminal 37 amino acid deletion of FlaX) and FlaI was studied in detail. In further experiments, FlaXc interacted with FlaI and FlaH, and FlaI interacted with FlaH. As both the NTD and CTD of FlaI were able to bind to the FlaX ring, we propose that the 14 nm diameter FlaI hexamer fits into the 30 nm diameter FlaX ring. Furthermore, a ternary complex comprising FlaI, FlaX and FlaH was also isolated *in vitro*, which further supports the hypothesis that the FlaX ring acts as a cytoplasmic scaffold in archaellum assembly.

Results

Archaella subunits have a high affinity towards each other

The stability of FlaX in *S. acidocaldarius* is dependent on the proteins FlaJ, FlaH, FlaI and also FlaB [8], which may therefore interact with FlaX. FlaX interaction with FlaI has been extensively demonstrated using cell lysate interaction, co-expression pull-down analysis and purified proteins [14]. However, using MST [15], we were able to assess the binding affinities of the two proteins. MST is a fluorescence-based biophysical technique that may be used to determine the affinities of molecular interactions (see Experimental procedures for details). The results confirmed interaction between FlaX and FlaI (Fig. 2A). In this experiment, 11 μM FlaI was labeled with 65 μM of an amine-reactive red *N*-Hydroxy succinamide fluorescent dye. Upon mixing varying concentrations of unlabeled FlaXc (40 μM–1.22 nM) with 7.5 nM labeled FlaI at a steady-state level under infrared light-mediated local heating, positive thermophoresis of monomeric FlaI was observed that was dependent on changes in the local hydrodynamic radius. The raw data for the capillary scan and the MST profiles of each experiment are shown in Fig. S1. The relative fluorescence was normalized using the mean of F_{max} (maximum fluorescence before infrared exposure, cold region). The nanotemper response (hot/cold) [15] after thermophoresis was plotted against the concentration of the unlabeled molecule, the raw data were analyzed using the dose-response sigmoidal fit function, and the dissociation constant (K_D) was determined using the half-maximum effective concentration (EC_{50}) or the half-maximum inhibitory concentration (IC_{50}) of the interaction (Fig. 2A–C). In this experiment, FlaXc bound FlaI with very high affinity, with a K_D value of 86 ± 32 nM and a Hill coefficient of 0.92 ± 0.33

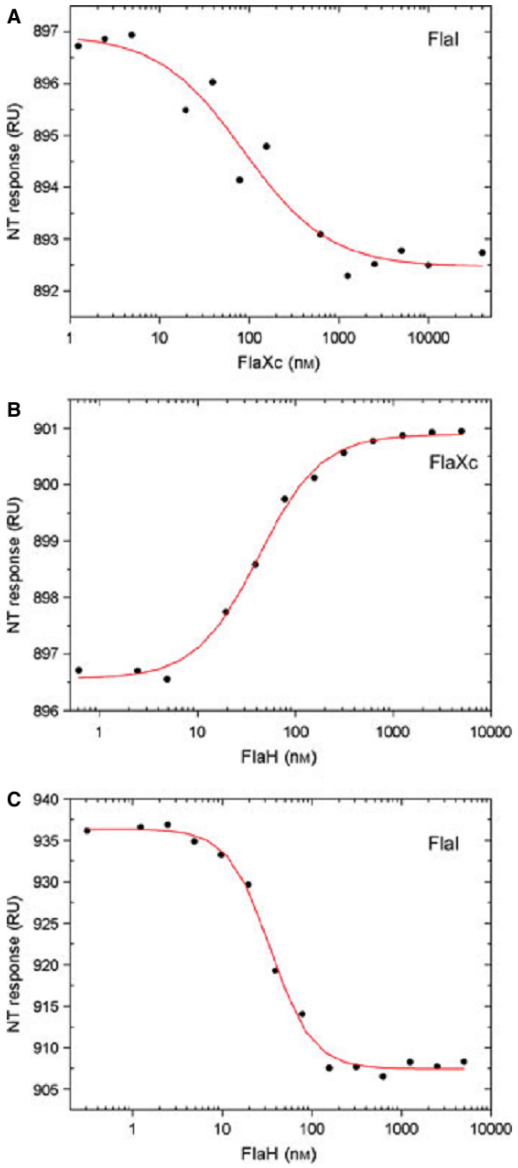


Fig. 2. Archaellum subunit interactions. (A) FlaXc interacted with the NT-647-labeled FlaI monomer in the MST experiment. Unlabeled FlaXc (40 μM–1.22 nM) was titrated in 7.5 nM labeled FlaI, and thermophoresis was performed. The data were analyzed by plotting the NT response (hot/cold) of thermophoresis against the FlaXc concentration using a non-linear fit (sigmoidal dose-response). FlaXc interacts with FlaI with a K_D of 86 ± 32 nM and a Hill coefficient of 0.92 ± 0.33. (B) Binding of FlaH and FlaXc. Labeled FlaXc (25 nM) was used in a thermophoresis experiment with 5 μM–0.15 nM unlabeled FlaH. The K_D was 42 ± 3.1 nM and the Hill coefficient was 1.36 ± 0.13. (C) MST experiment using 35 nM labeled FlaI monomer with 5 μM–0.31 nM unlabeled FlaH, showing that FlaI interacts with FlaH with an affinity (K_D) of 34.3 ± 2.18 nM and a Hill coefficient of 1.86 ± 0.19.

(Fig. 2A). In order to confirm our hypothesis that FlaX also interacts with FlaH, we performed a further MST experiment. Fluorescent labeling of FlaH was unsuccessful; therefore, purified FlaXc was dialyzed against 25 mM carbonate/bicarbonate pH 9.7, 200 mM NaCl buffer, for better efficiency of amine reactive labeling. The concentration of labeled FlaXc was kept constant at 25 nM, and a variable concentration of FlaH (5 μM–0.15 nM) was titrated at steady-state levels. Figure 2B shows that FlaH bound to FlaXc with a K_D value of 42 ± 3.1 nM and a Hill coefficient of 1.36 ± 0.13, indicating positive cooperativity of the interaction. As FlaXc interacts with both cytoplasmic proteins FlaI and FlaH, we investigated whether FlaH and FlaI interact with each other. We therefore performed MST analysis using 35 nM labeled FlaI and variable concentrations of FlaH ranging from 5 μM to 0.31 nM. These proteins showed interaction with the highest affinity measured so far (K_D 34.3 ± 2.18 nM, Fig. 2C). The Hill coefficient of 1.86 ± 0.19 also demonstrates positive cooperativity.

Molecular architecture of the FlaI hexamer and its position in the FlaX ring complex

The crown-shaped FlaI hexamer has a diameter of 14 nm. The base of the crown mainly comprises the CTD of FlaI, which contains the active sites for ATP hydrolysis, and the tip of the crown comprises the NTD [13]. FlaI contains several domains that may play important roles in the dual functionality of FlaI. The extreme N-terminal region was unstructured, and the first 29 amino acids were shown to be involved in the switch for assembly to rotation [13]. The unstructured region is followed by a triple helix domain (amino acids 61–127), which is believed to play a key role in interaction with the polytopic membrane protein FlaJ (Fig. 3A).

To identify the region of FlaI interacting with the FlaXc ring, we generated two truncated variants of FlaI, $FlaI^{\Delta N}$ and $FlaI^{\Delta C}$, i.e. CTD and NTD deletions of FlaI, respectively (Fig. 3A). The proteins were purified using either Ni^{2+} affinity or streptavidin-based affinity chromatography (see Experimental procedures for details). In the first sets of experiments, we used purified protein-mediated interaction analysis as described in Experimental procedures. Equal amounts of pure His_6-tagged FlaXc were mixed with Strep II-tagged $FlaI^{\Delta C}$ and loaded onto streptavidin beads, followed by washing and elution steps. The results showed that the NTD of FlaI interacts with the FlaXc ring (Fig. 3B). Interestingly, when we performed a cell lysate interaction analysis using His_6-tagged $FlaI^{\Delta N}$ and strep II-tagged FlaXc on Ni-NTA beads, we observed co-elution of both proteins (Fig. 3C), which indicates that FlaI also binds the FlaXc ring through its CTD. A control experiment using the same protocol was also performed for strep II-tagged FlaXc on Ni-NTA beads (Fig. 3D), showing that the protein did not bind to the column material.

It was previously shown that FlaI maintains a hexameric conformation upon binding to the non-hydrolyzable ATP analog adenosine 5′-(β,γ-imido) triphosphate [12]. Therefore, monomeric FlaI (N-terminal His_6-tagged) was incubated with 25 μM adenosine 5′- (β,γ-imido) triphosphate and then incubated with strep II-tagged FlaXc. Ni-NTA beads were used to bind hexameric FlaI, and Fig. 3E shows that FlaXc and hexameric FlaI co-elute.

The FlaXc ring is the scaffold for archaellum assembly

Archaellum rotation depends on the local ATP concentration. In *S. acidocaldarius*, the dual-function ATPase FlaI provides the energy for both archaellum assembly and swimming motility by hydrolyzing ATP [13]. Above, we have shown that both FlaH and FlaI interact with the FlaXc ring. Moreover, we showed that FlaI interacts with FlaXc via both the NTD and the CTD. We then wished to determine whether these three proteins can interact together. To isolate the ternary complex, we used strep II-tagged FlaXc, His_6-tagged FlaH and strep II-tagged FlaI. The three proteins were mixed and incubated with Ni^{2+} affinity beads (see Experimental procedures for details). After extensive washing, the bound protein fraction was eluted and further characterized using Coomassie-stained SDS/PAGE and western blot analysis using α-His and α-Strep antibodies. The results indicate that the three proteins FlaI, FlaX and FlaH, also forms a ternary complex *in vitro* (Fig. 4).

Discussion

Type II secretion systems (T2SS), type IV pili (T4P) and archaella assembly apparatuses are membrane-associated complexes that require ATP hydrolysis for their assembly and/or function. The energy required for all of these membrane-associated nanomachines is provided by soluble ATPases [4,7,16]. Association of soluble ATPases with membrane proteins is an essential step for proper assembly or secretion of proteins in T4P assembly or T2SS, respectively. For T2SS, it was shown that one of the membrane-integrated proteins e.g. *Vibrio cholerae* EpsL, recruits the ATPase, e.g. EpsE, to the membrane by interaction with its C-terminal domain [17]. Monotopic EpsL stimulated the activity of EpsE by 10 000-fold when EpsL and phospholipid

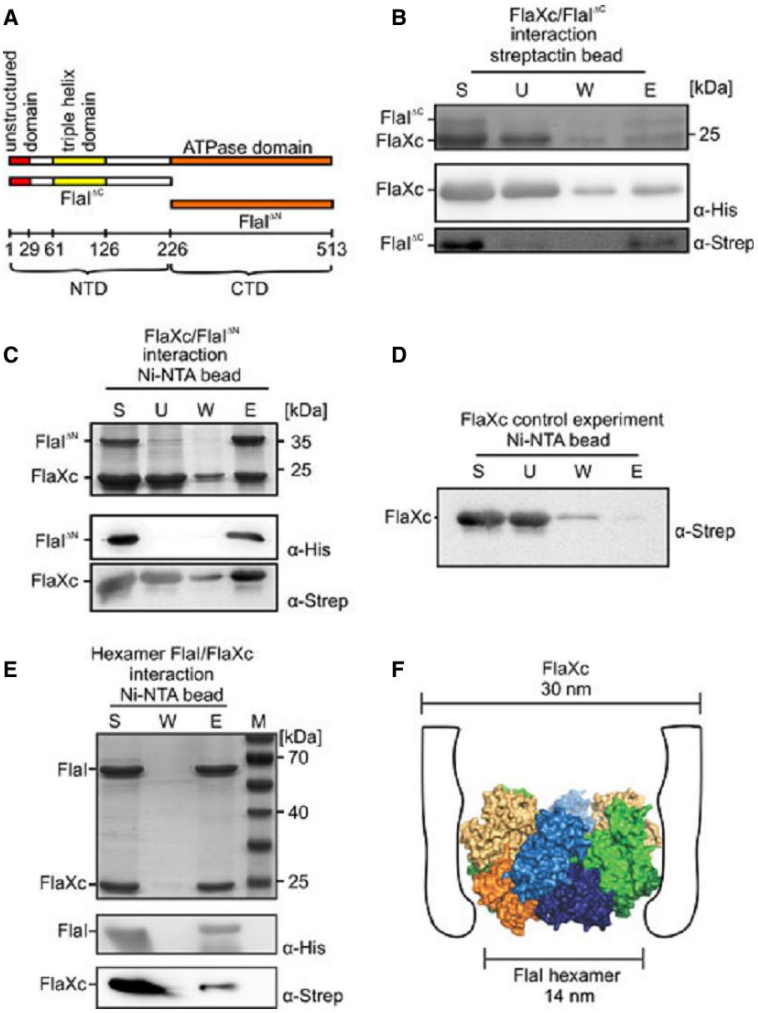


Fig. 3. Molecular architecture of the FlaI hexamer, and its position in the FlaX core complex. (A) Molecular architecture of FlaI monomer as determined previously [13] and of the truncated versions constructed in present study, $FlaI^{\Delta C}$ and $FlaI^{\Delta N}$. (B) Binary complex formation using pure protein interaction between FlaXc (C-terminal His_6-tagged) and $FlaI^{\Delta C}$ (deletion of amino acids 227–513 of FlaI) with an N-terminal Strep II tag. The co-elution of the proteins suggests direct interaction. (C) Refolded inclusion body of the N-terminal Strep II-tagged FlaXc and $FlaI^{\Delta N}$ (His_6-tagged) interaction was performed using Ni-NTA affinity beads. Co-elution of FlaXc and FlaI was confirmed using Coomassie-stained SDS/PAGE and immunoblotting with α-His and α-strep antibodies. (D) Control experiment for Strep II-tagged FlaXc (refolded inclusion bodies) using Ni-NTA beads. (E) *In vitro* interaction analysis of His_6-tagged FlaI hexamer and strep II-tagged FlaXc. His_6-tagged FlaI was hexamerized using the non-hydrolyzable ATP analog adenosine 5′- (β,γ-imido) triphosphate (25 μM). After mixing the proteins, a binary complex eluted through the Ni-NTA column. S, starting material; U, unbound fraction; W, pooled wash fractions; E, elution; M, marker. (F) Hypothetical spatial localization of the FlaI hexamer in the FlaXc ring.

membrane were mixed with EpsE [18]. Similar results were observed for the archaella assembly system, for which the activity of the ATPase FlaI was enhanced by *S. acidocaldarius* tetraether-linked lipids [12].

The data presented in this study show that FlaX, FlaH and FlaI interact with each other with high affinities (Fig. 5A), and we propose that the FlaXc ring may act as a platform during *S. acidocaldarius* archaellum assembly. Figure 5C shows our present model, in which the FlaX ring acts as a membrane-bound cytoplasmic assembly platform for subsequent assembly of the FlaI hexamer and FlaH. The affinities of the subunit interactions were in the nanomolar range, which strongly suggests that archaella-associated proteins bind their interacting partners immediately after expression. Interestingly, the FlaXc/FlaH and FlaI/FlaH interactions revealed significant evidence for cooperativity. Given that the apparent affinity of the FlaI/FlaX complex is approximately twofold lower than that for formation of the other two complexes, this may represent a rate-limiting step for basal body assembly. In the pull-down experiment (Fig. 4), FlaX and FlaH interacted with the same stoichiometry, whereas FlaI was present in lower amounts in the elution fraction. This indicates that FlaI and FlaH compete for similar binding sites in FlaX, as, in the FlaI/FlaX interaction assay, FlaI and FlaX eluted with a 1 : 1 stoichiometry (Fig. 3). FlaX may therefore serve as the priming protein for proper archaella assembly in *S. acidocaldarius*, but is also dependent on the presence of FlaI, FlaH and FlaJ, as FlaX is instable in Δ*flaI* and Δ*flaH* deletion strains [8]. Together, these proteins probably form the motor complex of the *S. acidocaldarius* archaellum.

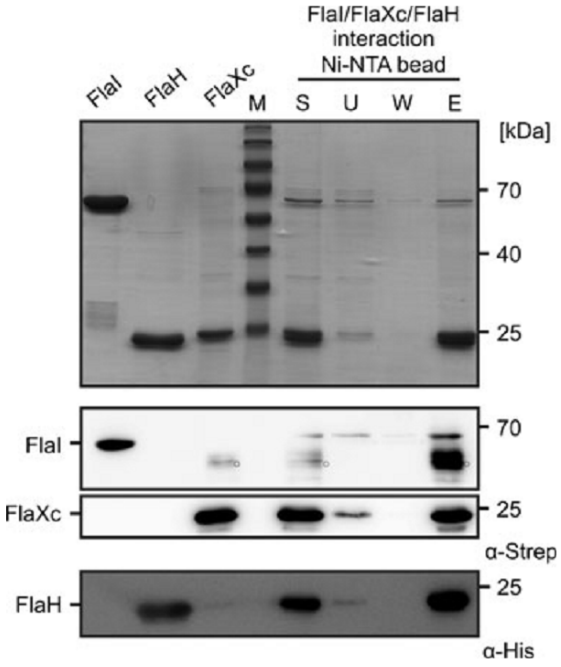


Fig. 4. FlaXc rings acts as a platform for archaella assembly. To form a ternary complex of FlaXc, FlaH and FlaI, His-tagged FlaH, Strep II-tagged FlaI and Strep II-tagged FlaXc were mixed and incubated with Ni-NTA affinity beads. The complex was eluted using 50 mM Tris/HCl pH 8, 150 mM NaCl, 500 mM imidazole buffer. The ternary complex was visualized by immunoblotting using α-His and α-strep antibodies. Individual proteins were loaded on the same SDS/PAGE to ensure the migration pattern of pure Strep II-tagged FlaXc, His_6-tagged FlaH and Strep II-tagged FlaI. S, starting material; U, unbound fraction; W, wash; E, elution; M, marker. 'O' represents the oligomer of FlaXc [14].

FlaX is a unique crenarchaeal archaellum component, and is not present in euryarchaeal genomes. However, euryarchaea possess three proteins (FlaC/D/E) that are not present in crenarchaeal genomes [4]. Therefore, we hypothesize that, in euryarchaea, these proteins may take over the function of FlaX in stabilizing the two conserved proteins FlaH and FlaI in the archaellum, but this remains an open question.

In T2SS and T4P assembly systems, it was predicted that the polytopic membrane protein GspF or PilC, respectively, interacts with the assembly and/or retraction ATPases through their positively charged cytoplasmic loops, termed cytoI or cytoII, which in turn facilitate membrane association of the ATPases [19,20]. FlaJ, the polytopic membrane protein is thought to be essential in providing the assembly platform for the archaellum in the membrane [7,21]. The polytopic membrane protein FlaJ shares homology to the T4P assembly platform-forming inner membrane protein, PilC. FlaJ of *S. acidocaldarius* has seven TMDs. The N-terminal cytoplasmic domain (cyto1) (amino acids 1–130) and the cyto2 domain (amino acids 233–359) have several stretches of positively charged amino acids, giving high theoretical pI values of 9.75 and 10.3 [13]. In contrast to crenarchaea, there is a putative Nterminal TMD present in euryarchaea that is followed by the cyto1 loop, a stretch of positively charged amino acids on the cytoplasmic side. Interestingly, the cytoplasmic loops show homology to each other [7], and the loops of FlaJ, PilC and GspF are similar [6]. Recently, it was proposed that these charged loops interact with the crown grooves of FlaI, which is partially negatively charged and has a theoretical pI of 5.75 (Fig. 5B). Previously, a model based on FlaI ATP hydrolysis and interaction with the cyto loops of FlaJ suggested generation of a possible 8–10 Å space in the membrane, which may lead to proximal incorporation of FlaB to a newly growing archaellum filament. [13]. This remains to be proven, but the *in vivo* stability of FlaX is also dependent on FlaJ and FlaB subunits. Therefore, we suggest that, in *S. acidocaldarius*, the FlaX ring may assemble around FlaJ, and that FlaI confers conformational change of FlaJ transfers to the membrane domain of FlaX ring, which leads to correct incorporation of archaellin (FlaB) into the newly growing archaellum filament.

This study provides an insight into how the cytoplasmic proteins FlaH and FlaI are assembled at the membrane in the archaella assembly complex, and will help to understand their function in assembly and rotation of the archaellum.

Experimental procedures

Plasmids and strains used in present study

The pET-based dual over-expression vector pET Duet1 (Novagen) was used for expression of the FlaXc, FlaI and FlaH proteins in *Escherichia coli* BL21-CodonPlus-RIL strain. Tables S1 and S2 list plasmid and primer details.

Protein expression and membrane isolation

The *E. coli* Codon Plus strain was freshly transformed with the pET-based vector and grown overnight in Luria-Bertani medium containing 50 $\mu g \cdot mL^{-1}$ ampicillin and 34 $\mu g \cdot mL^{-1}$ chloramphenicol. Then 1 mL of pre-culture was used to inoculate 1 L of Luria-Bertani medium supplemented with ampicillin (50 $\mu g \cdot mL^{-1}$) and chloramphenicol (34 $\mu g \cdot mL^{-1}$), and grown at 37 °C until an attenuance at 600 nm of 0.5–0.6 was reached. Then 1 mM isopropyl thio-β-D-galactoside was used to induce protein over-expression. All inductions were

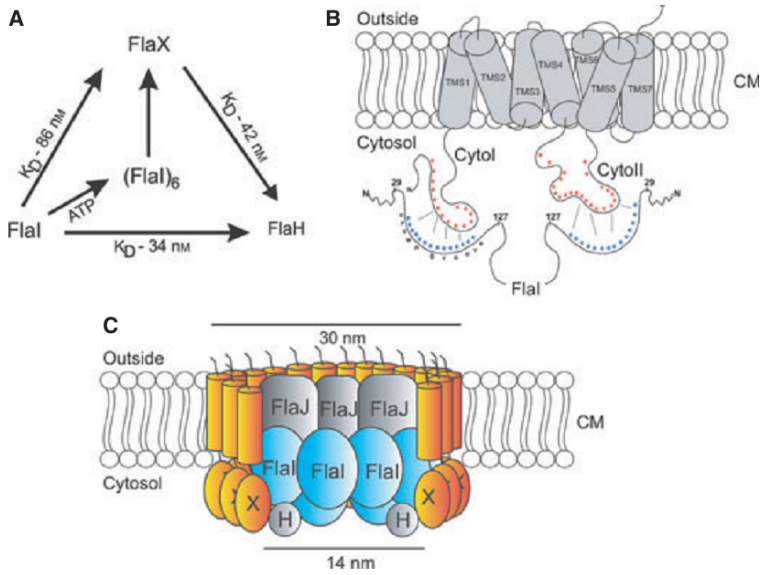


Fig. 5. FlaX is a membrane-bound cytoplasmic assembly platform. (A) Schematic overview of archaellum subunit interactions for building the motor complex. (B) The polytopic membrane protein FlaJ in *S. acidocaldarius* consists of seven transmembrane segments (TMS). The cytoplasmic loops cytoI and cytoII are highly positively charged (pI 9.75 and 10.3, respectively). The crystal structure of FlaI revealed that the crown grooves are negatively charged (pI 5.75) and constitute a triple helix domain. Reindl *et al.* [13] proposed that the positively charged cytoplasmic loops of FlaJ and the negatively charged crown grooves of FlaI interact with each other. CM, cell membrane. (C) Current working model of the archaellum motor complex. FlaX is a monotopic membrane protein, and Banerjee *et al.* [14] showed that the cytoplasmic domain (FlaXc) forms a 30 nm ring-like oligomer. FlaX binds to FlaI and FlaH with high affinities; FlaI and FlaH also interact with each other. The hexameric FlaI crystal exhibited a 14 nm diameter crown shape. As both the NTD and CTD of FlaI and also the hexameric FlaI interact with FlaXc, we propose that the FlaI hexamer is inside the FlaX ring.

performed at 16 °C overnight for controlled protein expression. The cells were collected by centrifugation at 10 000 ***g*** for 25 min at 4 °C, and were resuspended in 50 mM Tris/HCl pH 8, 150 mM NaCl, 1 x protease inhibitor cocktail mix. The resuspended cells were lysed using a Sonoplus HD3100 sonicator (Bandelin Sonorex Biotechnique, Berlin, Germany) with probe HD3100. Cell debris was removed by centrifugation at 10 000 ***g*** for 25 min at 4 °C, and the cell-free lysate was centrifuged at 100 000 ***g*** for 45 min at 4 °C to collect the membrane and the soluble fractions. The membrane fraction was solubilized using 50 mM Tris/HCl buffer, pH 8, 150 mM NaCl, 5% glycerol, 20 mM imidazole and 2% w/v *n*-dodecyl-β-maltoside, followed by incubation at 50 °C for 1–2 h for solubilization of the membrane proteins [14]. The fractions were analyzed by Coomassie-stained SDS/PAGE and immunoblot analysis using α-His, α-FlaX or α-FlaI antibodies.

Purification of recombinant proteins

His_6-tagged FlaH was purified using Ni-NTA beads. Purification was performed in 50 mM MES pH 6.2, 500 mM NaCl buffer, using a 20–500 mM imidazole gradient on a gravity column. The elution fraction was desalted in 50 mM MES pH 6.2, 500 mM NaCl buffer, concentrated and kept at −80 °C for future experiments. Strep II-tagged proteins (FlaI or $FlaI^{\Delta C}$) were purified using streptavidin beads. A concentration gradient of 2.5–5 mM desthiobiotin in 50 mM Tris/HCl, pH 8, 150 mM NaCl buffer, was used to elute Strep II-tagged FlaI from the column, and the pure sample was desalted and kept for future use. His_6-tagged FlaI and His_6-tagged $FlaI^{\Delta N}$ were purified as described previously [13]. Moreover, His_6-tagged FlaXc was purified using the protocol described previously [14] with a slight modification. His-tagged FlaXc was purified using an IMAC (immobilized metal ion affinity chromatography) column on a Bio-Rad (Berkeley, CA, USA) Profinia purification system, and the purified protein was desalted using a Profinia desalting column with 25 mM glycine/NaOH pH 10, 200 mM NaCl buffer and purified FlaXc was concentrated to 10 $mg \cdot mL^{-1}$.

Isolation and refolding of inclusion bodies

Upon heterologous over-expression, strep II-tagged FlaXc was present solely in inclusion bodies. To isolate strep II-tagged FlaXc, inclusion bodies were resuspended in buffer I (50 mM Tris/HCl pH 8, 150 mM NaCl, 3 M guanidium hydrochloride, 1% Triton X-100, 3% glycerol). Resuspended inclusion bodies were incubated on ice for 20–30 min. To ensure proper solubilization, resuspended inclusion bodies

were further sonicated and incubated under constant stirring on ice for 3 h or overnight. Solubilized inclusion bodies were separated from insoluble particles by centrifugation at 10 000 **g** for 30–60 min. Refolding of proteins from solubilized inclusion bodies was performed by three-step dialysis in buffer R (100 mM Tris/HCl pH 8, 500 mM L-arginine) by reducing the guanidium hydrochloride concentration (1, 0.25 and 0.125 M). To ensure proper refolding, dialysis was performed in buffer R only. To remove L-arginine, a final last dialysis step was performed using buffer A (100 mM Tris/HCl pH 8, 150 mM NaCl). Centrifugation at 13 000 ***g*** for 10 min was performed to ensure protein solubility. The soluble fraction was further dialyzed against 25 mM glycine/NaOH pH 10, 200 mM NaCl buffer, and concentrated to 5 mg·mL^{-1}. Refolded FlaXc was flash-frozen in liquid N_2 and kept at −80 °C for further use. Correct folding of FlaXc was tested by CD spectroscopy (Fig. S2).

Far-UV CD spectroscopy

The structural integrity of refolded strep II-tagged FlaXc was determined using CD spectroscopy, and CD spectra were recorded at room temperature using a Jasco (Easton, MD, USA) J-810 spectropolarimeter. Experiments were performed in 1 mm path length cells. Pure refolded inclusion body was dialyzed against 25 mM carbonate/bicarbonate pH 9.7, 200 mM NaCl buffer, to ensure no interference arising from the 'buffer only' spectrum. Confirmation of the structural integrity of the refolded inclusion body was obtained by acquisition of far-UV (200–280 nm) CD spectra. Far-UV spectra of the tested proteins are presented as the smoothed mean of three accumulations. Secondary structure prediction analysis was performed using the supplied J-810 software, using the Yang statistical algorithm [22]. All recorded CD spectra were baseline-corrected by subtraction of the 'buffer only' spectrum.

Binary or ternary complex formation

In vitro protein-protein complex formation was performed as described previously [14]. To form the binary or ternary complexes, pure proteins tagged with either His_6 or Strep II were incubated together in a 50 mM Tris/HCl pH 8, 150 mM NaCl buffer, and then applied to affinity beads (Ni-NTA; Sigma Aldrich, Seelze, Germany or streptavidin; IBA GmbH, Göttingen, Germany, respectively). Complexes were eluted with either 250–500 mM imidazole or 2.5 mM desthiobiotin, visualized by Coomassie-stained SDS/PAGE, and α-His or α-Strep antibodies.

Affinity determination using thermophoresis

Microscale thermophoresis (NanoTemper Technologies GmbH, Munich, Germany) is a fluorescence-based biophysical technique used to determine the binding affinities of proteins in solution by using infrared laser to induce a local temperature gradient in glass capillaries and thus to monitor thermophoretic movement. The detailed procedure has been described previously [15]. Purified His_6- tagged FlaI was precipitated using 80% ammonium sulfate to eliminate all bound nucleotides and to ensure that only monomeric FlaI was present. The precipitated protein was refolded in 50 mM HEPES pH 7.5, with 150 mM NaCl buffer. The ATPase activity of the monomeric FlaI was analyzed in order to confirm proper folding (data not shown). The FlaI monomer was fluorescently labeled using the amine-reactive Monolith NT™ protein labeling kit NT-647 (NanoTemper Technologies GmbH) according to the manufacturer's instructions. FlaXc was purified in glycine/NaOH buffer, and, as glycine interferes with the labeling efficiency of the dye, pure FlaXc was dialyzed against 25 mM carbonate/bicarbonate, pH 9.7, 200 mM NaCl buffer. FlaXc was labeled using the same protocol as FlaI. To determine the binding affinity of the FlaI/FlaH binary complex, 35 nM labeled FlaI was used and FlaH was titrated in a 1 : 2 serial dilution. For the FlaXc/FlaI interaction, 7.5 nM labeled FlaI was used, and for the FlaXc/FlaH interaction, 25 nM labeled FlaXc was used. The measurements were performed at 298 K, 80% LED power and 20% infrared-laser power, which induces a temperature increase of ~ 2 K. Samples were incubated for 10–20 min before measurements. The laser on and off times were adjusted to 30 and 5 s, respectively. The FlaI/FlaH interaction experiment was performed in 50 mM Tris/HCl pH 8 buffer, containing 150 mM NaCl and 0.05% Tween-20, the FlaXc/FlaI interaction experiment was performed in 25 mM glycine/NaOH pH 10 buffer, containing 200 mM NaCl and 0.05% Tween-20, and the FlaXc/FlaH interaction experiment was performed in 25 mM carbonate/bicarbonate pH 9.7 buffer, containing 200 mM NaCl and 0.05% Tween-20. The measurements were performed using a NanoTemper Monolith NT.115 instrument in standard treated capillaries, and analyzed using NANOTEMPER analysis software version 1.4.27 (NanoTemper Technologies GmbH). To generate representative curves, the raw data were extracted and plotted using ORIGIN 6.1 (OriginLabs, Northampton, MA, USA), and the K_D value and Hill coefficient were calculated using a non-linear sigmoidal (dose-response) fitting equation provided in the software.

Acknowledgements

A.B., P.T. and S.-V.A. were supported by intramural funds from the Max Planck Society. T.N. received support from European Research Council starting grant ARCHAELLUM/311523. We thank Anna-Lena Henche, Max Planck Institute for Terrestrial Microbiology, for providing the electron micrograph image of *Sulfolobus acidocaldarius*.

References

1 Henche AL, Koerdt A, Ghosh A & Albers SV (2012) Influence of cell surface structures on crenarchaeal biofilm formation using a thermostable green fluorescent protein. *Environ Microbiol* **14**, 779–793.

2 Henche AL, Ghosh A, Yu X, Jeske T, Egelman E & Albers SV (2012) Structure and function of the adhesive type IV pilus of *Sulfolobus acidocaldarius*. *Environ Microbiol* **14**, 3188–3202.

3 Lassak K, Ghosh A & Albers SV (2012) Diversity, assembly and regulation of archaeal type IV pili-like and non-type-IV pili-like surface structures. *Res Microbiol* **163**, 630–644.

4 Jarrell KF & Albers SV (2012) The archaellum: an old motility structure with a new name. *Trends Microbiol* **20**, 307–312.

5 Trachtenberg S & Cohen-Krausz S (2006) The archaeabacterial flagellar filament: a bacterial propeller with a pilus-like structure. *J Mol Microbiol Biotechnol* **11**, 208–220.

6 Peabody CR, Chung YJ, Yen MR, Vidal-Ingigliardi D, Pugsley AP & Saier MH Jr (2003) Type II protein secretion and its relationship to bacterial type IV pili and archaeal flagella. *Microbiology* **149**, 3051–3072.

7 Ghosh A & Albers SV (2011) Assembly and function of the archaeal flagellum. *Biochem Soc Trans* **39**, 64–69.

8 Lassak K, Neiner T, Ghosh A, Klingl A, Wirth R & Albers SV (2012) Molecular analysis of the crenarchaeal flagellum. *Mol Microbiol* **83**, 110–124.

9 Bayley DP, Florian V, Klein A & Jarrell KF (1998) Flagellin genes of *Methanococcus vannielii:* amplification by the polymerase chain reaction, demonstration of signal peptides and identification of major components of the flagellar filament. *Mol Gen Genet* **258**, 639–645.

10 Patenge N, Berendes A, Engelhardt H, Schuster SC & Oesterhelt D (2001) The *fla* gene cluster is involved in the biogenesis of flagella in *Halobacterium salinarum*. *Mol Microbiol* **41**, 653–663.

11 Chaban B, Ng SY, Kanbe M, Saltzman I, Nimmo G, Aizawa S & Jarrell KF (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.

12 Ghosh A, Hartung S, van der Does C, Tainer JA & Albers SV (2011) Archaeal flagellar ATPase motor shows ATP-dependent hexameric assembly and activity stimulation by specific lipid binding. *Biochem J* **437**, 43–52.

13 Reindl S, Ghosh A, Williams GJ, Lassak K, Neiner T, Henche AL, Albers SV & Tainer JA (2013) Insights into FlaI functions in archaeal motor assembly and motility from structures, conformations, and genetics. *Mol Cell* **49**, 1069–1082.

14 Banerjee A, Ghosh A, Mills DJ, Kahnt J, Vonck J & Albers SV (2012) FlaX, a unique component of the crenarchaeal archaellum, forms oligomeric ring-shaped structures and interacts with the motor ATPase FlaI. *J Biol Chem* **287**, 43322–43330.

15 Baaske P, Wienken CJ, Reineck P, Duhr S & Braun D (2010) Optical thermophoresis for quantifying the buffer dependence of aptamer binding. *Angew Chem Int Ed Engl* **49**, 2238–2241.

16 Korotkov KV, Sandkvist M & Hol WG (2012) The type II secretion system: biogenesis, molecular architecture and mechanism. *Nat Rev Microbiol* **10**, 336–351.

17 Sandkvist M, Keith JM, Bagdasarian M & Howard SP (2000) Two regions of EpsL involved in species-specific protein-protein interactions with EpsE and EpsM of the general secretion pathway in *Vibrio cholerae*. *J Bacteriol* **182**, 742–748.

18 Camberg JL, Johnson TL, Patrick M, Abendroth J, Hol WG & Sandkvist M (2007) Synergistic stimulation of EpsE ATP hydrolysis by EpsL and acidic phospholipids. *EMBO J* **26**, 19–27.

19 Py B, Loiseau L & Barras F (2001) An inner membrane platform in the type II secretion machinery of Gram-negative bacteria. *EMBO Rep* **2**, 244–248.

20 Karuppiah V, Hassan D, Saleem M & Derrick JP (2010) Structure and oligomerization of the PilC type IV pilus biogenesis protein from *Thermus thermophilus*. *Proteins* **78**, 2049–2057.

21 Jarrell KF, Stark M, Nair DB & Chong JPJ (2011) Flagella and pili are both necessary for efficient attachment of *Methanococcus maripaludis* to surfaces. *FEMS Microbiol Lett* **319**, 44–50.

22 Yang JT, Wu CS & Martinez HM (1986) Calculation of protein conformation from circular dichroism. *Methods Enzymol* **130**, 208–269.

23 Albers SV & Meyer BH (2011) The archaeal cell envelope. *Nat Rev Microbiol* **9**, 414–426.

24 Albers SV, Szabo Z & Driessen AJ (2003) Archaeal homolog of bacterial type IV prepilin signal peptidases with broad substrate specificity. *J Bacteriol* **185**, 3918–3925.

Supporting information

Additional supporting information may be found in the online version of this article at the publisher's web site:

Fig. S1. Microscale thermophoresis analysis to determine the interaction affinities.

Fig. S2. Far-UV CD spectroscopy.

Table S1. Plasmids used in the present study.

Table S2. Primers used in the present study.

Insights into subunit interactions in the *Sulfolobus acidocaldarius* archaellum cytoplasmic complex

Ankan Banerjee, Tomasz Neiner, Patrick Tripp and Sonja-Verena Albers

DOI: 10.1111/febs.12534

Supplementary material to " Insights into Subunit Interactions in the *Sulfolobus acidocaldarius* Archaellum Core Complex " by Banerjee et al.

Supplemental tables

Table S1. Plasmids used in present study

Plasmids	Relevant characteristics	Source
pETDuet-1	Amp^r ,Cam^r , expression plasmid containing replicon ColE1 (pBR322) and two MCS (MCS1 and MCS2)	Novagen
pSA4	Derivative of pET15b containing the multiple cloning site and C-terminal hexa-His tag	(1)
pSVA 251	Amp^r ,Cam^r , pETDuet-1 carrying untagged *flaI* in MCS2 using restriction sites NdeI-XhoI	(2)
pSVA 263	Amp^r ,Cam^r , pETDuet-1 carrying N-terminal His-tagged *flaI* in MCS1 using restriction sites EcoRI-HindIII	(2)
pSVA 338	Amp^r ,Cam^r , pETDuet-1 carrying N-terminal His-tagged ΔNTD (1-226aa) *flaI* in MCS1 using restriction sites EcoRI-HindIII	
pSVA1911	Del 37 FlaX (*flaXc*) gene in pSA4 using restriction sites NcoI-BamHI	(3)
pSVA1919	Amp^r ,Cam^r , pETDuet-1 Carrying Nterminal His_6 tagged delta 37 FlaX(*flaXc*) in MCS1 using restriction sites EcoRI-HindIII	(3)
pSVA1944	Amp^r ,Cam^r , pETDuet-1 Carrying Nterminal StrepII tagged delta 37 FlaX (*flaXc*) in MCS2 using restriction sites NdeI-XhoI	This study
pSVA1973	Amp^r ,Cam^r , pETDuet-1 carrying N-terminal StrepII-tagged *flaI* in MCS2 using restriction sites NdeI-XhoI	This study
pSVA1978	Amp^r ,Cam^r , pETDuet-1 carrying N-terminal StrepII-tagged ΔCTD (227-513 aa) *flaI* in MCS2 using restriction sites NdeI-XhoI	This study
pSVA2100	Amp^r, Cam^r, pETDuet-1 carrying N-terminal His_6 tagged s*flaH* in MCS1 using restriction sites EcoRI-HindIII	This study
pSVA2108	Amp^r, Cam^r, pETDuet-1 carrying N-terminal Strep tagged s*flaH* in MCS2 using restriction sites NdeI-XhoI	This study

Table S2. Primers used in present study

Primers	Sequence and characteristics	Source
P714	5'-GGGGGCTCGAGTCAAATATTACTGACATATTTTATC-3'; Reverse primer for *flaI* containing XhoI restriction site (underlined)	(2)
P786	5'-GGGGGCTCGAGTTATGCGCGAGACAGGC-3'; Reverse primer for *sflaH* containing a XhoI restriction site (underlined)	(2)
P2153	5'-CCCCGAATTCGAGAGTGCAAACCCAGCAAAG-3'; Forward primer for *flaXc* containing a EcoRI restriction site (underlined)	(3)
P2154	5'-CCCCAAGCTTTTAACCATTCACACCACCACCCTC-3'; Reverse primer for *flaXc* containing a HindIII restriction site (underlined)	(3)
P2168	5'-GGGGGCATATGTGGTCCCATCCTCAA TTTGAAAAGATTATCAGCACCGGCAACG-3'; Forward primer for *sflaH* containing an NdeI restriction site (underlined) and strep II tag.	This study
P1269	5'- GGGGGCATATGTGGTCCCATCCTCAATTTGAAAAGAGA GTGCAAACCCAGCAAAG-3'; Forward primer for *flaXc* containing an NdeI restriction site (underlined) and StrepII tag.	This study
P2170	5'- GGGGGCTCGAGTTAACCATTCACACCACCAC-3'; Reverse primer for *flaXc* containing a XhoI restriction site (underlined).	This study
P3612	5'-CCCCCCGAATTCAATGATTATCAGCACCGGC-3'; Forward primer for *sflaH* containing a EcoRI restriction site (underlined)	This study
P3613	5'-GGGGGGAAGCTTTTATGCGCGAGACAGGC-3'; Reverse primer for *sflaH* containing a HindIII restriction site (underlined)	This study
P3700	5'- GGGGCATATGTGGTCCCATCCTCAATTTGAAAAGAGCTT CGTGGAAGACTATTTAAC-3'; Forward primer for *flaI* containing an NdeI restriction site (underlined) and strep II tag.	This study
P3705	5'-GGGGCTCGAGTCATCTACTAAATTTCCTGACAG-3'; Reverse primer for *flaI* CTD (227-513 aa) deletion, containing XhoI restriction site (underlined).	This study

Supplementary figure

Figure S1.

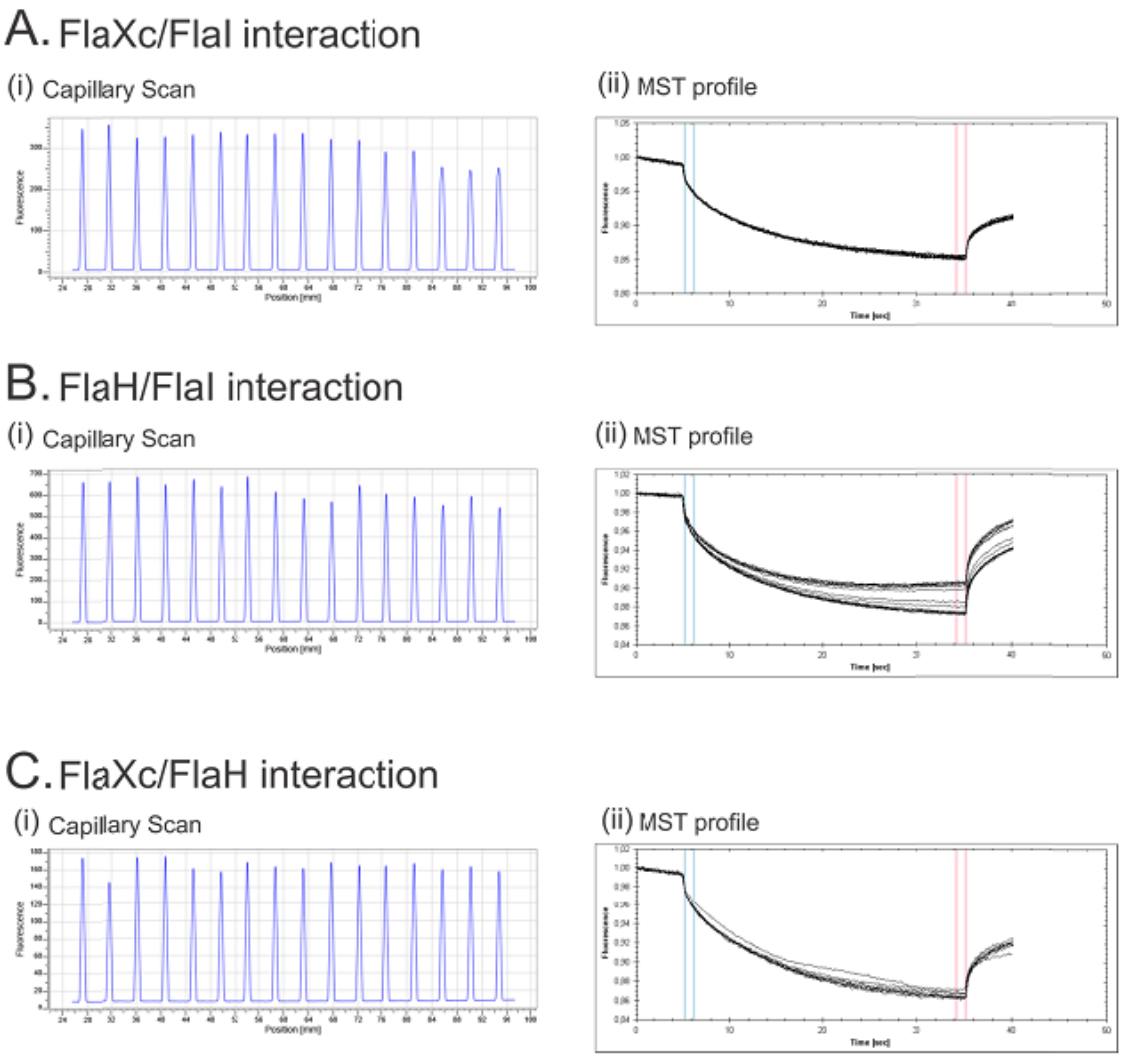


Fig. S1 Microscale thermophoresis analysis to determine the interaction affinities. A) 7.5nM of labeled monomeric FlaI was incubated with variable concentrations of FlaXc at steady state level. The capillary scan to check the initial fluorescence (RU) and area of interest (position mm) for LED exposure is shown in (i) and the MST profile of the experimental set is shown in (ii). The MST experiment was performed with 30s LED exposure and 5s LED off parameter. The blue bar area indicates the normalized fluorescence before infra red exposure (cold region) and the red bar area represents normalized fluorescence after infra red light exposure (hot region). The red region was predicted to have steady state conditions of the flourophore molecules and the diffusion of the fluorophore after interaction was

measured by normalizing the hot region fluorescence with cold region at specific position of the capillaries. The experimental raw data for FlaI/FlaH and FlaXc/FlaH was shown in **(B)** and **(C).**

Fig. S2

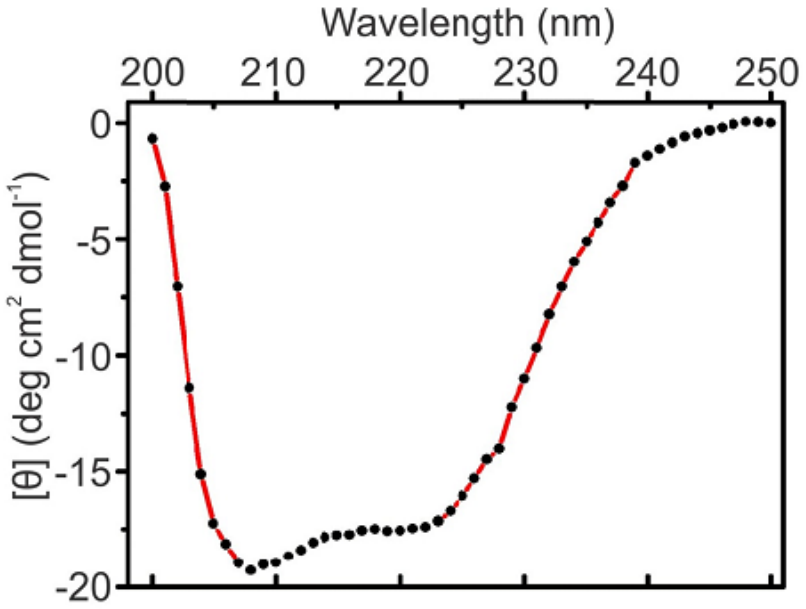


Fig. S2 Far-UV Circular Dichroism (CD) spectroscopy Far-UV CD spectrum (200nm-250nm) of 0.6 g/l refolded Strep-tagged FlaXc in 20mM Carbonate/Bicarbonate pH 10 buffer, 200mM NaCl. Experiments were conducted in 1mm path length cells. Far-UV spectrum of the tested protein is presented as the smoothed average of three accumulations.

Reference

1. **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**
2. **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**

3. **Banerjee, A., Ghosh, A., Mills, D. J., Kahnt, J., Vonck, J., and Albers, S. V. (2012) FlaX, a unique component of the crenarchaeal archaellum, forms oligomeric ring-shaped structures and interacts with the motor ATPase FlaI.** ***J Biol Chem***

3.4 Structural and functional characterization of FlaH: *an archaellum clock?*

Summary:

The archaeal motility organelle, the archaellum can be shortly described as an rotating type IV pillus (Shahapure et al., 2014a). It was mentioned in the previous sections that FlaI is the only functional ATPase in the archaellar system (3.2). It was also demonstrated that FlaI has a dual function, since it is involved in both, powering the T4P-like assembly process, as also propelling the rotation of the filament. However the *fla* operon encodes another protein defined *in silico* as an ATPase-like protein, containing a classical Walker A motif but an incomplete Walker B motif, namely FlaH (Ghosh & Albers, 2011a). The previous sections proved clearly that FlaH is an important component of the archaellar basal body, which not just interacts with the core components FlaX and FlaI, but also determinates the stability of FlaX in the cell of *S. acidocaldarius* (Lassak et al., 2012b, Banerjee et al., 2013). In this study we try to get deeper insight in the architecture of the archaellar basal body and expand our knowledge about the position of FlaH within this structure as also shed light on the putative function of this protein.

Using Föster resonance energy transfer experiments we have demonstrated that FlaH shows strong affinity towards ATP and has been even capable to bind any given nucleotide. *In vitro* mutants in both Walker motifs abolished ATP-binding completely and the corresponding mutation *in vivo* led to a non archaellated phenotype, linking ATP-binding of FlaH with proper archaellum assembly.

Moreover the crystal structure of FlaH was solved at 2.3 Å resolution, revealing the presence of an ATP molecule in the binding pocket, indicating that FlaH does not possess any ATPase activity, as crystallisation of active ATPase with a bound ATP molecule is a rather rare phenomenon.

In parallel we have confirmed the MST measured FlaH/FlaI interaction described in the previous section using pull-down analysis and small angle X-ray scattering. Purified protein interaction assays were further used to confirm the FlaH/FlaX interaction and in order to map the interacting part of FlaX, by testing the affinity of FlaH towards different truncated variants of FlaX. The experiment has revealed that the interaction with FlaH is mediated by the C-terminal part of FlaX, which is also crucial for FlaI interaction and FlaX ring closure. Additionally the FlaX/FlaH interaction was visualized using single particle analysis, which revealed the formation of dot-like substructures grouped into a second ring inside the FlaX ring, corresponding in size to monomeric FlaH.

Structural analysis of FlaH has shown explicit similarity to the CII domain of KaiC. KaiC is the core component in the cyanobacterial KaiABC circadian clock system and the CII domain is the functional part of this protein, regulating 24 hour oscillations in a phosphorylation dependent manner. Interestingly both FlaH and KaiC CII show the same spatial organization of the ATP-binding pocket, while they differ in terms of the orientation of the Walker B motif from all the others RecA family proteins. Based on the similarities to KaiC, we have tested FlaH for autophosphorylation activity, which was indeed demonstrated using radioactively labeled γ-P^{32}-ATP. Relying on structural similarities with the KaiC phosphorylation mechanism and a comparative alignment of FlaH sequences from different species, we selected potential phospho-acceptor groups, which we have subsequently mutated *in vivo*. All obtained mutants resulted in restricted or completely abolished motility, suggesting a phosphorylation dependent regulatory function of FlaH.

Contributions:

Tomasz Neiner purified FlaH, performed functional *in vitro* analysis including MANT-NTP binding assays and radioactive phosphorylation studies, constructed *in vivo* mutants and analyzed them phenotypically in terms of swimming behavior and archaella expression, performed Electron Microscopy and bioinformatical analysis. Ankan Banerjee purified proteins and performed interaction analyses. EdoardoD'Imprima did the Single Particle Analysis. Sophia Reindl performed SAXS analysis for FlaI-FlaH interaction and together with Andrew S. Arvai solved the crystal structure of FlaH. Abhrajyoti Ghosh constructed FlaH expression plasmid. The manuscript was written by Tomasz Neiner, Ankan Banerjee and Sophia Reindl (structural part). Sonja-Verena Albers supervised this study, designed experiments and revised the manuscript.

Structural and functional characterization of FlaH: *an archaellum clock?*

Tomasz Neiner[1#§], Ankan Banerjee[1#&], Edoardo D'Imprima[2$], Sophia Reindl[3£], Abhrajyoti Ghosh[1], Andrew S. Arvai[3], Deryck J. Mills[2], Janet Vonck[2], John A.Tainer[3] and Sonja-Verena Albers[1!]

[1]Molecular Biology of Archaea, Max Planck Institute for terrestrial Microbiology, Karl-von-Frisch- Strasse 10, 35043 Marburg, Germany

[2]Max Planck Institute of Biophysics, Department of Structural Biology, Max-von-Laue-Str. 3, D-60438 Frankfurt am Main, Germany

[3]Life Sciences Division, Department of Molecular Biology, Lawrence Berkeley National Laboratory, Berkeley, California 94720, USA.

Contribution:

# equal contribution of the authors.

§ who purified FlaH, performed functional assays, structure based analysis, Electron microscopy wrote manuscript and made figures.

& who purified proteins, performed interaction studies, wrote manuscript and prepared figures.

$ who performed single particle analysis, wrote manuscript and prepared figures.

£ who crystallized FlaH native protein, collected crystal diffraction data and SAXS data, analyzed crystal structure and SAXS data, and wrote paper and made figures

! Who wrote the manuscript, designed experiments and correspondence to S.V. Albers, e-mail: albers@mpi-marburg.mpg.de, tel.: +496421178426, fax: +496421178429

*Running title: FlaH an archaellum clock.

Keywords: archaellum/archaea/KaiC/circadian clock/archaellum core complex.

Abstract

To characterize the molecular architecture of the archaellum recent biochemical studies revealed that in *Sulfolobus acidocaldarius* FlaX forms a 30nm ring-like complex and interacts with FlaI and FlaH. FlaI is the only functional ATPase involved in archaellum assembly and subsequent rotation, whereas the role of FlaH remains unknown. FlaH is conserved in all archaellated archaeal species and with two other components FlaJ and FlaI were hypothesized to build the archaella motor complex. Here we present the first detailed study about the enigmatic archaella component FlaH. Combined structural, biochemical and genetic analysis of FlaH revealed that FlaH is an ATP binding protein, where both Walker A and Walker B are essential for ATP binding and archaella assembly. FlaH filled the FlaX rings, forming an inner ring *in vitro*. FlaH interacted with the C-terminus of FlaX, which also mediates interaction with FlaI and is essential for the FlaX ring formation. At the end we show that FlaH is homologue to KaiC CII domain, showing similar ATP-binding pocket architecture and autophosphorylation activity, suggesting a regulatory role of FlaH.

Introduction

Motility is evident in all three domains of life (Jarrell & McBride, 2008a, Margolis & Wilson, 1981). The archaeal motility structure, the archaellum, is a unique structure as it propels the cells forward in a rotary fashion like flagella, but its subunits share homologies with type IV pili (T4P) assembly system components. Archaella are composed of only 7-13 proteins and the archaella genes are usually found clustered in the genome. Deletion of most of these genes leads to non-archaellated strains, as demonstrated in *Methanococcales sp.*, *Sulfolobales sp.* and *Halobacteriaceae sp.* (Lassak et al., 2012b, Ken F. Jarrell, 2009). Although, archaella have been studied genetically, biochemical information about their subunits and the assembly process are still very scarce.

In *S. acidocaldarius*, seven different proteins build the archaellum complex and are essential for archaellation (Lassak et al., 2012b). Like in T4P, the structural component FlaB, the archaellin, exhibits a class III signal peptides that is processed by the membrane bound aspartic acid protease, PibD (Albers et al., 2003). Similar to T4P assembly it is so far unknown how processed archaellins or T4P are assembled into the filament. However, it is believed that the polytopic membrane protein of the archaella assembly system, FlaJ and the P-loop ATPase FlaI homologues to the bacterial T4P inner-membrane protein PilC and assembly/disassembly ATPases PilB/T, respectively, play a major role in this process (Ghosh & Albers, 2011b, Peabody *et al.*, 2003b, Ken F. Jarrell, 2009). FlaI from *S. acidocaldarius* formed ATP dependent hexamers that interacted specifically with tetraether lipids (Ghosh *et al.*, 2011b). Recently, the structure of FlaI was solved at 2Å resolution showing that the C terminal domain of FlaI holds the catalytic activity, while the N terminal domain was highly flexible. The first 29 amino acids show a high degree of flexibility and are predicted to play intrinsic role in archaellation. Intriguingly, a deletion of 29 amino acids *in vivo* exemplified the bifunctional role of FlaI as the deletion strain formed archaella, but swimming motility was abolished (Reindl et al., 2013). In parallel an extensive biochemical analysis of *S. acidocaldarius* FlaX suggested that the cytoplasmic domain of FlaX forms large ring like oligomeric structures of 30 nm diameter and that the last three predicted α-helices of FlaX were essential for ring closure (Banerjee et al., 2012b). This C-terminal domain of FlaX is also crucial for its interaction with the ATPase FlaI *in vitro* (Banerjee et al., 2012b). Further interactions with FlaH were demonstrated and formation of the FlaXHI ternary complex was confirmed *in vitro*, suggesting that these three proteins form the cytoplasmic motor complex of the archaellum. FlaH has a predicted Walker A motif (WA), but a non canonical Walker B motif (WB) and it was therefore predicted to be an ATP-binding protein, with a possible regulatory function towards FlaI. However, biochemical evidence regarding the role of FlaH have not been reported so far. In addition, it is not much known about the

archaella accessory proteins, FlaF, and FlaG. Both are predicted to be monotopic membrane protein with unknown function (Ghosh & Albers, 2011b, Banerjee et al., 2012b).

In this study we dissected and characterized FlaH structurally and biochemically. Using Förster resonance energy transfer (FRET), we showed that FlaH can bind any given nucleotide, but the affinity increased from diphosphate (ADP/GDP) to triphosphate (ATP/GTP) nucleotide species. A detailed structural analysis revealed that FlaH utilized both the Walker A and Walker B motif for ATP binding, however no ATP hydrolysis could be observed *in vitro*. Both *in vivo* and *in vitro* analysis of Walker A (K33A) and Walker B (D122N) mutants revealed that WA and WB motifs are important for ATP binding and also are essential for archaella assembly and motility. Using single particle analysis FlaH interaction with the FlaX rings was visualized. Moreover, domain mapping of FlaX/FlaH interaction revealed that the C-terminus of FlaX, which mediates FlaX-FlaI interaction and FlaX ring formation, is also essential for interactions with FlaH. Structural homology of FlaH to the KaiC CII domain implied autophosporylation activity, which could indeed be demonstrated *in vitro*. Moreover interference in the potential phosphorylation sites resulted in phenotypical defects. This study sheds light on the possible role of one of the most conserved archaellar components FlaH and increases our understanding of the assembly and interaction of archaella subunits.

Experimental Procedures

In silico analysis- FlaH protein sequences from different organisms were retrieved using Uniprot server (Consortium, 2012) and were subjected to ClustalW alignment analysis (Larkin *et al.*, 2007). Psipred and Jpred prediction servers (Buchan *et al.*, 2010, Cole *et al.*, 2008) were used to predict the secondary structure of all retrieved FlaH protein sequences. A comparative analysis of the secondary structure predictions were performed and presented in Fig.1. The DALI server (Holm & Rosenstrom, 2010) was used for identification of structural homologues of FlaH.

Strains and growth conditions- Sulfolobus acidocaldarius DSM639 was grown aerobically at 75^{o}C in Brock's basal medium (Brock et al., 1972) supplemented with 0.1% (w/v) tryptone (Roth) or NZ-Amine AS (Sigma) and 0.2% (w/v) dextrin with pH adjusted to 3.5 with sulphuric acid. The uracil auxotrophic strains MW001 (Wagner et al., 2012), MW156 (*S. acidocaldarius* Δ*aapF*) (Lassak et al., 2012b) and MW455 (*ΔflaH* in *S. acidocaldarius* Δ*aapF* background) (Lassak et al., 2012b) were grown in Brock medium supplemented with 10 μg ml^{-1} uracil. To prepare solid medium for *S. acidocaldarius* Brock medium was solidified with a final concentration of 0.6% (w/v) Gelrite (Sigma), 10 mM $MgCl_2$ and 3 mM $CaCl_2$ and pH was adjusted accordingly. The NEB ER1821^{kan} strain was used to methylate plasmids and *E. coli* BL21 (DE3) containing RIL plasmid (*E. coli* codon plus) was used for over expression of proteins.

Molecular cloning and mutagenesis- To construct recombinant FlaH (C-terminal His_6 tagged, N-terminal His_6 or N-terminal StrepII tagged) a pET based expression plasmid pETDuet1 was used. Full length *flaH* gene was amplified from pSVA284 (carrying a codon optimized *flaH* gene, *sflaH*) using primer pairs 771/772, 3612/3613 or 786/2168, respectively, and cloned as *NdeI/XhoI* or *EcoRI/HindIII* digested fragments into pETDuet1, yielding plasmids pSVA293, pSVA2100 and pSVA2108, respectively. FlaH point mutations were constructed using site directed mutagenesis via round PCR on the pSVA2100 vector backbone. FlaH^{Lys33}and FlaH^{Asp122}within the predicted Walker A and Walker B motifs were mutated to alanine (A) and asparagines (N) respectively, using primer pairs 3632/3633 and 3634/3635 yielding plasmids pSVA2130 (FlaH^{K33A}) and pSVA2131 (FlaH^{D122N}). Primers and plasmids used in this study are summarized in Table S2-S4.

Generation of FlaH in vivo point mutation- in vivo point mutations in *S.acidocaldarius* were generated following the protocol described by Lassak et. al (Lassak et al., 2012b) with a modification. The *flaH* gene with about 700bp

up-stream and down-stream regions was amplified by PCR from genomic DNA of *S. acidocaldarius* DSM639 with primers 3614/3617 and cloned as *ApaI/PstI* digested fragment into pSVA406 vector, yielding pSVA2126. Site directed mutagenesis of *flaH* was performed via round PCR on pSVA2126 plasmid using primer pairs 3615/3616 and 3627/3628 to obtain pSVA2113 ($FlaH^{K33A}$) and pSVA2124 ($FlaH^{D122N}$), respectively. The constructs were methylated using $ER1821^{kan}$ and transformed into MW455 (*S. acidocaldarius* Δ*aapF/ΔflaH*) to express the protein *in trans*. Strains carrying point mutations in potential phosphorylation sites were generated analogously to the Walker mutants described above. All the created vectors and used primer pairs are listed in the Supplementary information in Tables S2-S4.

Protein expression and purification- pET based overexpression constructs were transformed into *E. coli* codon plus cells and grown as preculture overnight at 37°C in LB medium containing ampicilin (50 μg/ml) and chloramphenicol (34 μg/ml). 2 liter of fresh LB-ampicilin-chloramphenicol medium was then inoculated with 2 ml preculture and grown at 37°C to an optical density (OD_{600}) of 0.6-0.8. Subsequently, the cultures were chilled on ice for 30 minutes and induced with 0.5 mM of isopropyl β-D-thiogalactopyranoside (IPTG). The induction was carried out at 17°C overnight (around 16 hours). The cells were collected by centrifugation and frozen in liquid nitrogen and stored at −80°C.

The frozen cells were slowly thawed on ice, resuspended in 50ml lysis buffer (50 mM Hepes-NaOH, pH 7.3, 150 mM NaCl) containing 1X complete EDTA-free protease inhibitor cocktail (Roche) and lysed by sonication with a Sonoplus HD3100 sonicator (Bandelin Sonorex Biotechnique, Germany) using the probe HD3100. Cell debris was removed by a low spin centrifugation (30 min, 4600 x g, 4°C) followed by a fast spin centrifugation with 20,000 x g at 4°C for 30 min to get rid of all the insoluble particles. The cell free lysate was then incubated with DNase I for one hour on ice followed by a 20% $(NH_4)_2SO_4$ (w/v) precipitation. The white fluffy pellet containing FlaH was collected by centrifugation with 4500 x g, at 4°C for 1 hour and resuspended into 10 ml lysis buffer and dialyzed overnight against 1 liter 50 mM Hepes pH 7.3, 300 mM NaCl buffer. Ni-NTA affinity or Streptactin affinity based purification was performed to purify FlaH variants, where FlaH was eluted with 50 mM MES, 500 mM NaCl, pH 6.0 containing 500 mM imidazol or 2.5 mM d-desthiobiotin. Further purification of FlaH was achieved with incubation at 50°C for 20 min. To remove the contaminating proteins the sample was centrifuged with maximum speed for 10-15mins in a table top centrifuge and the supernatant was dialyzed overnight against 1 liter 50 mM MES, 500 mM NaCl, pH 6.0 buffer. The protein aliquots were frozen in liquid nitrogen and stored at -80°C for further use. FlaI was purified following the protocol described by Ghosh et. al (Ghosh et al., 2011a). FlaXc and its variants were purified using the protocol described by Banerjee et. al (Banerjee et al., 2012b).

Crystallization and structure determination of FlaH- For crystallization experiments 2mM ATP and 5mM $MgCl_2$ was added to the eluate from the Ni-NTA affinity column and concentrated using Amicon centrifugal filter devices. Concentrated FlaH+ATP+$MgCl_2$ was applied to a Superose 6 size exclusion chromatography column in 20mM tris pH 7.0, 500mM NaCl and 5% Glycerol. FlaH in complex with ATP and $MgCl_2$ eluted as monomer from the column. Peak fractions were pooled, concentrated to 6mg/ml and crystallized in 100mM HEPES pH 7.0, 32.5% PEG 3350 and 150mM NaSCN. Crystals were flash-frozen in liquid nitrogen using 1,4-butanediol as cryo-protectant and diffraction data (Otwinowski & Minor, 1997) were collected at the ALS SIBYLS beamline (Berkeley, CA) and processed with HKL2000 (Table 1). The FlaH structure was solved using molecular replacement with the program Phaser (McCoy *et al.*, 2007) using the RecA superfamily ATPase PH0284 from *Pyrococcus horikoshii* (pdb code 2DR3) as search model. The structure was refined using PHENIX (Adams *et al.*, 2010) and manually fit using COOT (Emsley & Cowtan, 2004). One FlaH molecule was present in the asymmetric unit with clear density for an ATP molecule and a Mg^{2+} ion. Refinement statistics are shown in Table S1.

Small Angle X-ray Scattering (SAXS) - SAXS was performed to analyze a possible interaction of FlaH with FlaI in solution. SAXS data of proteins were collected at the SIBYLS beamline12.3.1 at ALS, LBNL (Hura *et al.*, 2009, Classen *et al.*, 2010). The wavelength λ of the incident X-ray beam was 1.11 Å and the sample-to- detector distance was 1.5 meter. The scattering vector *q* range is from 0.01 $Å^{-1}$ to 0.32 $Å^{-1}$. The *q* is defined as $(4\pi \sin\theta)/\lambda$, where 2θ is the scattering angle and λ is the wavelength. No radiation damage or aggregation was observed from the scattering data. All data were collected at room temperature (18–21 °C). The collected data were processed using PRIMUS program (Konarev *et al.*, 2003). The Guinier plot was used to calculate R_g (radius of gyration).

MANT-ATP binding assay- MANT-ATP [2'-/3'-O-(N'-methylanthraniloyl) adenosine-5'-O-triphosphate] binding was monitored by fluorescence resonance energy transfer (FRET) using a temperature controlled ISS PC1 spectrofluorometer. The slit-widths for the excitation and the emission beam were set to 1 nm. The excitation wavelength was set to 285 nm and the emission spectrum was recorded from 400 to 500 nm. The binding affinities were measured by titrating 1 μM protein with increasing MANT-ATP concentration (0-3 μM) at 25°C in binding buffer (20 mM MES, 500 mM Nacl, 5 mM $MnCl_2$, pH 6.2). Fluorescence spectra were corrected for the fluorescence of unbound MANT-ATP.

In vitro binary complex formation (pull down) - *in vitro* protein-protein complex formation was performed as described in (Banerjee et al., 2012a). To form the binary complex, pure proteins either with His_6 or strep II or untagged were incubated together on 50mM Tris-HCl pH8, 150mM NaCl buffer and pulled down using specific affinity beads (Ni-NTA or Streptactin). Complex formation was visualized on coomassie stained SDS-PAGE and confirmed by either α-His or α-strep or α-FlaI antibodies. In parallel a control experiment was performed with only His_6 tagged FlaX variants using the same protocol as mentioned above.

Electron microscopy and motility assay- Transmission electron microscopy of *S. acidocaldarius* cells, with image analysis and the swimming assay on semi solid plates were performed following the protocols described previously by Lassak et. al (Lassak et al., 2012b).

Cryo EM / Single particle analysis- Purified FlaX and FlaH were mixed in a molary equal ratio. For EM analysis FlaX/FlaH complexes were diluted to a final concentration of 0.05 mg/ml and negatively stained with 1% (w/v) uranyl acetate. Electron micrographs were collected using a Philips CM120 at 120 kV, at a magnification of 44,000 x on a Gatan 2kx2k CCD camera.
For cryo-EM, FlaX/FlaH samples were directly applied to glow-discharged Quantifoil grids (Quantifoil MicroTools) and vitrified by injection into liquid ethane using a Vitrobot plunge-freezing device (FEI). Images were recorded using a back-thinned FEI Falcon II direct electron detector. Images were recorded manually at a nominal magnification of 39,000x yielding a pixel size at the specimen of 2.74 Å. at a defocus of 2-6 μm. Particles were selected using the boxer module from EMAN (Ludtke, S. J, 1999). The CTF was determined using EMAN2 (Tang, G.,2007). A data set of 16,000 particles was processed with Imagic V (van Heel, M.,1996). The images were subjected to reference-free alignment and multivariate statistical analysis. The images were classified and images assigned to the same class were averaged.

In vitro phosphorylation assay- 0.5 μM purified proteins ware assayed in 20 μl of 25 mM MES pH 6.5, 150 mM NaCl, 5 mM $MgCl_2$ buffer containing 32 nM [g-32P]-ATP (222 TBq $mmol^{-1}$, Hartmann Analytics) and 0.8 mM ATP. The reaction was run at 55°C then stopped by adding 5x SDS loading dye. The samples were separated on SDS-PAGE and exposed to a Storage Phosphor Screen (Molecular Dynamics) for at least 2 hours. Images were scanned with Storm 840 scanner.

Results

FlaH contains an ATP binding motif- Two separate transcriptional units encoding

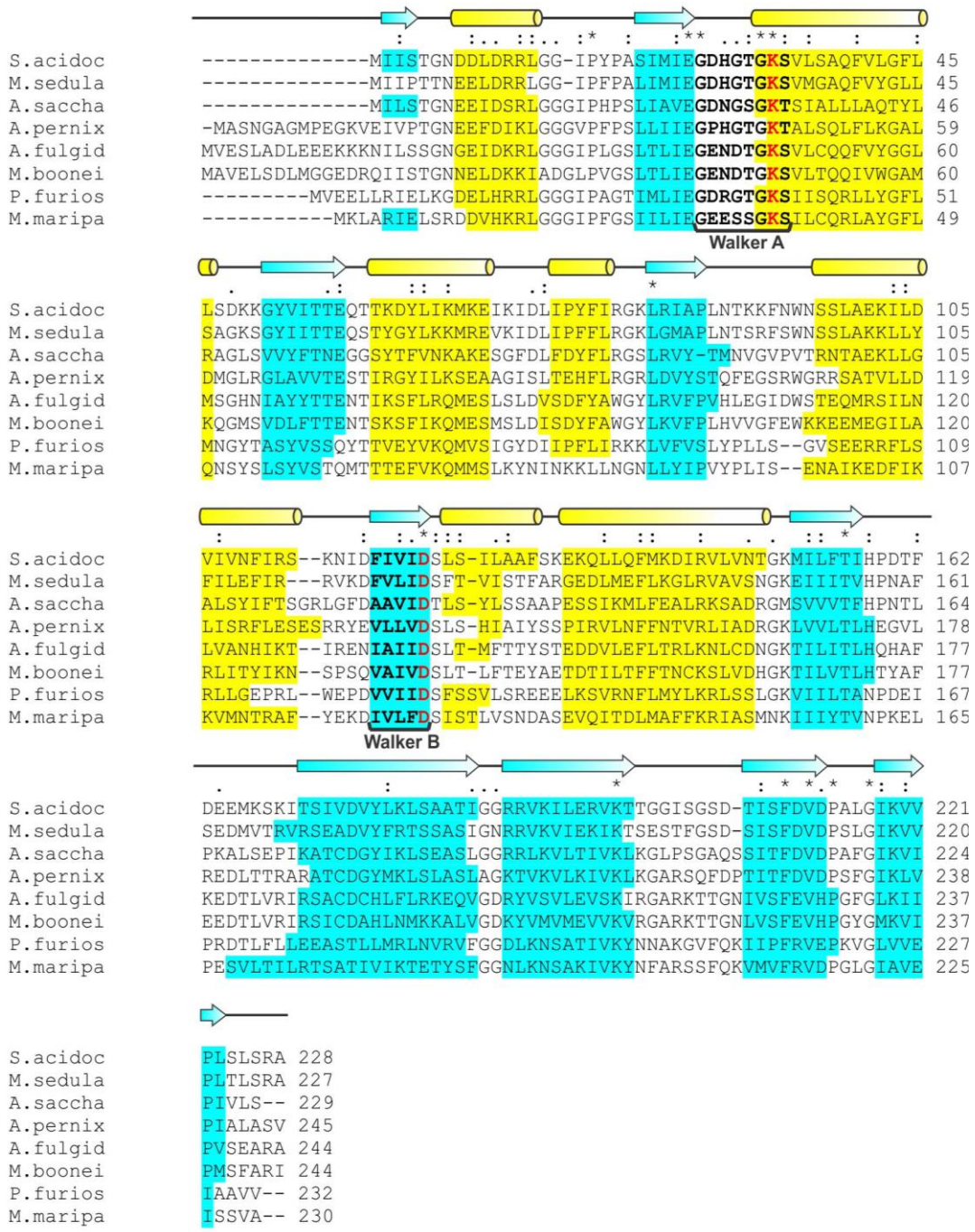


Fig.1. Multiple sequence alignment and secondary structures prediction of FlaH homologues from different archaeal species. *Sulfolobus* acidocaldarius FlaH homologues were identified with BLAST analysis and aligned using ClustalW. A secondary structure prediction for all aligned FlaH homologues was performed using PSIPRED. The received structural predictions were consistent for all the homologues: predicted α-helices are highlighted in yellow, β-shits in cyan. The identified Walker motif are marked with a bold font and the highly conserved lysine within Walker A and the aspartic acid within Walker B box are highlight in red. (Strains: *Metallosphaera sedula*, *Acidilobus saccharovorans*, *Aeropyrum pernix*, *Archaeoglobus fulgidus*, *Methanoregula boonei*, *Pyrococcus furiosus*, *Methanococcus maripaludis*)

archaella components have been described in *S. acidocaldarius*. The archaellin encoding gene *flaB* is transcribed separately, whereas the second transcriptional unit is comprised of the remaining *flaXGFHIJ* genes (Lassak et al., 2012b). In the archaellum assembly cluster FlaH was reported as an essential component for archaellation as well as motility in *S. acidocaldarius*, *M. maripaludis* and *M. voltae* (Lassak et al., 2012b, Ken F. Jarrell, 2009, Thomas *et al.*, 2001b).

FlaH, is highly conserved in all archaellated archaea (Jarrell & Albers, 2012). A ClustalW alignment of FlaH sequences from Euryarchaeota and Crenarchaeota showed a high degree of homology (Fig.1). All analyzed FlaH sequences share a conserved Walker A (WA) motif (Fig.1). The WA motif has the sequence GXXXXGK (T/S) and is also known as p-loop or phosphate binding loop (Hanson & Whiteheart, 2005). The WA box was reported to be essential for ATP binding (Ghosh et al., 2011a). On the other hand the Walker B (WB) box in FlaH exhibits non-canonical amino acid residues. It was suggested that hhhhDE is the signature sequence of an active WB box, where the aspartate functions in chelating the divalent cation and the glutamate mediates ATP hydrolysis (Hanson & Whiteheart, 2005). In FlaH, the aspartate moiety is conserved, but the glutamate is replaced by serine or threonine (Fig.1), which suggested that FlaH can bind ATP, but cannot hydrolyze it. A secondary structure prediction analysis revealed that the WB motif is present in a small β-sheet structure. FlaH from different archaeal species shares a similar secondary structure, as revealed from the secondary structure prediction analysis (Fig.1).

ATP binding to FlaH is essential for proper archaellation phenotype- ATP binding assays were performed with FlaH using the fluorescent ATP analogue MANT-ATP. 0-3μM MANT-ATP was titrated to 1μM FlaH and a FRET scan was monitored (Fig.2). As FlaH was reaching the ATP saturation point at ATP concentrations below the FlaH concentration, it was impossible to calculate the precise affinity (K_D). FlaH did not only bind MANT-ATP, but also MANT-GTP and MANT-ADP were strongly bound by FlaH (Fig.S1). To test whether FlaH posses additionally an ATPase activity a Malchite Green assay was performed, however no active ATP hydrolysis could been observed.

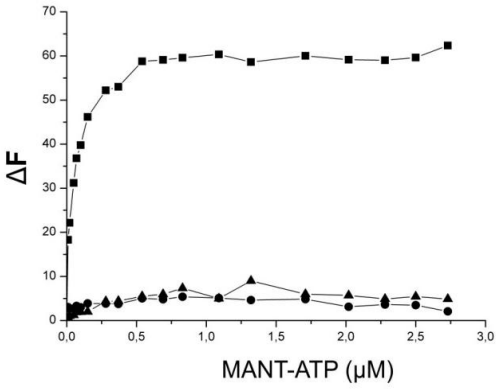


Fig.2. FlaH binds ATP and it requires both WA and WB boxes for binding. FRET analysis of the MANT-ATP binding to FlaH and FlaH Walker motif mutants. Change in fluorescence (ΔF) upon MANT-ATP binding is shown as a function of the MANT-ATP concentration. Closed square, $FlaH^{WT}$; closed circle, $FlaH^{K33A}$; closed triangle, $FlaH^{D122N}$. Fluorescence spectra were corrected for the fluorescence of unbound MANT-ATP.

To determine the role of the WA and WB motifs in nucleotide binding in FlaH, the lysine residue in the WA ($FlaH^{K33}$) and the aspartic acid in the WB ($FlaH^{D122}$) like motif were mutated to alanine ($FlaH^{K33A}$) and asparagine ($FlaH^{D122N}$), respectively. Both $FlaH^{WA}$ and $FlaH^{WB}$ mutants were expressed in *E. coli* and purified with a similar yield and purity as the $FlaH^{WT}$ (Fig. S2). FlaH variants were then tested for MANT-ATP binding. Surprisingly, the binding of ATP was not only abolished in the WA mutant, but also in the WB mutant (Fig2). These results suggested that in FlaH also the WB motif is involved in the coordination of ATP during binding of the nucleotide to the protein. To support these data *in vivo*, the *flaH* gene was mutated in both the WA ($FlaH^{K33A}$) and the WB ($FlaH^{D122N}$) motifs in the *S. acidocaldarius* genome. The mutant strains were analyzed in terms of their ability to swim on semi solid plates and whether archaella were assembled on their cell surface. The *S. acidocaldarius* FlaH mutant strains $K_{33}A$, $D_{122}N$, the motile control strain Δ*aapF* and the non-motile Δ*aapF*/Δ*flaH* strain were dropped on semi

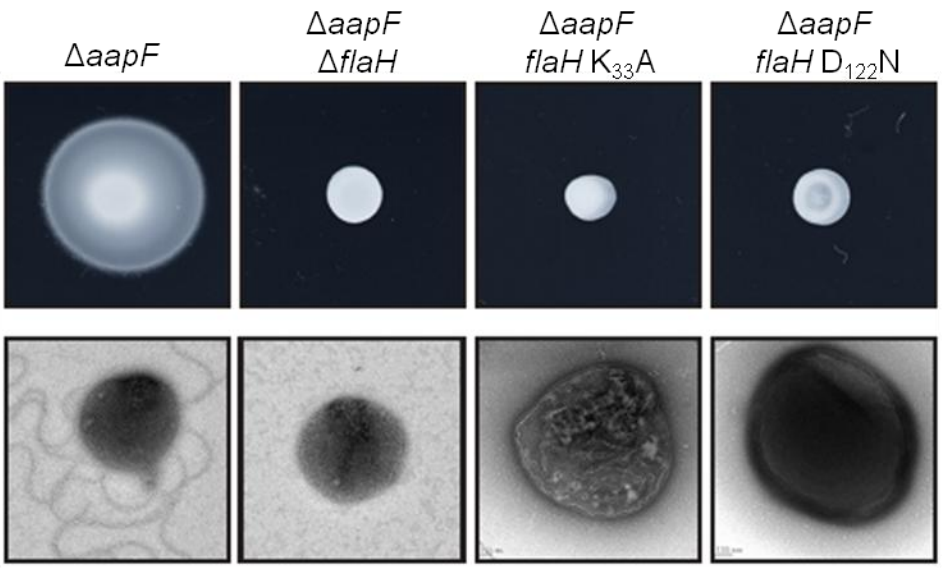


Fig.3. ATP binding to FlaH is essential for archaellation. Motility assay on semi solid swarm plates (upper row) and electron micrographs of negatively stained *S. acidocaldarius* cells (lower row). From left to right: As positive control the strong archaellated, hypermotile Δ*aapF* strain was used. Δ*aapF*/Δ*flaH* with abolished archaella expression was used as negative control. Both, the FlaH Walker A (K33A) and Walker B (D122N) mutations, resulted in suppressed archaella assembly and motility.

solid plates and incubated for seven days (Fig3, upper panel). In parallel, these strains were analyzed for archaellation by electron microscopy (Fig3, lower panel). Both the $FlaH^{K33A}$ and $FlaH^{D122N}$ mutant strains were non motile on semi solid plates and EM analysis revealed that these strains did not assemble archaella (Fig3). The results suggested that any change in the Walker motifs of FlaH leads to abolishment of archaellum assembly and therefore motility.

Crystal structure of FlaH bound to ATP and Mg^{2+}- In the absence of nucleotides FlaH was not stable at concentrations needed for crystallization, the protein started to precipitate at concentrations above approx. 1mg/ml. In the presence of ATP and Mg^{2+}, however, FlaH was successfully crystallized and the structure was solved by molecular replacement using the structure of the RecA superfamily ATPase PH0284 from *Pyrococcus horikoshii* (pdb code 2DR3) as search model. The final model was refined to 2.3 Å resolution with free-R and working-R factors of 24.8% and 19.3%, respectively and good stereochemistry (Table S1). The crystal asymmetric unit contains one molecule of FlaH, one ATP and one Mg-ion. The final model encompasses residues 1-226, missing the last 2 residues and the C-terminal His-tag. (Fig. 4). FlaH contains the typical RecA superfamily ATPase domain of a large twisted central β-sheet, sandwiched by α-helices on both sides. The central β-sheet consists of nine strands and no additional domains besides the classical RecA fold. Crystal packing does not reveal any biologically relevant quaternary structure.

Structural comparison of FlaH to other RecA superfamily proteins- Using the DALI search (Holm & Rosenstrom, 2010) for structural homologies confirmed FlaH as a close homolog to other members of the RadA/Rad51/RecA family (Fig 4.). The first hit is the RecA domain of the circadian clock protein kinase KaiC (3S1A.pdb). The next closest hits were the archaeal RadA Sso2452 paralog with antirecombinase activity (2W0M.pdb), and the archaeal RecA superfamily ATPase PH0284 (2DR3.pdb) that was used for molecular replacement as input model, the archaeal Rad51 protein (1PZN.pdb) followed by other members of the RecA superfamily of ATPases (Fig. 4C). An additional helical domain that is present in some of those ATPases (shown in light green in Fig. 4C) is missing in the FlaH structure, which consists of the plain RecA domain. The

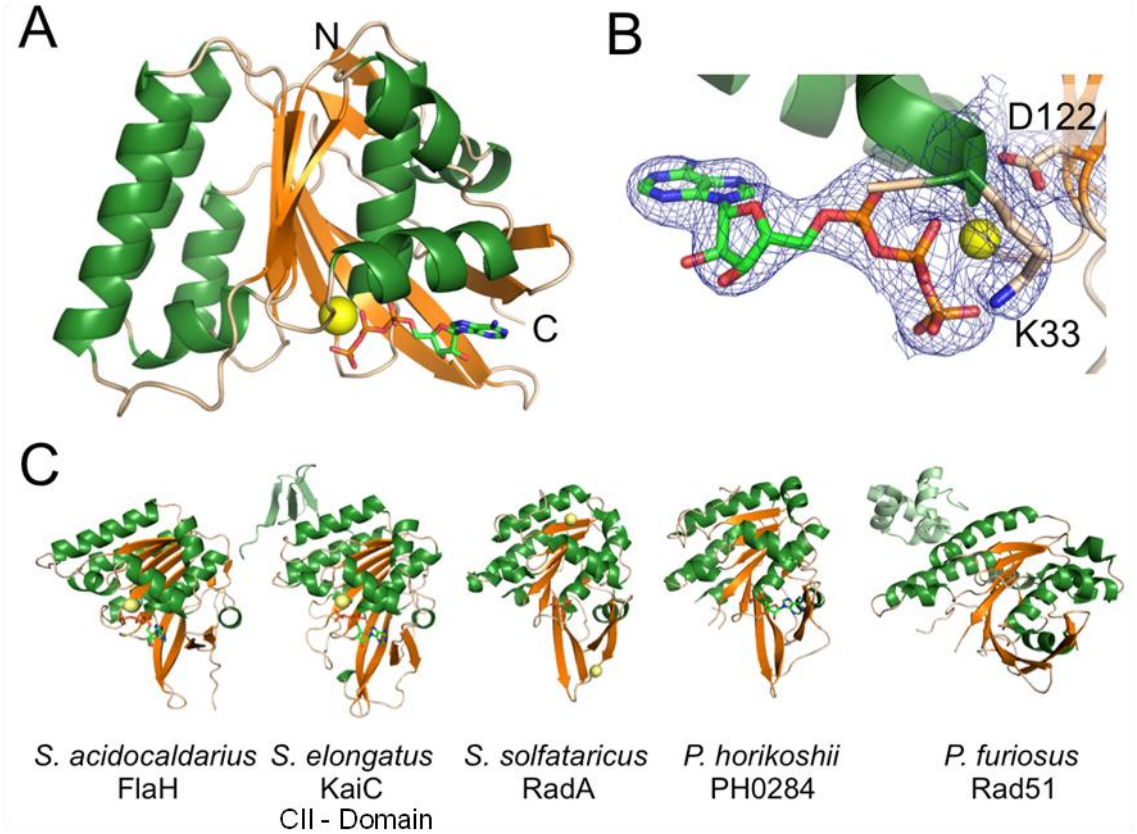


Fig.4. Crystal structure of FlaH with ATP. A) The *S. acidocaldarius* FlaH monomer is shown in ribbon representation. ATP bound to the active site is shown as sticks and the Mg^{2+} ion as yellow sphere. N- and C-termini are indicated. **B)** Active site with electron density (2|Fo|-|Fc| map at 1σ) for the bound ATP and Mg^{2+}-ion. The Walker A and Walker B residues K33 and D122. **C)** FlaH compared to other RecA superfamily ATPases. The first hits from a DALI (1) search for structural homologs are shown. Additional subunits besides the RecA domain are shown in light green. Nucleotides or phosphate bound to the active sites are shown in sticks, ions as spheres.

high concentrations of FlaH necessary for crystallization experiments could only be achieved in the presence of ATP and $MgCl_2$. ATPase activity could not be detected for purified FlaH, therefore it is not surprising that clear electron density for a bound ATP molecule was found in the structure (Fig. 4B).

In vitro interaction of FlaH with FlaI and FlaX- It was shown recently that FlaX together with FlaH and FlaI form the cytoplasmic motor complex of the crenarchaeal archaellum, where FlaX was predicted to act as a molecular scaffold (Banerjee et al., 2013). The formation of a ternary FlaXHI complex was confirmed *in vitro* and the three proteins interacted with very high affinities in the nanomolar range (Banerjee et al., 2013). Here we additionally examined the affinity of FlaH towards the other archaella core components FlaI and FlaX using pull-down assays with purified proteins *in vitro*. Unfortunately, during the course of experimental procedure we observed that FlaH invariably bound unspecifically to any affinity matrix (Fig.S3), therefore we decided to use the affinity tag of FlaH for purified protein pull down analysis. N-terminal His_6 tagged FlaH was mixed with strep tagged FlaI and incubated with Ni-NTA beads to bind FlaH. Both Coommassie stained SDS-PAGE and immunoblot analysis showed that FlaI eluted together with FlaH (Fig.5A). As FlaI did not bind to the Ni-NTA beads, FlaI and FlaH specifically interacted. An analogues experiment was performed in order to exam the FlaX/FlaH interaction, wherein a truncated mutant of FlaX, FlaXc, was used. FlaXc lacks the first 37 amino acid residues and therefore its transmembrane domain (Fig. 4C) (Banerjee et al., 2012a). His_6-FlaXc and Strep tagged FlaH were mixed and incubated

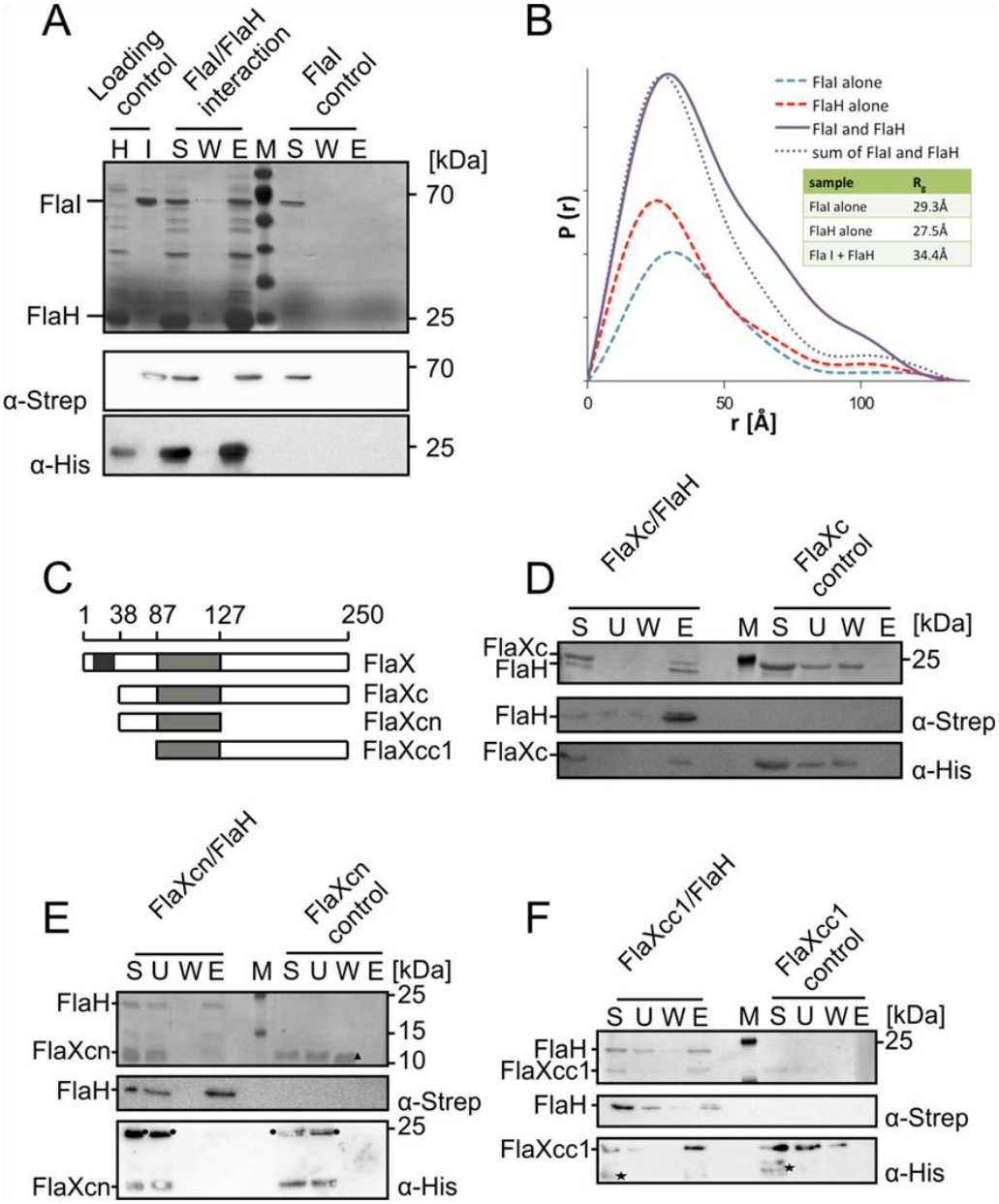


Fig.5. Archaellum interactome analysis. A) *In vitro* FlaI and FlaH interaction was tested using pure StrepII tagged FlaI and partially pure His_6 tagged FlaH. Proteins were mixed and incubated together at 4°C under constant stirring condition and the binary complex was pulled down using Ni-NTA affinity beads. A control experiment was performed with StrepII tagged FlaI. H, partially pure FlaH and I, pure FlaI shown as loading control. The result was visualized using coomassie stained SDS-PAGE and immunoblots using α-Strep and α-His antibodies. **B)** SAXS analysis of FlaI and FlaH interaction: The P(r) distribution is shown for FlaI (blue) and FlaH (red) alone and the sum of both curves (dashed, purple) is compared to the data of a 1:1 mixture of both proteins (solid, purple). Inset table: Radii of gyration derived from the Guinier plot**. C)** A schematic of FlaX truncates used in this study, where the dark grey box highlights the membrane domain of FlaX and the light grey box is the coiled coil region of FlaX. **D)** FlaH interacts with FlaXc. A binary complex formation experiment was performed using StrepII tagged FlaH and His_6 tagged FlaXc. Proteins were mixed and pulled down using streptactin beads and 5 mM d-desthiobiotin. Result visualized using coomassie stained SDS-PAGE and immunoblots using α-Strep and α-His antibodies. **E,F)** Domain mapping of FlaH-FlaXc interaction. Two separate *in vitro* protein complex formation experiments using either His_6 tagged FlaXcn or FlaXcc1 and StrepII tagged FlaH were performed and results visualized accordingly. Control experiments using FlaXc, FlaXcn and FlaXcc1 were performed following the same protocol of *in vitro* binary complex formation on Streptactin beads.

S, starting material; U, unbound fraction; W, wash fraction and E, elution of the binary complex; M, PAGE ruler Marker.

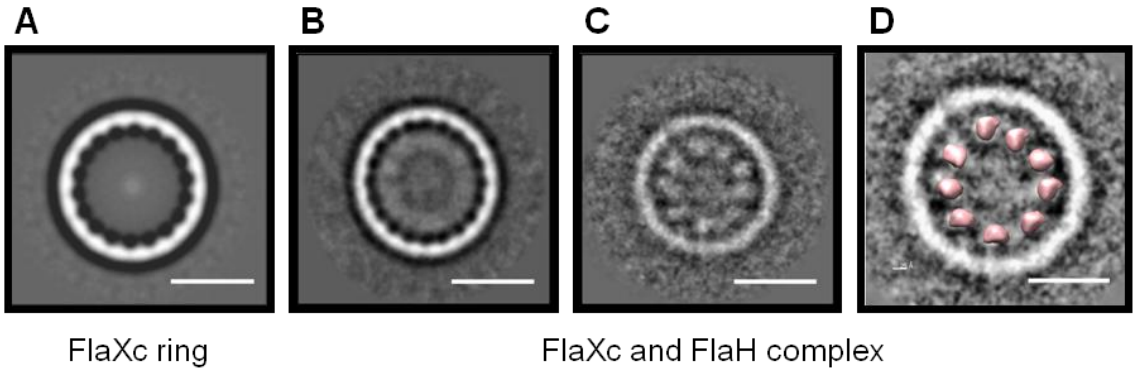


Fig.6. FlaH is forming stable substructures inside the FlaX ring *in vitro* *(Banerjee et al., 2012b).* **(A)** Single particle analysis of FlaX rings; **(B)** Analysis of FlaX and FlaH complexes aligned against the ring rim; **(C)** particle belonging to class average on B aligned against the inner part of the ring revealed densities that match the x-rays map of FlaH filtered at 60A resolution. These densities were smeared out in B resulting in an inner thicker ring; **(D)** The densities inside the FlaX rings correspond to the size of the monomer structure of FlaH. Bar = 20 nm.

with Streptacin beads. FlaXc was only present in the elution fraction when it was incubated in the presence of FlaH, but did not bind to the Streptactin beads itself. Therefore it could be confirmed that FlaXc and FlaH interact *in vitro* (Fig.5D).

The FlaH/FlaI interaction was tested additionally using Small angle X-ray scattering. We measured the scattering curves for both FlaH and FlaI separately and compared them with the data for the mixture of the two proteins. The best way to visualize any interaction is the pair distribution function (P(r)) that shows a distribution of all distances present in the sample. Without any interaction, the sum of the P(r) function of FlaH and FlaI would look identical to the data from the mixture. Fig. 5B clearly shows an occurrence of larger distances for the mixture in comparison to the sum of the single curves. A second measure is the radius of gyration R_g that can be extracted from the low angle scattering data: for the mixture the R_g is 34.4Å whereas the R_g for FlaH and FlaI alone is 27.5Å and 29.3Å, respectively. Although an *ab initio* structure determination of the complex is not possible, because the sample consists of a mixture of the protein complex with separate molecules of FlaH and FlaI, the presence of an interaction was clear.

Domain mapping of FlaX/FlaH interaction- FlaX has a triple helix coiled coil region, separating the C-terminus and N-terminus (Banerjee et al., 2012a). Coiled coil domains are often involved in protein-protein interactions. To identify the domain of FlaX interacting with FlaH, we have used FlaXcn (a deletion of the C-terminal domain) and FlaXcc1 (a deletion of N-terminal domain) (Fig. 5C) (Banerjee et al., 2012a). In a pull down assay Strep tagged FlaH was mixed with His tagged FlaXcn or FlaXcc1 and then bound to Streptactin beads. The formation of the complex was confirmed using Coomassie stained SDS-PAGE, α-His and α-Strep immunoblot analysis (Fig. 5E/F). The results indicated that the N-terminal half (38-86 aa) and the coiled coil region (87-127 aa) of FlaX do not interact with FlaH, similarly to FlaX's interaction with FlaI (Banerjee et al., 2012a).

Single particle analysis of the FlaX/FlaH complex- It has been demonstrated that FlaX forms highly stable ring-like oligomeric structures (Fig6A) and shows strong affinity towards FlaH (Banerjee et al., 2012a, Banerjee et al., 2013). In order to visualize the FlaX/FlaH interactions we mixed the purified proteins *in vitro* in an equal molar ratio and analyzed them with electron microscopy. Almost every FlaX ring was filled with an inner ring (Fig6B/C). Further Cryo-EM analysis indicated that in presence of FlaH the dominating symmetry was 18-19 fold in contrast to the empty FlaX rings which were found to have symmetries ranging between

16-23 fold (Banerjee et al, 2013). The analysis of a data set composed of more than 16 000 images and the absence of any stain allowed to improve the resolution to such an extent, that the finer details of the formation of substructure within the FlaX rings could be visualized (Fig.6C). After imposing the crystal structure of FlaH low pass filtered at 60 Å on the obtained FlaX-FlaH ring model the visualized substructures within the ring corresponded to monomeric FlaH in size (Fig. 16D). The FlaX-FlaH complex was a large, but unfortunately a relative flat structure that tended to remain attached to the liquid/air interface on the sample grids, so that only rarely side views were found. Any attempt of increasing the number of side views of the rings like addition of detergents and tilting the specimen failed, thus a further 3D structural model of the FlaX/FlaH rings could not be obtained with this method.

FlaH shows auto-phosphorylation activity- The nearest homologue of FlaH is the circadian clock protein of cyanobacteria, KaiC. KaiC consists of two domains, CI and CII, both show structural homology towards each other, differ however significantly in performed functions (Hayashi et al., 2004). The CI domain was reported to have ATPase activity and mediate the KaiC hexamerisation, whereas the CII domain is the regulatory part of KaiC (Hayashi et al., 2003, Hayashi et al., 2004). We analyzed the structural homologies between FlaH and KaiC in order to elucidate functional aspects of FlaH. Structural alignment revealed that FlaH resembles much more the CII domain of KaiC, than the CI part. The superimposed structures of FlaH and KaiCII represent a unique spatio-structural similarity (Fig.8A). Interestingly the ATP binding pockets of both structures show a well preserved sterical overlay. The conserved Lys 33 of Walker A and Asp 122 of Walker B show in FlaH and KaiCII a very similar spatial location (Fig. 8B). Whereas the structural alignment with other RecA family proteins revealed a remarkably conserved sterical placement of the WA motif, the important Asp residue within the WB motif of FlaH pointed into a different direction compared to the other RecA proteins (Fig. 7).

It is known that the circadian rhythms regulation is strictly coupled with the autophosphorylation activity of KaiC (Rust et al., 2007). The phosphorylated residues in KaiC were proven to be S431 and T432 located in the CII domain (Nakajima et al., 2005). Interestingly, these conserved residues are also present in FlaH of *S. acidocaldarius*. The T171 and S172 of FlaH exhibits an inverted order, relatively to S431T432, show however a close spatial location (Fig. 9A). In order to investigate whether the structural

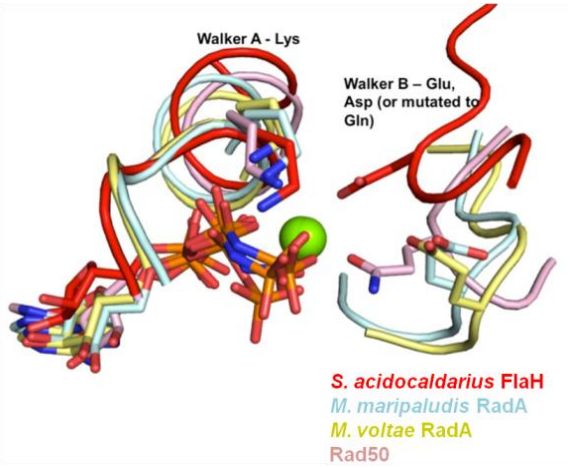


Fig.7. **ATP-binding pocket analysis**. Structural alignment of ATP-binding domains of different RecA family proteins: *M. maripaludis* RadA, cyan; *M. voltae* RadA, yellow and Rad50, pink with *S. acidocaldarius* FlaH, red. The lysine (K) within the Walker A motif is conserved in all analyzed structures. However the aspartate (D) of FlaH deviates in its spatial location from the other RecA family proteins.

similarity to the phospho-acceptor part of KaiC CII domain had functional implications, we performed a radioactive autophosphorylation assay. Purified *S. acidocaldarius* FlaH was incubated with γ-P^{32}-ATP at 55°C. Samples were collected at increasing time intervals from 0 to 120 minutes, then mixed with an SDS containing loading dye, which should remove the non-covalently bound ATP. The samples were next separated by a SDS-PAGE and analyzed for presence of radioactive labeled phosphate groups in FlaH. Indeed, a phosphorylation signal could be detected starting from after 30 min incubation (Fig. 9B). The signal intensity increased with prolonged incubation time, indicating FlaH autophosphorylation in a time dependent manner. The absence of a radioactive signal in the samples with no or a

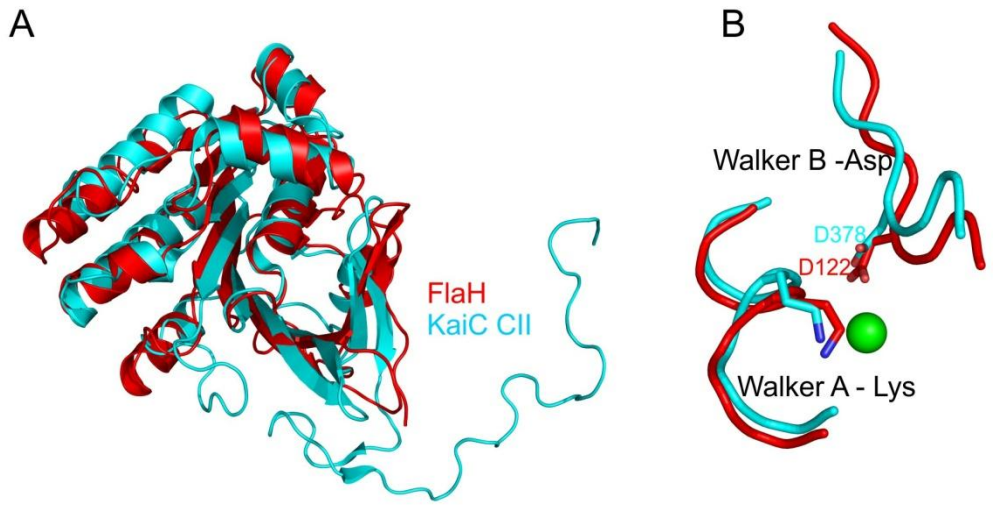


Fig.8. FlaH is similar to KaiC CII domain. KaiC is the circadian clock ATPase and the CII domain is responsible for regulating the circadian cycle in a phosphorylation dependent manner. A) The structural alignment of FlaH, red and KaiC CII domain, cyan. The aligment suggests that both of proteins share high structural similarity. B) Analysis of Walker B motifs of KaiC CII and FlaH, interestingly the aspartate (D) residues are aligned on the same spatial location.

shorter incubation times excludes a nonspecific signal by none covalently bound *y*-P_{32}-ATP.

In vivo analysis of potential phosphorylation sites of S. acidocaldarius FlaH-Based on both sterical and sequence similarities to the phosphorylated domain of KaiC, we assumed that residues T171 and S172 might be involved in the phosphotransfer within FlaH (Fig. 9A). However, these residues are not fully conserved among Archaea in contrast to T156 which is preserved in all analyzed archaeal FlaH homologues (Fig. 10). All the three considered residues were then mutated *in vivo* and a set of *S. acidocaldarius flaH* point mutation strains was created. At first the amino acids predicted as phosphorylation sites of FlaH were substituted by alanines, which resulted in $T_{156}A$, $T_{171}A$, $S_{172}A$ single mutants, a $T_{171}A/S_{172}A$ double mutant and finally a triple *flaH* mutant. The mutated strains were then analyzed phenotypically regarding their ability to swim on semi-solid gel ride plates. All the analyzed *flaH* mutants, together with Δ*aapF* and Δ*aapF*/Δ*flaH*, the positive and negative control strains respectively, were spotted on the same motility plate and incubated for 5-7 days at 76°C. After the incubation, the motility radius of the different mutants was compared to the swimming radius of the background strain Δ*aapF*. In all the analyzed strains a phenotypical defect could be detected as all of them showed significantly decreased swimming (Fig. 11 upper row). The strongest motility defect was observed in the *flaH* $S_{172}A$ mutant, which showed a 34% motility reduction. Surprisingly the $T_{171}A/S_{172}A$ double mutation had less influence on the swimming behavior then the $S_{172}A$ single mutation alone. The $T_{171}A/S_{172}A$ mutant showed around 15% swimming radius reduction, which equals the *flaH* single $T_{171}A$ mutant. The conserved T156 also seems to be strongly involved in the function of FlaH, because the alanine substitution of this residue reduced the motility radius around 20%. Interestingly, the mutagenesis of all the analyzed residues did not abolished the motility completely, but the ability to swim was reduced comparable to the $S_{172}A$ mutant to around 66% (fig. 11).

Because the alanine substitution resulted in diminished motility, we decided to perform an experiment to elucidate the effect of the phosphorylation of the analyzed FlaH residues. To do this we created *in vivo* mutations mimicking the phosphorylated state

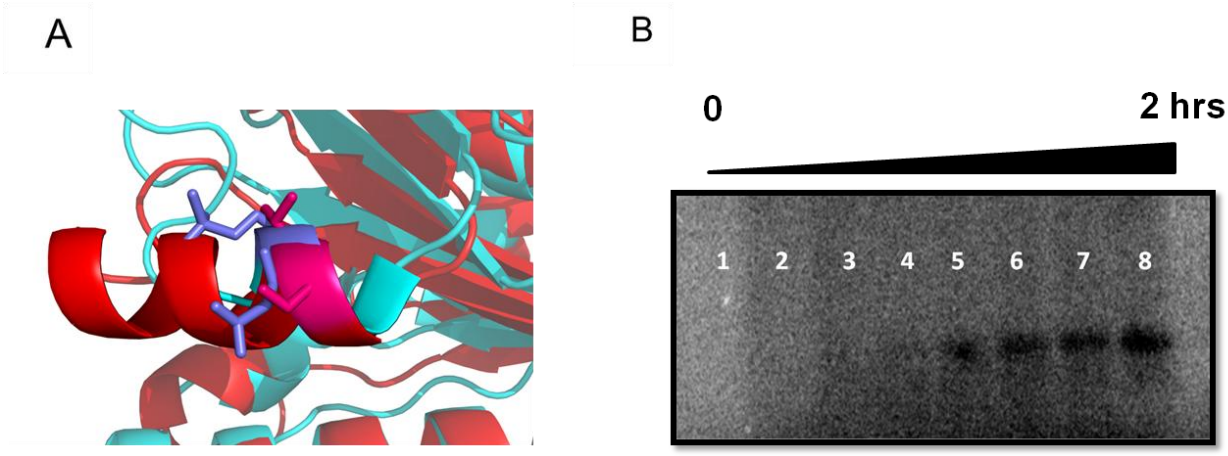


Fig.9. FlaH can autophosphorylate. **A)** The phosphorylation site of KaiC shows similarities in sequence and structure to FlaH. The phosphate acceptor residues S431T432 of KaiC CII (orange) are superimposing with the T171 or S172 of FlaH (yellow). **B)** Autophosphorylation assay showing that FlaH auto-phosphorylates. FlaH was incubated with γ-P_{32}-ATP at 55°C for 0, 5, 10, 15, 30, 60, 90 and 120 minutes corresponding to the samples 1-8, respectively.

of the chosen residues. Therefore as previously described a new set of *flaH* point mutants was generated: $T_{156}E$, $T_{171}E$, $S_{172}D$ single mutants, $T_{171}E/S_{172}D$ double mutant and$T_{156}E/T_{171}E/S_{172}D$ triple mutant. The mutated strains were again examined for the swimming behavior on semi-solid gelrite plates and compared to the background strain Δ*aapF*. Surprisingly, most of the generated *flaH* mutants here, turned out to be completely immotile. Only the *flaH* $T_{171}E$ mutant remained able to swim, but its motility radius was dramatically reduced to a level of 60% in comparison to the background strain Δ*aapF* (Fig.11 middle row).

In order to clarify if the abolished motility in the phosphorylation mimicking mutants is a consequence of targeting biophysical parameters or interruption of the archaellum assembly mechanism, we performed electron microscopy. The investigated strains were induced on starvation medium and analyzed by EM. We observed that these mutations are affecting already the archaellum assembly, since only in the motility restricted $T_{171}E$ mutant strain archaella could be detected, whereas all the other strains revealed to be not archaellated (Fig.11 lower row)

Due to the lack of functional antibodies against FlaH we could not confirm the cellular stability of FlaH mutants by immunoblot analysis. Also our efforts to detect tagged *in cis* or *in trans* complementation of FlaH failed so far. However based on the knowledge that FlaH is crucial for the archaella assembly, we can conclude that the effect observed is rather due to the mutation and not because of the mutated FlaH are not expressed or are not stable (Lassak et al., 2012b). Thus most of the analyzed FlaH mutants show a reduced and not an abolished swimming behavior.

Discussion

The three genes *flaHIJ* are conserved in all *fla* operons among Archaea and are proposed to form the archaellum assembly core complex (Ghosh & Albers, 2011b, Banerjee et al., 2013). In order to understand the archaellum assembly we characterized FlaH in the present study. Our structural and functional analysis of FlaH revealed that it is an ATP binding protein. The ATP bound structure of FlaH suggests involvement of both Walker A and Walker B box, while *in vivo* and *in vitro* analysis revealed that both WA and WB are essential for ATP binding and subsequently for archaellation. FlaH shows significant homology to the CII domain of the circadian clock protein KaiC. Both display structural

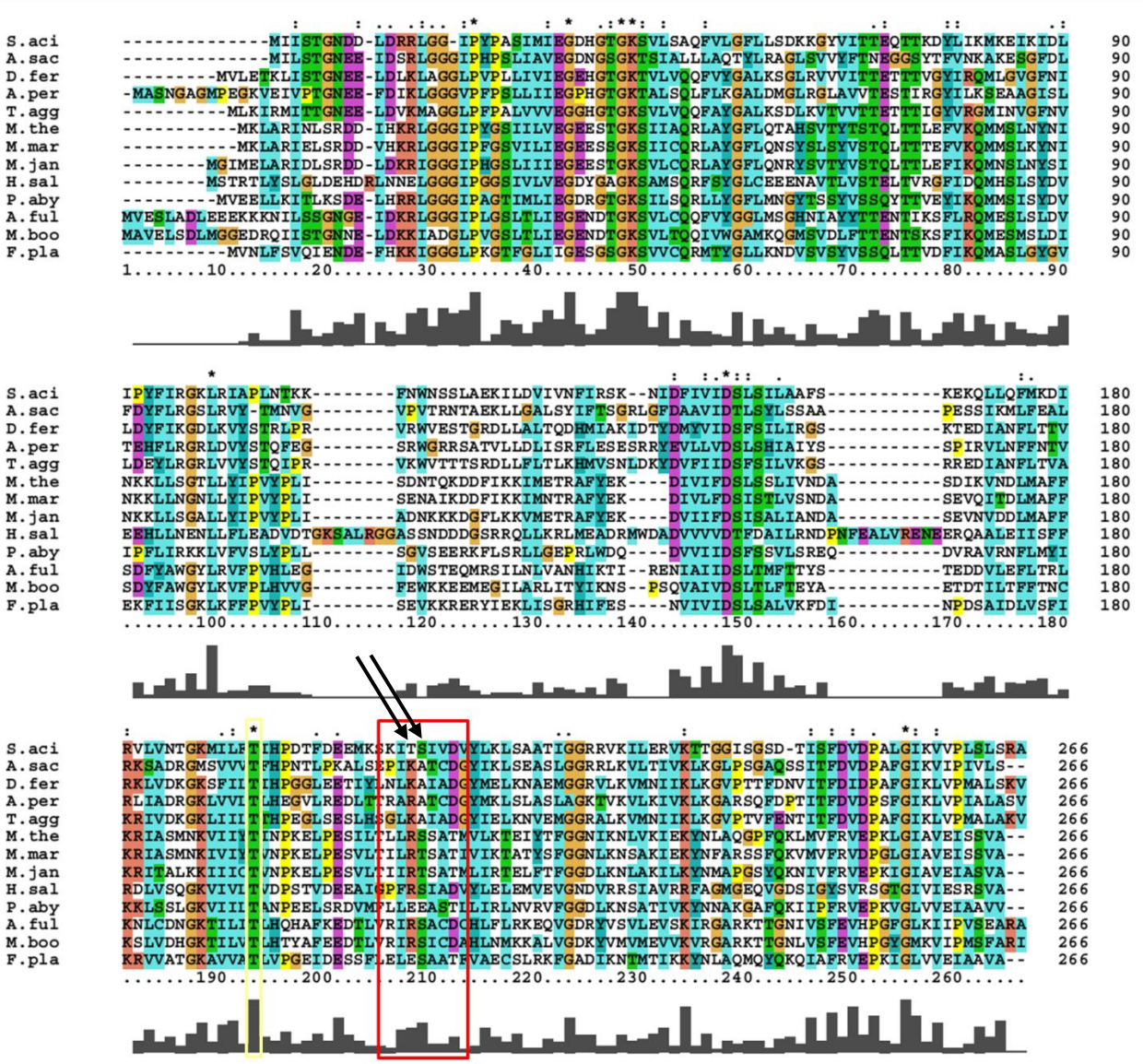


Fig.10. Multiple sequence alignment of FlaH homologues from different archaeal species generated in ClustalX. The conserved threonine 156, considered as involved in FlaH autophosphorylation is marked with a yellow frame. The T171 and S172 are indicated by black arrows. Although these residues are not directly conserved in archaea, they are considered as *S.acidocaldarius* FlaH phosphorylation sites based on the homology to KaiC. However the week conservation of these residues does not exclude a preserved pshosphorylation mechanism within FlaH, because all the analyzed FlaH homologues include potential phosphate acceptors like serines, threonine or threonines within this structural part (marked with a red frame).

similarities in the sterical organization of the ATP-binding pocket, in contrast to other RecA family proteins, where the conserved Asp within the Walker B motif is present at a different location relatively to FlaH. Although other RecA family members, like the DNA-repair protein RadA show high ATPase activity, the CII domain of KaiC is known to hydrolyze a limited number of ATP molecules per 24 hour cycle (Qian *et al.*, 2007, Egli et al., 2012), while for FlaH no ATPase activity has been observed so far. Thus probably FlaH binds ATP to assist in archaella assembly/rotation (regulatory role), not necessarily to hydrolyze it. Further structural comparison between FlaH and KaiC revealed high similarities in the region containing the already well characterized KaiC phosphorylation sites. First, we tested and confirmed FlaH autophosphorylation

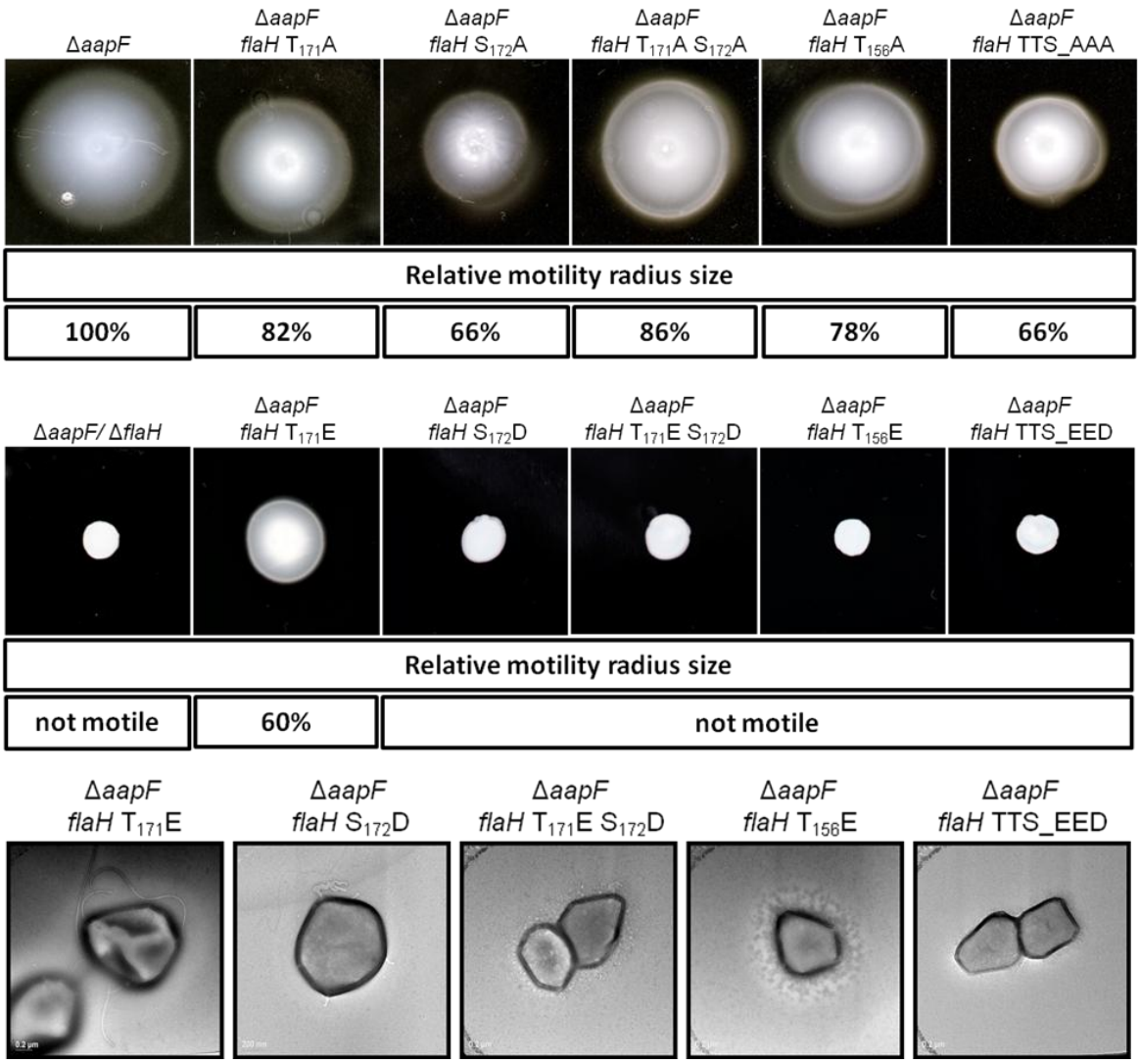


Fig.11. Motility assays of *S.acidocaldarius flaH* point mutants on semi solid gelrite plates. Potential phosphorylation sites were mutated *genomically* causing an effect on the swimming radius of the cells. **Upper row:** Substitution of the predicted phosphorylation sites with alanine reduced the motility significantly. The S172 to A mutation resulted in a 34% reduction of the swimming radius, but in the T_{171}/S_{172} double mutant the motility radius was comparable with the single $T_{171}A$ mutant. Substitution of all the predicted phosphorylation sites reduced the motility to the $S_{172}A$ single mutation level. **Midle row:** In contrast the mutants mimicking the phosphorylated state revealed mostly a non-motile phenotype. Only the *flaH* $T_{171}E$ mutant was still able to move however its swimming radius was reduced to 60% in comparison with the background strain. **Lower row:** Electron microscopy showed that all the non-motile FlaH phosphorylation mimicking mutants lack archaella, confirming that the introduced mutations affect already the archaellum assembly.

activity. We have made efforts to identify the phosphorylated sites of FlaH by mass spectrometry, but unfortunately all our attempts failed so far. For this reason we employed a bioinformatic approach to select the potential autophosphorylation sites of FlaH. Interestingly, the residues T171 and S172 in FlaH of *S. acidocaldarius*, which corresponds to the phosphorylated sites S431T432 of KaiC, are not conserved in Archaea. However after bioinformatic analysis we recognized that this particular FlaH region is a flexible loop. Although the mentioned residues are not conserved, all the aligned proteins contain in this region threonines, serines or tyrosines which could be targets for phosphorylation (Fig. 10). In addition we included T156 in our analysis, because on contrary to the previously described T171 and S172, the T156 is highly

conserved in FlaH homologues among all archaellated archaea, both in representatives of crenarchaeota as well as euryarchaeota (Fig. 9). Mutation in these particular sites resulted in reduced motility, clearly showing the importance of the T156, T171 and S172 residues for function of *S. acidocaldarius* FlaH. Interestingly, all the mutations of FlaH affect the motility in a different manner, moreover mutation of T171 to alanine could suppress the effect of the S172 substitution, indicating that all the here analyzed residues might be involved in a functional or regulatory pathway within the archaellum biosynthesis pathway. Looking again at the conservation of *flaHIJ* in the archaeal *fla* operons it appears unlikely that the phosphorylation activity of FlaH is strain specific and restricted to *S. acidocaldarius*. It seems rather that FlaH homologues differ in the phosphorylation pathway and final phosphate acceptors. However the mutated residues can be also part of the phosphorylation pathway, but not necessarily final phosphate acceptors. Nevertheless to definitively connect the obtained phenotypical effects with the auto-phosphorylation activity, a direct proof needs to be obtained. For this purpose MS/MS analysis of phosphorylated proteins will be performed in further experiments. In order to confirm our hypothesis that the autophosphorylation activity of FlaH is a global phenomenon, we plan to exam the FlaH homologue from the euryarchaeaon *Pyrococcus furiosus*. Our first attempts to purify this protein were successful. We want to perform a comparative study on this euryarchaeal FlaH homolog, which belongs to an archaellum system with a different accessory protein composition, in order to elucidate how strong is the functional invariability of FlaH in Archaea in general.

It has long been postulated that FlaH might play a regulatory role towards FlaI, redirecting its activity from archaellum assembly into rotation. The recent results obtained with the crenarchaeal model organism *S. acidocaldarius* would tend toward this hypothesis. It was shown already that FlaX, FlaH and FlaI interact with each other building the archaellar basal body, where FlaX forms a ring around FlaI, acting as a structural scaffold (Banerjee et al., 2013). In present study we established that the C-terminal part of FlaX is crucial for the interaction with FlaH. Thus both FlaI and FlaH bind to the same site of FlaX, which is additionally essential for the formation of the ring-like oligomers. Moreover using single particle analysis we could visualize that FlaH fill the FlaX rings, forming a second, inner ring. The space left is however to narrow to incorporate a 14 nm FlaI hexamer, suggesting that FlaX-FlaH-FlaI are not organized in one plane (Banerjee et al., 2013). Because both domains of FlaI interact with FlaX a direct contact between these proteins must be preserved. Thus it is reasonable to belief, that FlaI is integrated inside the FlaX tube, whereas FlaH is "sitting" on the interface between FlaI and FlaX. This potential architecture would also strongly support the regulatory role of FlaH towards FlaI, which seems to be based on FlaH autophosphorylation activity. It might be that FlaH is acting as the molecular switch, redirecting the activity of FlaI between archaella assembly and rotation, what is in a line with the *S. acidocaldarius* FlaH *in vivo* analysis. If this would be the case the swimming phenotypes could be explained as result of to early or to late switching of FlaI into rotation mode. Thus in this scenario the cell appendages should differ significantly in size and the swimming radius reduction would appear as result of hydrodynamically disadvantageous dimensions of the archaella. To answer this question we intend to perform Electron Microscopy, subsequently analyze statistically and compare the archaella lengths in all the created *in vivo* mutants. Further live imaging of swimming behavior of these strains will be used to investigate the changes of biophysical parameters of their motility. The results of this analysis could give us a detailed view in the exact effect of the introduced mutations. However, to conclusively link the phosphorylation of FlaH with the activity mode of FlaI, more detailed studies need to be done.

References

Adams, P.D., P.V. Afonine, G. Bunkoczi, V.B. Chen, I.W. Davis, N. Echols, J.J. Headd, L.W. Hung, G.J. Kapral, R.W. Grosse-

Kunstleve, A.J. McCoy, N.W. Moriarty, R. Oeffner, R.J. Read, D.C. Richardson, J.S. Richardson, T.C. Terwilliger & P.H. Zwart, (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta Crystallogr D 66: 213-221.

Albers, S.V., Z. Szabo & A.J. Driessen, (2003) Archaeal homolog of bacterial type IV prepilin signal peptidases with broad substrate specificity. J Bacteriol 185: 3918-3925.

Banerjee, A., A. Ghosh, D.J. Mills, J. Kahnt, J. Vonck & S.V. Albers, (2012a) FlaX, a unique component of the crenarchaeal archaellum, forms oligomeric ring-shaped structures and interacts with the motor ATPase FlaI. J Biol Chem.

Banerjee, A., A. Ghosh, D.J. Mills, J. Kahnt, J. Vonck & S.V. Albers, (2012b) FlaX, a unique component of the crenarchaeal archaellum, forms oligomeric ring-shaped structures and interacts with the motor ATPase FlaI. The Journal of biological chemistry 287: 43322-43330.

Banerjee, A., T. Neiner, P. Tripp & S.V. Albers, (2013) Insights into subunit interactions in the Sulfolobus acidocaldarius archaellum cytoplasmic complex. The FEBS journal 280: 6141-6149.

Brock, T.D., K.M. Brock, R.T. Belly & R.L. Weiss, (1972) Sulfolobus: a new genus of sulfur-oxidizing bacteria living at low pH and high temperature. Archiv fur Mikrobiologie 84: 54-68.

Buchan, D.W.A., S.M. Ward, A.E. Lobley, T.C.O. Nugent, K. Bryson & D.T. Jones, (2010) Protein annotation and modelling servers at University College London. Nucleic Acids Research 38: W563-W568.

Classen, S., I. Rodic, J. Holton, G.L. Hura, M. Hammel & J.A. Tainer, (2010) Software for the high-throughput collection of SAXS data using an enhanced Blu-Ice/DCS control system. Journal of synchrotron radiation 17: 774-781.

Cole, C., J.D. Barber & G.J. Barton, (2008) The Jpred 3 secondary structure prediction server. Nucleic Acids Research 36: W197-W201.

Consortium, U., (2012) Reorganizing the protein space at the Universal Protein Resource (UniProt). Nucleic Acids Research 40: D71-D75.

Egli, M., T. Mori, R. Pattanayek, Y. Xu, X. Qin & C.H. Johnson, (2012) Dephosphorylation of the core clock protein KaiC in the cyanobacterial KaiABC circadian oscillator proceeds via an ATP synthase mechanism. Biochemistry 51: 1547-1558.

Emsley, P. & K. Cowtan, (2004) Coot: model-building tools for molecular graphics. Acta Crystallogr D 60: 2126-2132.

Ghosh, A. & S.V. Albers, (2011) Assembly and function of the archaeal flagellum. Biochem Soc T 39: 64-69.

Ghosh, A., S. Hartung, C. van der Does, J.A. Tainer & S.V. Albers, (2011a) Archaeal flagellar ATPase motor shows ATP-dependent hexameric assembly and activity stimulation by specific lipid binding. Biochem J 437: 43-52.

Ghosh, A., S. Hartung, C. van der Does, J.A. Tainer & S.V. Albers, (2011b) Archaeal flagellar ATPase motor shows ATP-dependent hexameric assembly and activity stimulation by specific lipid binding. The Biochemical journal 437: 43-52.

Hanson, P.I. & S.W. Whiteheart, (2005) AAA+ proteins: Have engine, will work. Nat Rev Mol Cell Bio 6: 519-529.

Hayashi, F., N. Itoh, T. Uzumaki, R. Iwase, Y. Tsuchiya, H. Yamakawa, M. Morishita, K. Onai, S. Itoh & M. Ishiura, (2004) Roles of two ATPase-motif-containing domains in cyanobacterial circadian clock protein KaiC. The Journal of biological chemistry 279: 52331-52337.

Hayashi, F., H. Suzuki, R. Iwase, T. Uzumaki, A. Miyake, J.R. Shen, K. Imada, Y. Furukawa, K. Yonekura, K. Namba & M. Ishiura, (2003) ATP-induced hexameric ring structure of the cyanobacterial circadian clock protein KaiC. Genes to cells : devoted to molecular & cellular mechanisms 8: 287-296.

Holm, L. & P. Rosenstrom, (2010) Dali server: conservation mapping in 3D. Nucleic acids research 38: W545-549.

Hura, G.L., A.L. Menon, M. Hammel, R.P. Rambo, F.L. Poole, S.E. Tsutakawa, F.E. Jenney, S. Classen, K.A. Frankel, R.C. Hopkins, S.J. Yang, J.W. Scott, B.D. Dillard, M.W.W. Adams & J.A. Tainer, (2009) Robust, high-throughput solution structural analyses by small angle X-ray scattering (SAXS). Nat Methods 6: 606-U683.

Jarrell, K.F. & S.V. Albers, (2012) The archaellum: an old motility structure with a new name. Trends in microbiology 20: 307-312.

Jarrell, K.F. & M.J. McBride, (2008) The surprisingly diverse ways that prokaryotes move. Nature Reviews Microbiology 6: 466-476.

Ken F. Jarrell, D.J.V.a.J.W., (2009) Archaeal Flagella and Pili. Caister Academic Press, UK.

Konarev, P.V., V.V. Volkov, A.V. Sokolova, M.H.J. Koch & D.I. Svergun, (2003) PRIMUS: a Windows PC-based system for small-angle scattering data analysis. Journal of applied crystallography 36: 1277-1282.

Larkin, M.A., G. Blackshields, N.P. Brown, R. Chenna, P.A. McGettigan, H. McWilliam, F. Valentin, I.M. Wallace, A. Wilm, R. Lopez, J.D. Thompson, T.J. Gibson & D.G. Higgins, (2007) Clustal W and clustal X version 2.0. Bioinformatics 23: 2947-2948.

Lassak, K., T. Neiner, A. Ghosh, A. Klingl, R. Wirth & S.V. Albers, (2012) Molecular analysis of the crenarchaeal flagellum. Molecular microbiology 83: 110-124.

Margolis, R.L. & L. Wilson, (1981) Microtubule Treadmills - Possible Molecular Machinery. Nature 293: 705-711.

McCoy, A.J., R.W. Grosse-Kunstleve, P.D. Adams, M.D. Winn, L.C. Storoni & R.J. Read, (2007) Phaser crystallographic software. Journal of applied crystallography 40: 658-674.

Nakajima, M., K. Imai, H. Ito, T. Nishiwaki, Y. Murayama, H. Iwasaki, T. Oyama & T. Kondo, (2005) Reconstitution of circadian oscillation of cyanobacterial KaiC phosphorylation in vitro. Science 308: 414-415.

Otwinowski, Z. & W. Minor, (1997) Processing of X-ray diffraction data collected in oscillation mode. Method Enzymol 276: 307-326.

Peabody, C.R., Y.J. Chung, M.R. Yen, D. Vidal-Ingigliardi, A.P. Pugsley & M.H. Saier, Jr., (2003) Type II protein secretion and its relationship to bacterial type IV pili and archaeal flagella. Microbiology 149: 3051-3072.

Qian, X.G., Y.J. He & Y. Luo, (2007) Binding of a second magnesium is required for ATPase activity of RadA from Methanococcus voltae. Biochemistry 46: 5855-5863.

Reindl, S., A. Ghosh, G.J. Williams, K. Lassak, T. Neiner, A.L. Henche, S.V. Albers & J.A. Tainer, (2013) Insights into FlaI functions in archaeal motor assembly and motility from structures, conformations, and genetics. Molecular cell 49: 1069-1082.

Rust, M.J., J.S. Markson, W.S. Lane, D.S. Fisher & E.K. O'Shea, (2007) Ordered phosphorylation governs oscillation of a three-protein circadian clock. Science 318: 809-812.

Thomas, N.A., C.T. Pawson & K.F. Jarrell, (2001) Insertional inactivation of the flaH gene in the archaeon Methanococcus voltae results in non-flagellated cells. Mol Genet Genomics 265: 596-603.

Wagner, M., M. van Wolferen, A. Wagner, K. Lassak, B.H. Meyer, J. Reimann & S.V. Albers, (2012) Versatile Genetic Tool Box for the Crenarchaeote Sulfolobus acidocaldarius. Frontiers in microbiology 3: 214.

Supplementary to **'Structural and functional characterization of FlaH: *an archaellum clock?'***

Tomasz Neiner[#], Ankan Banerjee[#], Edoardo D'Imprima, Sophia Reindl, Abhrajyoti Ghosh, Andrew S. Arvai, Janet Vonck, John A.Tainer and Sonja-Verena Albers

Spplementary figures

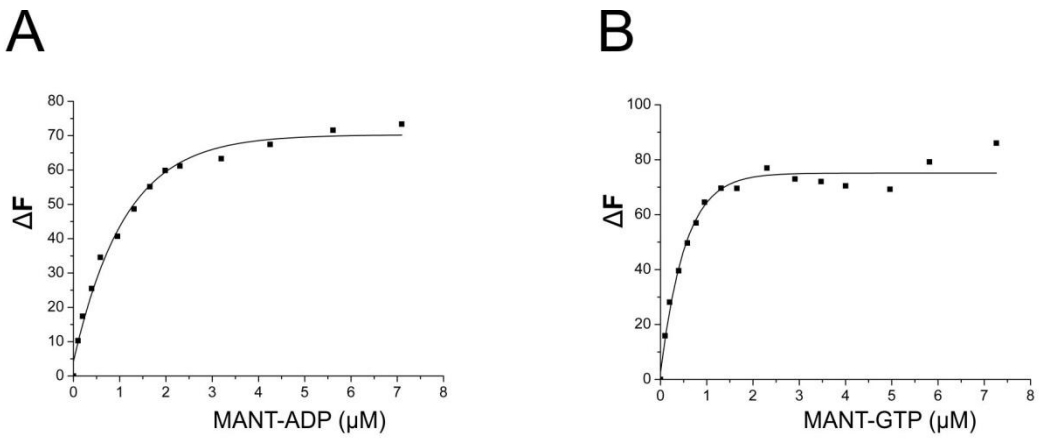


Fig.S1 FlaH can bind to any given nucleotide. A) FRET analysis of the MANT-ADP binding to FlaH. **B)** FRET analysis of the MANT-GTP binding to FlaH.

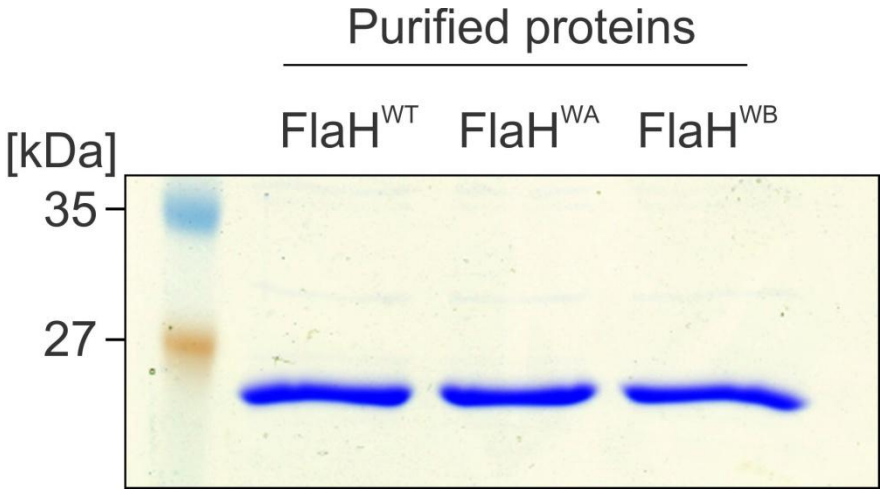


Fig.S2 Pure FlaH Variants. SDS-PAGE of pure FlaH variants (FlaH^{WT}, FlaH^{WA}, FlaH^{WB}) used in MANT-ATP binding.

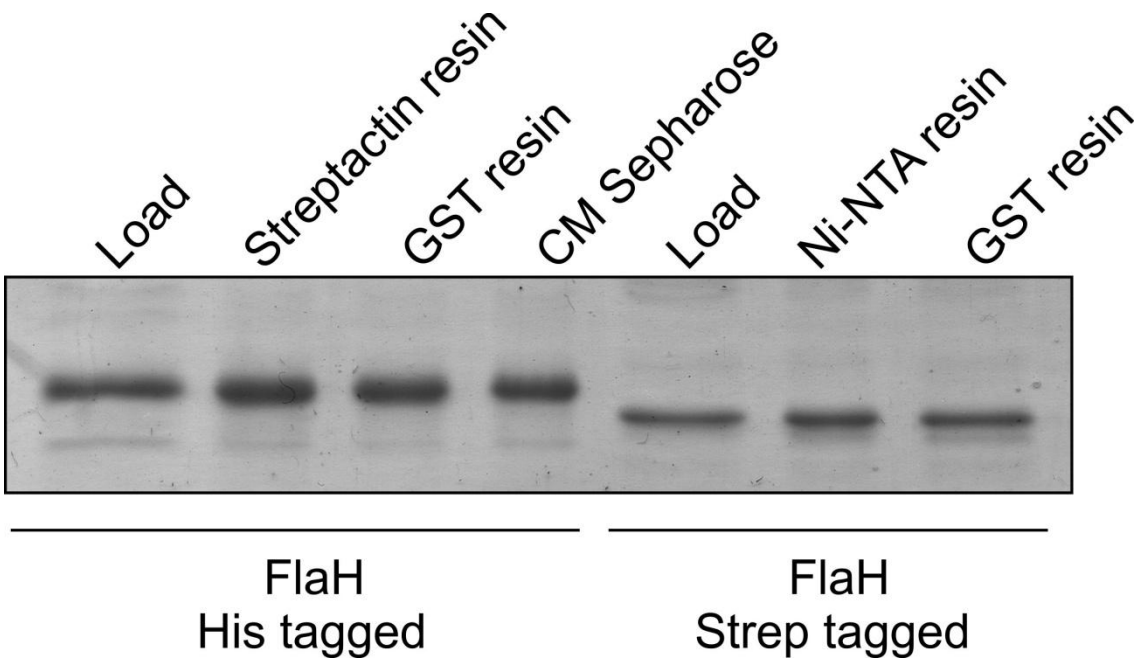


Fig.S3 FlaH invariably binds to affinity beads. Either His_6 tagged or StrepII tagged FlaH was incubated with streptactin or GST-resin or Ni-NTA or CM-sepharose (cation exchange) beads. After an extensive washing step the bound fraction was collected using 2500x g centrifugation for 5 min and loaded on to a denaturing PAGE. SDS-PAGE was visualized using coomassie staining.

Supplementary tables

Table S1. Crystallographic data collection and refinement statistics

Data collection	
Space group	C 2
Cell dimensions (Å/°)	a=103.4; b=52.9; c=73.6
Wavelength (Å)	1.0
Resolution range (Å)	50 – 2.3
Completeness (%)[1]	85.5/43.2
Observed reflections	28,367
Unique reflections	11,612
Multiplicity[1]	2.4/1.9
R_{sym} (%)[1]	2.8/33.3
mean I/σ[1]	19.3/2.4
Refinement statistics	
Resolution range (Å)	42.9 – 2.3
R_{work} (%)/R_{free} (%)	24.8/19.3
Average B-factor (Å^2)	60.1
protein	59.8
solvent	64.6
Number of atoms	1878
protein	1765
ligand	32
water molecules	81
RMSD bond lengths	0.005
RMSD bond angles	0.98
Ramachandran favored (%)	95
Ramachandran allowed (%)	0.45
PDB ID	xxx.pdb

[1] Overall/last shell
[2] R_{sym} is the unweighted R value on I between symmetry mates.
[3] R_{free} is the cross-validation R factor for 10% of reflections against which the model was not refined.

Table S2. Strains used in present study

Strains	Relevant characteristics	Source
E.coli		
DH5α	l2 f80d/lacZ DM15 D(lacZYA-argF)U169 recA1 endA1 hsdR17 (rK2 mK1) supE44 thi-1 gyrA relA1	Gibco
Bl21(DE3)-RIL	B F-ompThsdS(rB − mB −) dcm + Tetr E. coli gal λ (DE3) endAHte [argUileYleuWCam^{r}]	Stratagene
ER1821	F^{-}*glnV44 e14^{-}(McrA^{-}) rfbD1? relA1? endA1 spoT1? thi-1* Δ*(mcrC-mrr)114::IS10*	NEB
S.acidocaldarius		
DSM639	Wild-type *Sulfolobus acidocaldarius*	DSMZ
MW001	DSM 639 Δ*pyrE*	(17)
MW156	MW001 Δ*aapF* (*Saci_2318*)	(4)
MW455	MW156 Δ*flaH* (*Saci_1174*)	(4)
MW466	MW156 *flaH* K33A	This study
MW467	MW156 *flaH* D122N	This study
MW469	MW156 *flaH* T171A	This study
MW470	MW156 *flaH* S172A	This study
MW471	MW156 *flaH* T171A S172A	This study
MW472	MW156 *flaH* T156A	This study
MW473	MW156 *flaH* T156A T171A S172A	This study
MW474	MW156 *flaH* T171E	This study
MW475	MW156 *flaH* S172D	This study
MW476	MW156 *flaH* T171E S172D	This study
MW477	MW156 *flaH* T156E	This study
MW478	MW156 *flaH* T156E T171E S172D	This study

Table S3 Plasmids used in present study

Plasmids	Relevant characteristics	Source
pETDuet-1	Amp^r ,Cam^r , expression plasmid containing replicon ColE1 (pBR322) and two MCS (MCS1 and MCS2)	Novagen
pUC57	Amp^r, *E. coli* cloning vector	
pSVA251	Amp^r ,Cam^r , pETDuet-1 carrying untagged *flaI* in MCS2 using restriction sites NdeI-XhoI	(Ghosh et al., 2011a)
pSVA284	Amp^r codon optimized *flaH (sflaH*) in pUC57	
pSVA293	Amp^r, Cam^r, pETDuet-1 carrying C-terminal His_6 tagged s*flaH* in MCS2 using restriction sites NdeI-XhoI	This study
pSVA406	Amp^r, gene targeting plasmid, pGEM-T Easy backbone, pyrEF cassette of S.solfataricus	(Wagner et al., 2012)
pSVA1911	Del 37 FlaX (*flaXc*) gene in pSA4 using restriction sites NcoI-BamHI	(Banerjee et al., 2012a)
pSVA1934	Amp^r, Cam^r, pETDuet-1 Carrying N-terminal His_6 tagged *flaXcn*(37^{th}-127^{th} amino acid FlaX) in MCS1 using restriction sites EcoRI-HindIII	(Banerjee et al., 2012a)
pSVA1935	Amp^r, Cam^r, pETDuet-1 Carrying N-terminal His_6 tagged *flaXcc1*(87^{th}-250^{th} amino acid FlaX) in MCS1 using restriction sites EcoRI-HindIII	(Banerjee et al., 2012a)
pSVA2100	Amp^r, Cam^r, pETDuet-1 carrying N-terminal His_6 tagged s*flaH* in MCS1 using restriction sites EcoRI-HindIII	This study
pSVA2108	Amp^r, Cam^r, pETDuet-1 carrying N-terminal Strep tagged s*flaH* in MCS2 using restriction sites NdeI-XhoI	This study
pSVA2113	Amp^r, *flaH* K33A cloned into pSVA406 with ApaI-PstI	This study
pSVA2124	Amp^r, *flaH* D122N cloned into pSVA406 with ApaI-PstI	This study
pSVA2126	Amp^r, *flaH* cloned into pSVA406 with ApaI-PstI	This study
pSVA2130	Amp^r, Cam^r, pETDuet-1 carrying N-terminal His_6 tagged s*flaH* K33A in MCS1 using restriction sites EcoRI-HindIII	This study
pSVA2131	Amp^r, Cam^r, pETDuet-1 carrying N-terminal His_6 tagged s*flaH* D122N in MCS1 using restriction sites EcoRI-HindIII	This study
pSVA2149	Amp^r, *flaH* T171A cloned into pSVA406 with ApaI-PstI	This study
pSVA2150	Amp^r, *flaH* S172A cloned into pSVA406 with ApaI-PstI	This study
pSVA2151	Amp^r, *flaH* T171A S172A cloned into pSVA406 with ApaI-PstI	This study
pSVA2152	Amp^r, *flaH* T156A cloned into pSVA406 with ApaI-PstI	This study
pSVA2169	Amp^r, *flaH* T171E cloned into pSVA406 with ApaI-PstI	This study
pSVA2170	Amp^r, *flaH* S172D cloned into pSVA406 with ApaI-PstI	This study
pSVA2171	Amp^r, *flaH* T171E S172D cloned into pSVA406 with ApaI-PstI	This study
pSVA2172	Amp^r, *flaH* T156E cloned into pSVA406 with ApaI-PstI	This study
pSVA2173	Amp^r, *flaH* T156E T171E S172D cloned into pSVA406 with ApaI-PstI	This study
pSVA2174	Amp^r, *flaH* T156A T171A S172A cloned into pSVA406 with ApaI-PstI	This study

Table S4. Primer details

Primers	Sequence and characteristics	Source
771	5'-GGGGGCATATGATTATCAGCACCGGCAACG-3'; Forward primer for *sflaH* containing an NdeI restriction site (underlined)	This study
772	5'-GGGGGCTCGAGTTAATGATGATGATGATGATGTGCGCGAGACAGGC-3'; Reverse primer for *sflaH* containing a XhoI restriction site (underlined)	This study
786	5'-GGGGGCTCGAGTTATGCGCGAGACAGGC-3'; Reverse primer for *sflaH* containing a XhoI restriction site (underlined)	This study
2168	5'-GGGGGCATATGTGGTCCCATCCTCAATTTGAAAAGATTATCAGCACCGGCAACG-3'; Forward primer for *sflaH* containing an NdeI restriction site (underlined)	This study
3612	5'-CCCCCCGAATTCAATGATTATCAGCACCGGC-3'; Forward primer for *sflaH* containing a EcoRI restriction site (underlined)	This study
3613	5'-GGGGGGAAGCTTTTATGCGCGAGACAGGC-3'; Reverse primer for *sflaH* containing a HindIII restriction site (underlined)	This study
3614	5'-CCCGGGGGGCCCGGGGGAAACAACGATCTC-3'; Forward primer for *flaH* down-stream containing a ApaI restriction site (underlined)	This study
3615	5'-CACTCGCACCCGTGCCATGATCTCC-3'; Forward primer for *flaH*K33A site directed mutagenesis	This study
3616	5'-GCGAGTGTGTTATCTGCTCAATTCG-3'; Reverse primer for *flaH*K33A site directed mutagenesis	This study
3617	5'-CCCCCCCTGCAGCCATCATAGAGGATAATGTTCC-3'; Reverse primer for *flaH* up-stream containing a PstI restriction site (underlined)	This study
3627	5'-GACAAACTATTGATTACGATGAAGTCTATATTT-3'; Forward primer for *flaH* D122N site directed mutagenesis	This study
3628	5'-GTAATCAATAGTTTGTCAATACTAGCTGCTTTCTC-3'; Reverse primer for *flaH* D122N site directed mutagenesis	This study
3630	5'-GCTCCATTGAATACGGAGGAATTTAATTGGAACTC-3'; Forward primer for *flaH* KK91-92EE site directed mutagenesis	This study
3631	5'-GAGTTCCAATTAAATTCCTCCGTATTCAATGGAGC-3'; Reverse primer for *flaH*KK91-92EE site directed mutagenesis	This study
3632	5'-GATCATGGCACGGGTGCAAGTGTGCTGAGCG-3'; Forward primer for s*flaH* K33A site directed mutagenesis	This study
3633	5'-CGCTCAGCACACTTGCACCCGTGCCATGATC-3'; Reverse primer for s*flaH* K33A site directed mutagenesis	This study
3634	5'-CGATTTCATTGTTATCAATAGTCTGAGCATTCTG-3'; Forward primer for s*flaH* D122N site directed mutagenesis	This study
3635	5'-CAGAATGCTCAGACTATTGATAACAATGAAATC-3'; Reverse primer for s*flaH* D122N site directed mutagenesis	This study
3645	5'-ATGAAGAGTAAAATAGCTAGTATAGTTG-3'; Forward primer for s*flaH* T171A site directed mutagenesis	This study

3646	5’-CACATCAACTATACTAGCTATTTTAC-3’; Reverse primer for s*flaH* T171A site directed mutagenesis	This study
3647	5’-GAAGAGTAAAATAACTGCTATAGTTGATG-3’; Forward primer for s*flaH* S172A site directed mutagenesis	This study
3648	5’-ATACACATCAACTATAGCAGTTATTTTAC-3’; Reverse primer for s*flaH* S172A site directed mutagenesis	This study
3649	5’-GAAGAGTAAAATAGCTGCTATAGTTGATGTG-3’; Forward primer for s*flaH* T171A S172A site directed mutagenesis	This study
3650	5’-CACATCAACTATAGCAGCTATTTTACTCTTC-3’; Reverse primer for s*flaH* T171A S172A site directed mutagenesis	This study
3651	5’-GATTCTATTCGCTATTCATCCAGATAC-3’; Forward primer for s*flaH* T156A site directed mutagenesis	This study
3652	5’-GTATCTGGATGAATAGCGAATAGAATC-3’; Reverse primer for s*flaH* T156A site directed mutagenesis	This study
3691	5’-GAGTAAAATAGAAAGTATAGTTGATGTGTATCTTAAGC-3’; Forward primer for s*flaH* T171E site directed mutagenesis	This study
3692	5’-CACATCAACTATACTTTCTATTTTACTCTTCATTTCC-3’; Reverse primer for s*flaH* T171E site directed mutagenesis	This study
3693	5’-GAGTAAAATAACTGATATAGTTGATGTGTATCTTAAGC-3’; Forward primer for s*flaH* S172D site directed mutagenesis	This study
3694	5’-CACATCAACTATATCAGTTATTTTACTCTTCATTTCC-3’; Reverse primer for s*flaH* S172D site directed mutagenesis	This study
3695	5’-GGAAATGAAGAGTAAAATAGAAGATATAGTTGATGTGTATCTTAAGC-3’; Forward primer for s*flaH* T171E S172D site directed mutagenesis	This study
3696	5’-GCTTAAGATACACATCAACTATATCTTCTATTTTACTCTTCATTTCCTC GTC-3’; Reverse primer for s*flaH* T171E S172D site directed mutagenesis	This study
3697	5’-GATTCTATTCGAAATTCATCCAGATACTTTTGACG-3’; Forward primer for s*flaH* T156E site directed mutagenesis	This study
3698	5’-GTATCTGGATGAATTTCGAATAGAATCATTTTACCCG-3’; Reverse primer for s*flaH* T156E site directed mutagenesis	This study

4. Discussion

Motility is a highly important feature of the microbial universe. This form of active locomotion is often mediated by effective and complex molecular machineries, like the eukaryotic cilia or bacterial flagellum and type IV pili. All these eukaryotic and bacterial structures were intensively studied during the last decades, where in contrast the archaeal motility organelle remained scarcely investigated. Because of the exterior similarities to the bacterial flagellum, the motility organelle of Archaea was initially assumed as structurally related and described as archaeal flagellum. However, with the increasing number of sequenced genomes of representatives of the third domain it became clear that Archaea lack homologues of the bacterial flagellum and eukaryotic cilia. Thus Archaea must exhibit a distinct motility system. Further genomic analysis has identified a locus encoding genes classified as pilins (Szabo *et al.*, 2007b). Interestingly this locus was highly conserved among Archaea and could be found in all motile species. Surprisingly this identified "*fla*" operon exhibits remarkably similarities to the bacterial T4P. The pilin related proteins FlaA and FlaB include a class III signal peptide as the bacterial T4P. Moreover the *fla* locus encodes the ATPase FlaI and the polytopic membrane protein FlaJ, which resemble the T4P components PilB/T and PilC, respectively. The main limitation in archaeal research is the limited number of available genetic systems within this domain. However *fla* deletion strains could been created in the euryarchaea *Halobacterium salinarum* and *Methanococcus voltae*, resulting in non-motile strains, which conclusively linked the *fla* operon with the motility structure (Patenge et al., 2001, Thomas *et al*., 2002). During these studies a class III signal recognizing peptidase was identified as part of this system, representing another overlay with the T4P (Bardy & Jarrell, 2002). Moreover, the archaeal motility structure propels the cells in a rotary fashion, similar to the bacterial flagella, but completely unprecedented for a bacterial T4P structure. Hence the archaeal motility structure can be described as "rotating type IV pilus" (Shahapure et al., 2014a). Finally the appearing contrasts between the archaeal motility organelle and both, the architecture of the bacterial flagellum and mode of operation of T4P resulted in renaming the archaeal flagellum into the "archaellum" (Jarrell & Albers, 2012). The renaming of the archaeal organelle was supposed to avoid the misleading link to the bacterial flagellum, since the archaellum is a unique structure and represents a completely different, third way to move. Up to date the archaeal motility investigation was restricted to phylogenetic, genetic and physiological analysis of *Halobacterium sp.* and *Methanococcus sp*, which are all representatives of the subkingdom of Euryarchaeota. Using *Sulfolobus acidocaldarius* as model organism, we have demonstrated here the first systematic analysis of a crenarchaeal archaellum. The archaellum of *S. acidocaldarius* includes only seven different components: the archaellin FlaB; the motor protein FlaI; polytopic membrane protein FlaJ and 4 accessory proteins FlaX, FlaG, FlaF, FlaH. Hence it is not surprising that all these encoding proteins revealed to be essential for archaella assembly. Moreover all analyzed *fla* deletion strains exhibited completely non-motile phenotypes which could be, in every case, restored by *in trans* complementation of the deleted gene. These results conclusively linked the archaellum with swimming motility in Crenarcheota (Lassak et al., 2012b).

This study has set its focus in unraveling the structure and the assembly model of the archaellum of *S. acidocaldarius*. First, we tried to identify interaction partners by analyzing the *in vivo* stability of Fla proteins by Western blot analysis. However, we were strongly limited in this approach, because we had only antibodies against two out of seven *S. acidocaldarius* Fla proteins, namely FlaB and FlaX. Nevertheless we could show for the first time interaction between the archaellar components. We were unable to detect FlaX in strains lacking FlaHIJ, which was the first suggestion for possible interactions between these proteins. FlaB is building the filament,

which is most likely assembled as the final step, after all other Fla component occupied already their places within the archaellum assembly machinery. However it seems that FlaB is stably expressed independently from the other Fla components, since it was missing only in the *flaB* deletion strain (Lassak et al., 2012b). Thus in absence of other Fla proteins FlaB still accumulates in the cell, but it cannot be assembled into a functional filament.

Bioinformatical approaches suggested that FlaB shares structural similarities with the bacterial T4P pilins. Both structures possess a long α-helical N-terminus extending from the packed β strands C-terminus. Another common feature of the T4P pilins and the archaellins are the class III signal peptides. Previous studies on *S. solfataricus* have already confirmed the presence of a specific aspartic acid protease PibD, which recognizes and processes the class III signal peptides in FlaB (Albers et al., 2003). Recently we could obtain similar results in our model organism *S. acidocaldarius* (Dr. Ankan Banerjee, personal communication). A very interesting feature is the *N*-linked glycosylation of the archaellar filament components, highly ubiquitous in Archaea (Meyer *et al.*, 2011, Tripepi *et al.*, 2012). The deletion of N-glycosylation pathway components of *M. voltae* and *M. maripaludis* resulted in decreased molecular weight of the archaellin subunits which could been visualized by faster migration of these proteins in SDS-PAGE and western blot analysis (Chaban *et al.*, 2006b, VanDyke *et al.*, 2009, Jarrell *et al.*, 2010). In both *M. voltae* and *M. maripaludis* mutations disrupting the addition of terminal sugar into the glycan resulted in archaellated, but less motile cells (Jarrell et al., 2010). Whereas the limitation of glycan compositions to only one sugar abolished also the archaella assembly in this species (Jarrell et al., 2010). A current study in *H. volcani* could go even further and prove that the *N*-linked glycosylation of its major archaellin FlaA1 is essential for the stability and functionality of its archaella (Tripepi et al., 2012). So far for *S. acidocaldarius* it was only shown that the deletions of certain components of the glycosylation pathway resulted in reduced molecular weight of FlaB on SDS-PAGE (Meyer et al., 2011). Therefore we decided to perform a more detailed study on the correlation between the glycosylation and motility of *S. acidocaldarius.* FlaB has six predicted glycosylation sites: N^{145}, N^{173}, N^{195}, N^{214}, N^{218}, N^{229}. We constructed systematic *in vivo* mutants, where the predicted asparagines as glycan tree acceptors where substituted by alanines, generating strains carrying single or different combination of multiple FlaB mutations. One of the constructed strains was lacking all six predicted asparagines. First we analyzed the motility of glycosyltransferase mutants we had from other studies, where we observed that mutations affecting the glycan three compositions led to abolished archaella biosynthesis and revealed non-motile phenotypes (Dr. Benjamin Meyer, personal communication). When we examined the effect of FlaB glycosylation on motility, we have observed that all the *flaB* point mutants remain still motile. Even after deleting all the predicted glycosylation sites of FlaB, motility was only slightly reduced, but not completely abolished (Dr. Benjamin Meyer, personal communication). Our results show that an intact glycosylation pathway is indeed crucial for proper archaellum assembly and motility in *S. acidocaldarius*; however the posttranslational modification of FlaB directly plays only a secondary role for the archaella stability in *S. acidocaldarius.*

The second protein encoded in the *fla* operon of *S. acidocaldarius* is FlaX, which is a unique archaella component and can be found exclusively in crenarchaeota. Previous studies demonstrated that FlaX can form higher order ring shaped oligomers, with diameters around 30 nm, revealing a dominating 18-fold symmetry (Banerjee et al., 2012b). It was also shown that the of three α-helices composed C-terminus of this protein plays an essential role in the formation of the ring-like structures. We proved furthermore that the C-terminal part of FlaX is decisive for the

interactions with both, FlaI (Chapter 3.3) and FlaH (Chapter 3.4). Moreover we demonstrated that both termini of FlaI can interact with FlaX independently. Taking into account that hexameric FlaI is approximately 14 nm in diameter, we postulated that FlaI may function inside the FlaX ring complexes. In addition we have been able to demonstrate high affinity interactions between FlaX, FlaH and FlaI which form a ternary complex *in vitro*. Therefore, we concluded that FlaX may serve as the molecular scaffold for the crenarchaeal archaellum assembly. So far archaellar basal bodies could be isolated exclusively from few representatives of Euryarchaeotas. Structural analysis of archaella obtained from *Methanococcus voltae* revealed a knob-like architecture, distinct from the ring based basal body typical for Bacteria (Kalmokoff et al., 1988). Further analysis of archaella isolated from other methanogens, like *Methanospirillum hungatei* or representatives of *Methanococcus* spp. showed a similar basal body structure like *M. voltae (Southam et al., 1990a, Faguy et al., 1994)*. Moreover knob-like basal bodies could been also observed in *Halobacterium salinarium* (Kupper et al., 1994), suggesting that this basal body architecture might be a common feature of euryarchaeal archaella.

In case the crenarchaeal archaella shows a similar knob-like architecture one could assume that the FlaX rings could represent the envelope of this structure, giving the shape of its basal body. Recently it has been suggested that crenarchaeal FlaX and euryarchaeal FlaC might be orthologous. Despite these proteins do not share any sequence similarities, *in silico* analysis revealed structural homology between FlaC and FlaX. Moreover first attempts of FlaC characterization showed similar biochemical behavior for these proteins (Prof. Dr. Sonja-Verena Albers, personal communication). Thus it seems that the crenarchaeal and euryarchaeal archaellum might be more similar than previously assumed.

FlaCDE have been already linked with the archaeal chemotaxis system, which is present in euryarchaeata but not crenarchaeota (Schlesner et al., 2009). This is supported by the fact that we could so far not observe any directed movement under the conditions assayed so far. However a recent study has reported that the switching of rotation direction of *S. acidocaldarius* archaella is not a so rare phenomenon as previously believed and appeared in around 10% of the analyzed cells (Shahapure et al., 2014a). Interestingly FlaX shows distant homology to the methyl accepting chemotaxis protein (MCP). It was shown that MCPs are methylated at the E or Q residues present in the characteristic A/S-X-X-E-[E/Q]-X-[A/T/S]-A-[A/T/S] sequences (Hagman *et al.*, 1997). ClustalW analysis revealed that FlaX carries a similar peptide sequence S^{112}-A-V-E-E-L-A-L-S^{120}, where the residue E^{116} might be a potential methylation site. Thus it might be that FlaX initiates the switch between clockwise (CW) to counterclockwise (CCW) rotation of the crenarchaeal archaellum. However, the methylation of FlaX has not been investigated so far and still reminds an open question. Nevertheless *in vivo* point mutations within this domain would elucidate the functional importance of this part of FlaX.

The archaella accessory components FlaG and FlaF both include domains referring to the archaellin super family, however they lack the characteristic class III signal peptide and are therefore not considered as functional archaellins (Ghosh & Albers, 2011b). Preliminary data indicated strong interactions between these proteins (Dr. Ankan Banerjee and Patrick Tripp, personal communication). So far not much is known about the biochemical and functional properties of FlaG and FlaF, however recently significant progress was made in terms of FlaF. The crystal structure of the soluble domain of FlaF has been solved, revealing eight closely packed anti parallel β-sheets forming a β-sandwich structure (Dr. Ankan Banerjee and Patrick Tripp, personal communication). In addition FlaF forms dimers and strongly binds to the S-layer, most probably recognizing the highly abundant glycan trees (Dr. Ankan Banerjee and Patrick Tripp, personal communication). These preliminary results indicate functional similarities to

MotB, which is a stator component of the bacterial flagellum. MotB dimers bind to the bacterial peptidoglycan, forming a larger ring complex which fixes the bacterial flagella into the cell wall acting thereby as a stator (Kojima *et al.*, 2009). FlaF may perform an analogues role in the archaellum, stably fixing the entire complex into the cell envelope, additionally creating a pore that would enable the assembly of the archaellum. Considering that in *S. acidocaldarius* the gaps between the single S-layer lattices are not wider then 4-5 nm, the existence of a broader molecular gate appears to be conditioning the penetration and elongation of a 10-14 nm in diameter filament. This scenario would additionally explain the non-archaellated phenotype observed in the mentioned above *S. acidocaldarius* strains with impaired glycosylation pathway. These mutants displayed incomplete glycans, what would support the idea that FlaF targets directly the glycan tree, not the S-layer proteins alone. Although to confirm this we would need to isolate the S-layer from mutants displaying aberrant glycan tree composition and perform interaction studies with FlaF *in vitro*.

The so far best characterized component of *S. acidocaldarius* archaellum is the motor ATPase FlaI. It belongs to the "secretion ATPase superfamily", sharing homology with ATPase from bacterial type II/IV secretion and type IV pili systems. *S. acidocaldarius* FlaI functions as a hexamer, reaching its maximal ATPase activity at pH 6.5 at 75^{o}C (Ghosh et al., 2011a). In this work we focused on the structural analysis of FlaI (Chapter 3.2). First we obtained the crystal structure of FlaI in its hexameric form. It was solved at 2.0 Å resolution, showing a crown-like structure with a 14 nm diameter and a 7.5 nm thickness (Reindl et al., 2013).

Since long time it has been believed that the archaella assembly as its rotation mediated motility is ATP driven (Streif *et al.*, 2008). The fact that all known archaella systems include only one functional ATPase, which is additionally close related to motor proteins powering completely distinct bacterial machineries, was always a paradox. It was hard to imagine, that FlaI may power at the same time a T4P like assembly model and generate a completely different filament rotation based type of movement. However in this work we validated this controversial model. By deleting the first 29 amino acid at the N-terminus of FlaI *in vivo*, we obtained a *S. acidocaldarius* strain which was still able to synthesize archaella, but remained completely non-motile. Thus we were able to uncouple the archaella assembly and rotation functions from each other, proving clearly the dual role of FlaI in the archaella system. Furthermore we were able to explain this intriguing phenomenon. During structural analysis of the molecular dynamic of FlaI we identify a clear difference in the ATP-processing induced molecular motion of this protein compared to its bacterial homologues. We observed that FlaI moves directed parallel to the plane of the hexamer, whereas the motions of the secretion ATPase GspE are directed perpendicular (Yamagata & Tainer, 2007). Hence this difference in rotatory conformational change between FlaI and GspE proves FlaI to be potentially suitable to drive the archaella assembly and rotation. An interesting point for further investigation would be the elucidation of the biomechanical aspect of FlaI rotation in terms of understanding the importance of the first 29 amino acids of FlaI for the movement generation. We intent to perform *in vivo* point mutations within this domain in order to identify crucial functional residues for the archaellum rotation. Moreover, the here proposed mechanism of movement generation will require a stator complex. Considering the architecture of the basal body proposed in Chapter 3.3, FlaX is most probably a part of the archaellar stator complex. It would be interesting to confirm if the flexible N-terminal domain is involved in FlaX-FlaI interactions. In order to answer that question, we would need to perform domain mapping using only the N-terminal region of FlaI, since we showed that both domains of FlaI can interact with FlaX independently. These approaches could help us to fully understand the mechanism in which the archaellum rotation is generated.

Similarly to the bacterial T4P system, the archaellum assembly machinery includes a polytopic membrane protein. Although bacterial PilC and archaeal FlaJ show homology, they significantly differ in the number of transmembrane domains (Peabody et al., 2003b). PilC has usually three transmembrane helices, whereas their number is much more variable for FlaJ (Karuppiah *et al.*, 2010). Euryarchaeal FlaJ homologues usually have nine predicted transmembrane domains, whereas crenarchaeal FlaJ has between five and seven (Peabody et al., 2003a). FlaJ of *S. acidocaldarius* is predicted to have seven transmembrane domain and two big cytoplasmic loops, termed CytoI and CytoII (Ghosh & Albers, 2011a). Long time ago it was postulated that these cytosolic domains are the functional regions of this protein. To exam the importance of the Cyto loops, we recently constructed mutants, where the amino acid region 41-121 of CytoI and 240-351 of CytoII were deleted individually and together. Phenotypical analysis of these strains revealed that all lack archaella and are therefore non-motile, suggesting an essential role of the Cyto domains in the archaellum assembly in *S. acidocaldarius* (Dr. Ekatarina Kapotova, personal communication). As we discussed previously in Chapters 3.2 and 3.3, the charge distribution in the Cyto domains of FlaJ presents a strong overlay with the charged crown grooves of FlaI, indicating an electrostatic interaction between these proteins. Analogous interaction between polytopic membrane proteins and assembly ATPases were strongly discussed for bacterial systems and could even be shown for *Pseudomonas aeruginosa* type II secretion and its T4P (Arts *et al.*, 2007, Takhar et al., 2013). Interestingly the structural analysis of FlaI showed that the top surface moves oscillatory up and down in an 8–10 A˚ range during ATP processing. This shift would correspond to the theoretic size of the archaellin FlaB. Thus our model assumes that this conformational change in FlaI is transferred through the interaction with FlaJ Cyto domains to incorporate new subunits by pushing the growing archaellum filament outside the cell. A similar mechanism was also proposed for the filament assembly in the T4P system, where the assembly ATPase induces upon ATP hydrolysis conformational changes in the inner membrane core protein leading to filament extraction (Melville & Craig, 2013). Moreover Nivaskumar et al. showed recently that the T4P of *Klebsiella oxytoca* assembles in a slightly rotational fashion (Nivaskumar *et al.*, 2014). In the presented work they suggested that the ATPase GspE (PilT) promotes a gentle counter clock-wise rotation of the membrane protein GspF (PilC), which strongly indicates a direct interaction between these proteins.

In order to confirm its hypothesized interactions with FlaI *in vitro,* we intensively tried to express FlaJ. However, all our multiple attempts of homologues and heterologues protein expression methods were unsuccessful so far. Unfortunately, bacterial FlaJ homologues could not be expressed and purified. So far only two cytoplasmic domains PilC/GspF family proteins could be crystallized, the $EpsF^{56-171}$ from *V. cholerae* and $PilC^{1–168}$ from *Thermus thermophilus* (Abendroth *et al.*, 2009, Karuppiah et al., 2010). Thus, we employed a similar approach that involved the expression of single Cyto domains, however also here homologues or heterologous expression could not be obtained. Expression of FlaJ and FlaI from a different organism and exploration of their interaction in a parallel archaellum system could be an interesting alternative to overcome the difficulties that we encountered in *S. acidocaldarius. Pyrococus furiosus* is an organism often used for archaeal structural characterization, because its proteins often show astonishing stability not comparable with other homologues. *P. furiosus* is additionally highly archaellated and habitats similar environments as *Sulfolobus*. Therefore first attempts of heterologous expression of *P. furiosus* Fla proteins have been already initiated.

The main focus of my PhD project was the biochemical and functional characterization of the archaella accessory protein FlaH. In chapter 3.4 we reported the solved crystal structure of FlaH, which revealed that FlaH exhibit a single RacA domain composition lacking any site

domains. FlaH was predicted to be an ATPase-like protein due to the presence of a Walker A and Walker B box. Whereas the Walker A motif of FlaH is well conserved, Walker B motif however lacks a glutamate and shows a shifted spatial location compared to other RecA family proteins with ATPase activities. Despite multiple experimental trials, we were unable to detect any ATP hydrolysis in case of FlaH. Moreover the crystallized protein revealed the presence of a non-hydrolyzed ATP molecule in the nucleotide binding pocket of FlaH, suggesting that FlaH simply does not have any ATPase activity. Nevertheless we could show that both the Walker motifs are directly involved in the binding of ATP, since mutation of lysine 33 in Walker A box as well asparagine 122 in Walker B abolished MANT-ATP binding in the FRET assay. Hence it seems that FlaH is an ATP-binding protein and not a functional ATPase. The corresponding mutations *in vivo* confirmed moreover that ATP binding is crucial for the archaella assembly in *S. acidocaldarius*.

The biggest question regarding FlaH *in vivo* studies was the cellular stability of the FlaH mutants. Unfortunately we were unable to obtain functional antibodies against this protein so far, therefore we could not use classical Western blot analysis to confirm that the introduced mutations do not affect FlaH stability in the cell. The majority of the FlaH point mutants analyzed in chapter 3.4 can however still synthesize their archaella, what would be impossible without stable FlaH, which was already confirmed as an essential element of the archaellum puzzle. Based on this knowledge an indirect test of FlaH stability would be possible. Since we have functional antibodies against FlaX and we know that it requires the presence of FlaH to exist in the cell, we can simply check the presence of FlaX in all the FlaH mutants. This indirect test seems to be not fully satisfactory in order to confirm if mutated FlaH remains stable, however it may help to at least preselect mutant strains where it do not.

Another option would be *in trans* or *in cis* expression of Strep or His tagged versions of FlaH point mutations and subsequently performed Western blot analysis targeting the introduced tags. Although it is a possibility to confirm the presence of FlaH in the cell, it remains always a risk that the tag may influence the archaellum assembly, leading to an unwanted phenotypical effect. Thus these tagged mutants could then be used to confirm the cellular stability exclusively, losing its value for the phenotypical characterization of the mutants. Nevertheless we have already made new efforts to generate functional antibodies against the remaining Fla proteins of *S. acidocaldarius* including FlaH. Therefore a new set of freshly purified Fla proteins was sent for antibodies production and hopefully those can be generated in this trial. In parallel we are determining the expression level of FlaH mutants via RT-qPCR. However expression of a gene do not guarantees the stability of the corresponding protein, therefore this approach is not solving this issue.

Here, we have been able to elucidate the interaction network around FlaH, confirming its strong affinity towards FlaI and FlaX. We could demonstrate that these proteins are involved in the basal body formation of the archaellum of *S. acidocaldarius.* However the biggest achievement of this study is the partial reconstitution of this basal body *in vitro.* We were capable to obtain stable FlaX-FlaH complexes *in vitro,* which could be further used for single particle analysis. This approach was our first step towards the characterization of the architecture and stoichiometry of the archaellum complex. Moreover it allowed us to obtain a first structural visualization of internal archaellar components interactions, what is a milestone in archaeal motility research.

However, it was still important to proof the FlaX-FlaH-FlaI interaction model. The main obstacle has been the diversity of buffers, in which the three components are purified. The FlaX-FlaH interaction could be visualized in a high pH which blocks the FlaI hexamer formation. After series of trials with different conditions we could observe the appearance of distinctive,

completely filled rings (Fig.1.) Unfortunately these rings represent trace amounts in the solution, therefore we need to increase the amounts dramatically to obtain a structural model of this triple complex. Nevertheless these partial results confirm our hypothesis, that FlaX serves as a molecular scaffold in the archaellum assembly and together with FlaH and FlaI forms the basal body of the archaellum.

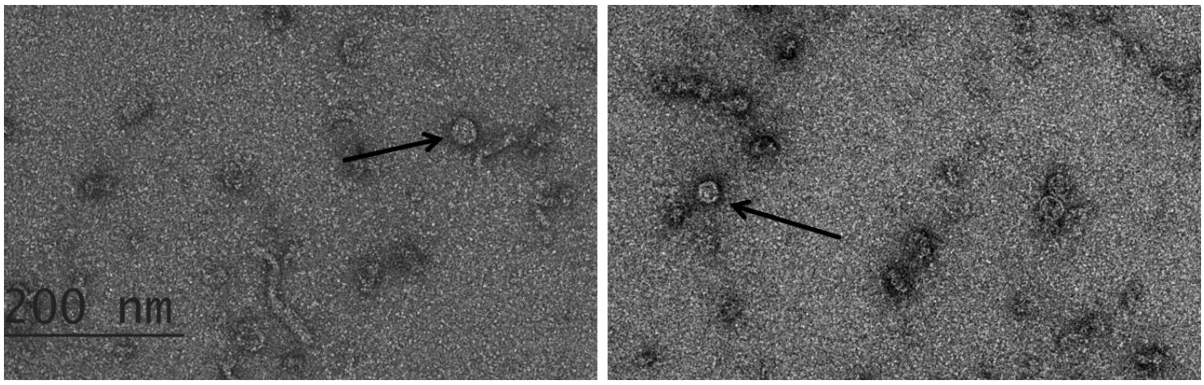


Fig.1. After *in vitro* mixing purified FlaX, FlaH and FlaI in presence of 2mM ATP and 5mM Mg^{2+} distinctive, completely filled ring-shaped structures could be observed.

During our analysis of *S. acidocaldarius* FlaH structure, we observed a remarkably homology to the CII domain of the circadian clock protein KaiC. Both proteins show astonishing similarities in the organization of their ATP binding pockets, deviating significantly from other RecA family proteins. The CII domain is the functional part of KaiC, which undergoes cyclic autophosphorylation and autodephosphorylation events according to the daily rhythms (Nakajima et al., 2005). Therefore, it was tempting to determinate if FlaH also exhibits autophosphorylation activity. We could observe the appearance of phosphate groups covalently bound to FlaH after incubating it with γ-P_{32}-ATP for at least 30 minutes, which confirmed that FlaH has autophosphorylation activity.

The located in the CII domain residues S431 and T432 have already been established as the final phosphate acceptors of KaiC (Xu *et al.*, 2004). Furthermore Xu et al. proved the importance of a third threonine (T426) in the phosphorylation pattern of KaiC, which is most probably an intermediate phosphate-acceptor (Xu *et al.*, 2009, Pattanayek *et al.*, 2009). Mutations of all the three residues T426, S431 and T432 identified as phosphorylated within KaiC resulted in disrupted rhythmicity of this circadian clock (Xu et al., 2004, Xu et al., 2009). The phosphorylation status of alanine substitution of these residues was also analyzed *in vitro*. The $T_{426}A$ and $T_{432}A$ strongly decreased the KaiC phosphorylation, whereas the $S_{431}A$ mutant in contrast increased the phosphorylated/non-phosphorylated ratio of KaiC (Xu et al., 2004). A triple mutation of these residues abolished the phosphorylation of KaiC completely, proving that they represent all the phosphate acceptors within this circadian oscillatory system (Xu et al., 2004).

We were able to predict potential phosphorylation sites within FlaH: T156, which is highly conserved in all archaeal FlaH and T171 and S172, which refer to the phosphorylation sites of KaiC are, however, *S. acidocaldarius* specific. Interestingly the T156 in FlaH of *S. acidocaldarius* precedes the T171 and S172 similar like T426 precedes the final phospho-acceptors S431 and T432 of KaiC, suggesting a further analogy in the phosphate transfer pattern. For this reasons, we decided to investigate the role of all those amino acids in FlaH function and mechanism using *in vivo* and *in vitro* approaches. Surprisingly, all variations of *in vivo* alanine substitution performed

on these residues resulted in phenotypical defects. Our *in vitro* investigation of the autophosphorylation activity of FlaH is an ongoing project. Unfortunately, we still cannot confirm if we already identified the entire phosphorylation sites within this protein. Thus our future plans include the *in vitro* autophosphorylation activity analysis of the FlaH T_{156}A/E, T_{171}A/E and S_{172}A/E mutations. *E. coli* expression vector carrying appropriate FlaH point mutations have been already prepared and will be analyzed in the nearest future.

Although KaiC can phosphorylate itself, it requires the presence of KaiA and KaiB to fully restore the circadian rhythms (Vakonakis & LiWang, 2004, Pattanayek et al., 2013, Egli et al., 2013). The interaction partners of FlaH within the archaellum have been already identified, however their influence on its phosphorylation activity was not investigated yet. Thus future experiment will focus on studying the autophosphorylation activity of FlaH in presence of other Fla components, especially its known interaction partners, FlaX and FlaI.

As mentioned before, one strategy to overcome some of the difficulties that we encounter during the archaellum characterization is to express the components of *P. furiosus* archaellum in a heterologous system. Preliminary results suggest that the FlaH of *P. furiosus* shows similar characteristic to its *S. acidocaldarius* homologue. We have been able so far to confirm the involvements of both Walker motifs of *P. furiosus* FlaH in the ATP binding and its autophosphorylation activity (see appendix). Moreover, using MST and pull down analysis we could also confirm interaction between *P. furiosus* FlaI and FlaH (data not shown, Paushali Chaudhury and Prof. Dr. Sonja-Verena Albers, personal communication). We decided to include the Walker mutants of *P. furiosus* FlaH in this interaction analysis. Surprisingly we could observe a strongly reduced affinity of the Walker mutated FlaH towards FlaI (Paushali Chaudhury, personal communication). Thus, it seems that the intact ATP-binding pocket is crucial for the FlaH-FlaI interactions. It will be interesting to confirm this unexpected finding in the *S. acidocaldarius* archaellum system, which definitely needs to be investigated.

As previously mentioned, for long time it has been postulated that FlaH can influence the function/activity of FlaI, but beyond the interaction between these two proteins, nothing supporting this hyphothesis has been demonstrated. Our preliminary results suggest that the ATPase activity of *P. furiosus* FlaI is reduced in presence of *P. furiosus* FlaH (Paushali Chaudhury personal communication). Despite our multiple attempts, we did not observe a similar effect for *S. acidocaldarius* proteins. However, *S. acidocaldarius* FlaI is around 100 times less active *in vitro* than *P. furiosus* homologue, what could lead to the assumption that a similar effect could be not detected due to limitations in the sensitivity of the technique employed. Additionally, the high temperature (around 70^{o}C) at which the assays were performed causes ATP degradation, what leads to analytical discrepancies, especially at a relative slow enzymatic activity of the examined protein.

The Archaellum assembly model:

Summarizing all available information considering the crenarchaeal archaellum components we created a detailed assembly model of the *S. acidocaldarius* archaellum (Fig.2). FlaJ is most probably the starting point of this assembly process (Fig.2 A) The Western blot analysis of *ΔaapFΔflaB-J* strains presented in chapter 3.1 show that FlaJ is required for the stability of FlaX, thus it must be the initiator of archaella assembly. After FlaJ takes its place in the membrane, it recruits the hexameric motor ATPase FlaI (Fig.2 B). The interaction between FlaJ and FlaI was discussed in chapters 3.2 and 3.3. After the FlaIJ complex is set, FlaH is integrated into the

formed core complex (Fig.2 C). Although FlaI and FlaX are known to interact with each other, we can conclude from the previously cited Western blot, that only the presence of all three FlaJ, FlaI and FlaH guarantees the cellular stability of FlaX. Thus it seems that FlaX cannot exist alone in the cell membrane and has a highly specified place in the archaellum assembly order, after FlaJ, FlaI and FlaH, respectively. At this stage FlaX surrounds the already build FlaHIJ complex forming the molecular scaffold of the archaellar basal body (Fig.2 D). FlaX forms ring-like structures with a diameter of approximately 30 nm that would enable FlaHIJ to fit inside (Banerjee et al., 2012a). The position of FlaX in the archaella assembly might be due to its role. Thus FlaX could be part of the stator, whereas FlaI and FlaJ are mobile elements generating torque and transmitting it on the rotating filament, respectively. FlaH in this situation creates a bridge between the motor and the stator complexes of the archaellum. This in-out assembly model seems to be justified, because a fixed stator would strongly impede the assembly of the inner motor complex. FlaF and FlaG are the only two components which are not affecting the cellular stability of FlaX, which indicates that they are incorporated into the archaellum after FlaX, when the entire core complex is already assembled (Fig.2 E). Preliminary results showed that FlaF and FlaG can interact with each other and FlaF can bind to the glycan tree present in the S-layer. Since the archaellum filament is too wide to penetrate the tightly packed S-layer, FlaF and FlaG would form a pore complex in order to overcome this obstacle. Additionally the anchoring into the S-layer would lead to a very stable placement of this complex. Thus we think that the FlaGF complex is directly linked to the FlaX ring via their N-terminal trans membrane domain and that all together form a tube, which is solidly fixed to the S-layer and represents the static part of the archaellum. A similar scenario is present in the bacterial flagellum, where the stator component MotB is fixed to the peptidoglycan, guaranteeing the high stability of the entire molecular machinery (Kojima et al., 2009). However, so far we could not confirm any interaction between FlaG or FlaF and FlaX. All proteins we have been expressed and purified as truncated soluble variants, lacking the transmembrane domains, which are most probably crucial for these putative interactions.

After the basal body is formed, the filament can be synthesized (Fig.2 F). The archaellum of *S. acidocaldarius* is formed by only one type of archaellin, FlaB. This archaellin is expressed as a pre-protein form and gets located in the membrane, where it undergoes a set of posttranslational modifications. FlaB gets activated by the peptidase PibD and is *N*-linked glycosylated to be subsequently incorporated in the bottom of the growing filament. It seems that FlaB must be present in the membrane already from the early stages of the archaellum assembly, because it also appears to be crucial for the stability of FlaX. We cannot state if the presence of FlaB is needed already for the proper FlaJ membrane incorporation or if it can affect FlaX stability directly. However a direct interaction between FlaX and FlaB would explain how the archaellins are recruited into the archaellum assembly machinery, when the movable core is already closed in the stator.

The role of FlaH within the Archaellum

Taking together all the presented results, FlaH might act as a molecular switch, directing the energy generated by FlaI between the assembly and the rotation of the archaellum. In theory the highest hydrodynamic effectiveness can be reached only when the archaellum is within a certain range of length. A too short filament would have an insufficient surface to convert its rotation into swimming movement. However a too long archaellum would at some point generate to high resistance to propel the cells forwards. Looking at the complex and highly energetically optimized

regulation of the archaella expression it seems unlikely that this aspect could be evolutionary overlooked. Thus we hypothesize that the archaellum elongation is stopped at a specified point, in which FlaI gets redirected from assembly into generating filament rotation. In our working model, FlaH would be responsible for switching between the modes of FlaI and thereby it is directly involved in the archaellum length determination (Fig.2 G).

A strong interaction between FlaH and FlaI was confirmed in this work with help of multiple methods. Moreover we could place FlaH as part of the basal body in the crenarcheal archaellum system, directly "sitting" on the interface between FlaX and FlaI. Additionally, MST measurement revealed differences in the interaction affinities among these proteins. The direct FlaX-FlaI interaction is much weaker then these between FlaX-FlaH and FlaI-FlaH. This suggested that FlaH may act as a bridge, strengthening the FlaX-FlaI interaction. This position between the motor protein FlaI and the potential stator FlaX, would strongly privilege the postulated role of FlaH as molecular switch.

Hypothetically FlaH may immobilize FlaI during the archaellum assembly, directing all the generated energy into the filament elongation by incorporating FlaB into it. On the other hand after reaching a particular point the FlaH-FlaI interaction might be loosened, freeing the FlaI hexamer and enabling its free rotation. Our model assumes that this transition is due to the change in the phosphorylation status of FlaH. In this case, the relaxation of the FlaH-FlaI interaction is caused by conformational changes within FlaH as result of its autophosphorylation. This scenario would fully explain the phenotypical defects caused by the mutations introduced into the putative phosphorylation sites of FlaH. We observed reduced motility in all the *flaH* phosphorylation mutants. This effect could be caused by shifting the assembly/rotation switching moment, resulting in archaella outside the optimum filament length range. In contrast the phosphorylation mimicking mutations show mostly a completely non-motile phenotype, what may indicate that in this case FlaH is from the beginning directed into promoting the rotary mode of FlaI. This situation would lead to cells lacking archaella, what we could indeed observe using electron microscopy. According to our model FlaH is holding FlaI during the archaellum assembly and releasing it in order to promote filament rotation. Thus the changing phosphorylation status of FlaH, directly impacts the FlaH-FlaI interaction force. Therefore as future plan we intend to express the corresponding FlaH phosphorylation mimic mutants *in vitro* and test them for FlaI interaction.

Unfortunately, at this stage we are unable to explain the exact mechanism responsible for this FlaI mode switching event. It appears reasonable to believe, that the analyzed residues of FlaH: T156, T171, S172 are all part of a complex phosphorylation pathway, since various mutation combinations impact the motility in a different manner. It seems that the Serine 172 plays a highly important role within this pathway, because it shows the most emphatic effect on motility, when mutated alone. However, involved in a double or triple mutation their impact on motility is weakened. Moreover our preliminary EM analysis indicated that the archaella of *S. acidocaldarius* carrying the *flaH* $S_{172}A$ mutation appeared much shorter then in wild type, or even in the *flaH* $T_{171}A$ or $T_{171}A/S_{172}A$ mutants. This would suggest that the S172A mutation lead to a too early switching of the archaellum assembly towards movement generation, before the filament can reach the mature size guaranteeing enough surface for sufficient energy transition. However, a conclusive statistical analysis is still missing.

FlaI and FlaH are well conserved among Archaea. Additionally our preliminary results showed also a strong relation between FlaI and FlaH in *P. furiosus*, including data indicating that FlaH might influence FlaI activity. Hence we believe that our hypothesis considering FlaH as a molecular switch might be true for all the archaella systems in the entire domain of Archaea. Since most of the archaella components except FlaX are conserved in Archaea, one could expect

that also the mode and hierarchy of the archaellum assembly is highly preserved in the entire third domain. Because of the limited chemotaxis need, Crenarchaeota evolved FlaX as a "simplified" stator component, which is most probably a replacement of FlaCDE found in other Archaea.

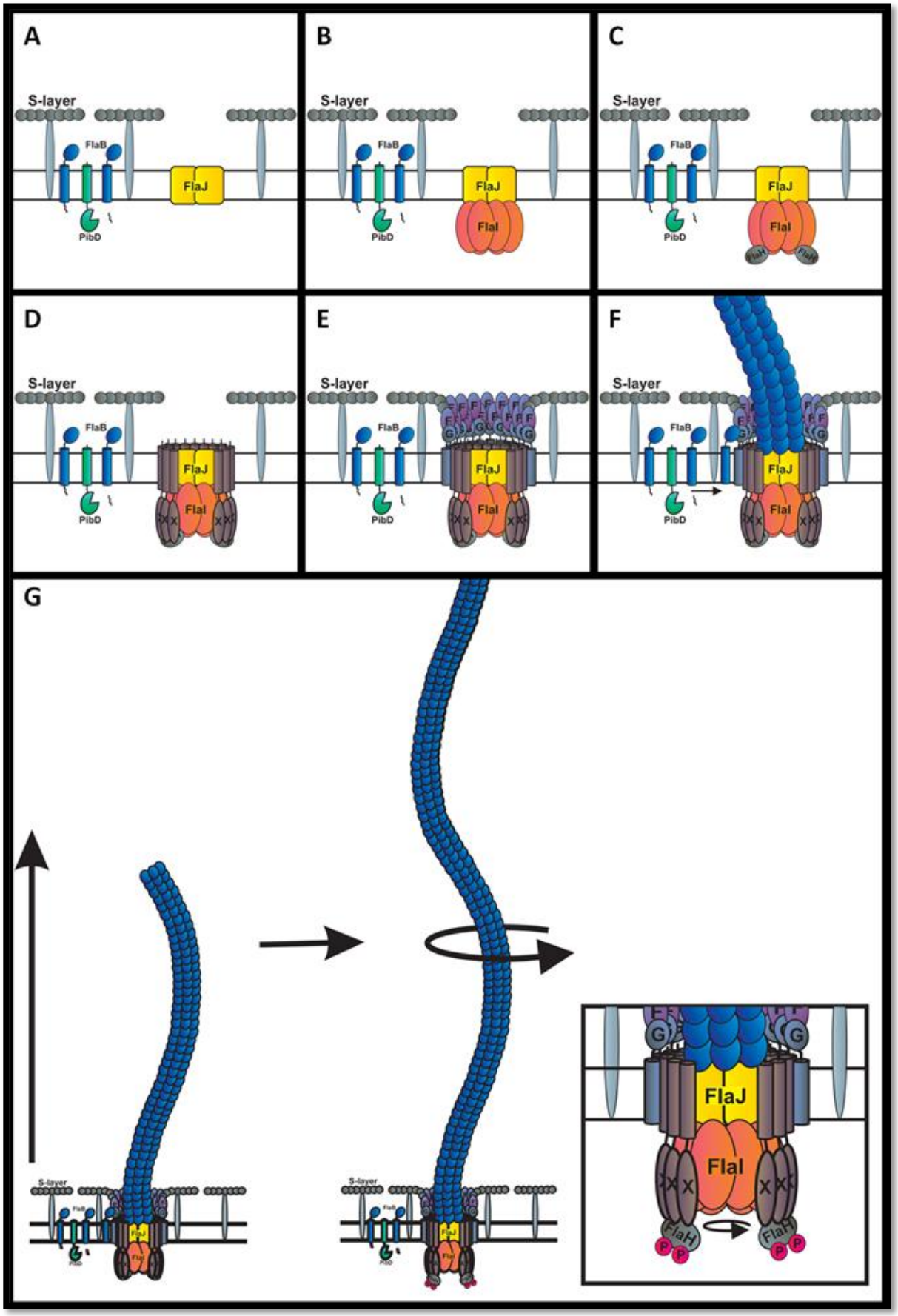
As the archaella biosynthesis is highly energy consuming it seems reasonable to restrict this process only to certain circumstances. In this study we could elucidate the conditions which promote the archaella expression in *S. acidocaldarius*. We were able to demonstrate, that this process is a specific response to nutrient limitation, which provided us a powerful tool for the further detailed structural and functional analysis of the crenarchaeal motility.

Since the archaella is expressed in “harder times” it appears reasonable that the entire biosynthesis machinery is strictly energetically optimized. Archaellation density is imposed by the number of assembled archaeal basal bodies. Assuming that the crenarchaeal archaellum basal body is in scale of the FlaX *in vitro* ring, it would require dozens, up to maybe hundreds of copies of different components, whereas the assembly of a single several cell lengths long filament consumes already thousands of subunits. Here we showed that the *fla* locus is divided into two transcriptional units, where the basal body components and the archaellins are expressed independently (Lassak et al., 2012b).

During this study we observed that Δ*aapF* strain, which lacks the Aap pilus, reacts much stronger to the starvation stress and exhibit often a multiple archaellated phenotype, resulting, moreover, in accelerated motility (Henche et al., 2012a, Lassak et al., 2012b). The reason for this phenotype is so far unclear, however it can indicate a crosstalk between the regulation of the Aap pili and the archaella. Both structures perform opposite functions; therefore silencing the Aap-pilus while increasing archaella biosynthesis seems to be reasonable.

As the archaella biosynthesis takes place under energy scarce conditions, the existence of a highly economically optimized regulatory system seems to be reasonable. Recently an archaella regulatory network was found in *S. acidocaldarius*, where two transcriptional activators ArnR and ArnR1 and a couple of repressors ArnA and ArnB could be identified (Reimann et al., 2012, Lassak et al., 2013). ArnR and ArnR1 are homologues; both contain a helix-turn-helix motif (HTH) and two single transmembrane domains. Both have been shown to localize in the cell membrane (Lassak et al., 2013). Deletion of *arnR* and *arnR1* resulted in reduced *flaB* expression and subsequently in impaired motility (Lassak et al., 2013). On the other hand the FHA-containing domain ArnA and the vWA domain carrying ArnB, have been characterized as repressors of the archaella biosynthesis in *S. acidocaldarius* (Reimann et al., 2012). It was shown that both have strong affinities towards each other and control the archaellum expression in a phosphorylation dependent manner (Reimann et al., 2012).

Fig.2. Current assembly model of the archaellum of *S. acidocaldarius*: A) FlaJ represents the core of the archaellum structure. It incorporation in the membrane is the starting point of the archaellum assembly process; B) FlaJ recruits the hexameric motor ATPase FlaI into the membrane; C) FlaH is interacting with FlaI and gets associated into the FlaHIJ core complex; D) FlaX is forming membrane bound ring-like structures surrounding the FlaHIJ core complex; E) FlaF and FlaG are forming a pore within the S-layer, enabling the further filament assembly. The FlaFG pore interacts with the FlaX rings forming a stator complex, which is stably fixed in the cell envelope due to FlaF direct binging to the glycan trees of the S-layer; F) The glycosylated and processed by PibD archaellins are incorporated on the bottom of the filament, coursing its elongation; G) After the archaellum reaches its mature size, FlaH is switching the FlaI activity from assembly into torque generation. The autophosphorylation of FlaH causes a conformational changes leading to release of FlaI, enabling its free rotation.


A
S-layer
FlaB
PibD
FlaJ
B
S-layer
FlaB
PibD
FlaJ
FlaI
C
S-layer
FlaB
PibD
FlaJ
FlaI
FlaH
FlaB
D
S-layer
FlaB
PibD
FlaJ
FlaI
E
S-layer
FlaB
PibD
FlaJ
FlaI
F
S-layer
FlaB
PibD
FlaJ
FlaI
G
S-layer
FlaB
PibD
FlaJ
FlaI
S-layer
FlaB
PibD
FlaJ
FlaI
FlaJ
FlaI
FlaH
FlaB
P

The exact regulation mechanism is however not fully understood yet. ArnR and ArnR1 are thought to be one-component systems, which can recognize the *flaB* promoter region and activate its expression in response to diverse stimuli (Lassak et al., 2013). The specific binding to the *flaB* promoter could be shown, but whether ArnR/R1 act as self-sufficient sensor is not so undeniable. The fact that ArnR itself express in response to starvation, preceding the FlaB expression implies already the need of further, hierarchically higher regulators. ArnR1 is expressed in a constant rate independent from the starvation stress (Lassak et al., 2013). However its role as ArnR expression activator can be excluded, since the *arnR1* deletion only slightly reduced the FlaB and did not abolished the motility completely (Lassak et al., 2013). Moreover ArnR1 is present exclusively in *S. acidocaldarius*, what further suggest and outlier role in the archaellum expression activation. It was demonstrated that ArnR and ArnR1 are located in the membrane, what may suggest that both these proteins are coupled with the same or different membrane associated sensory mechanisms.

It seems likely that ArnR is expressed in response to an already sensed stimuli and acts as a link between this detection system and the transcription machinery. So far it is also not clear how the membrane located proteins can activate the *flaB* transcription; however a cleavage mechanism appears here more likely than DNA recruitment. Thus it would be an interesting aspect to test whether an overexpression of ArnR/R1 lacking the transmembrane domains may directly induce the archaellum biosynthesis in *S. acidocaldarius*.

The ArnA and ArnB are repressors, since the deletion of corresponding genes resulted in increased accumulation of Fla proteins. Whereby the amount of FlaB was only slightly increased, FlaX however was significantly overproduced (Reimann et al., 2012). Moreover the *arnA* and *arnB* deletion strains revealed to be hyperarchaellated and hypermotile, resembling phenotypically the previous described Δ*aapF* strain. It was shown in *S. tokodai* that ArnA can in its unphosphorylated state directly bind to the *flaX* promoter (Duan & He, 2011). Therefore previous models of *S. acidocaldarius* archaealla regulation assumed a constant repression of the *flaX* promoter due to ArnA binding. This affinity of ArnA or ArnB towards the *flaX* promoter sequence could not be confirmed so far for *S. acidocaldarius.* However it was clearly shown, that the archaellum assembly must be initiated by the activator proteins. Thus only in presence of both ArnR and ArnR1 the archaellum of *S. acidocaldarius* can be fully induced, what would eliminate the need of constant *flaX-J* repression. It is more likely that the repressor proteins ArnA and ArnB are gradually silencing the transcription of *flaX-J* during the activated archaellum biosynthesis focusing it on the filament assembly. It was shown with q-PCR analysis that after starvation the *flaB* transcription level is significantly higher than these of *flaX-J* (Chapter 3.1), what we explained as different induction response of *flaB* and *flaX* promoter (Lassak et al., 2012b). However considering the recent data, the difference in the analyzed expression level might be because of the earlier down regulation of FlaX-J expression compared to that of FlaB. This would be in line with the promoter assay results, where no direct activation of the *flaX* promoter upon starvation could be observed (Lassak et al., 2012b). Thus it might be that ArnA/B and ArnR/R1 are indeed part of one highly complex regulatory pathway, but do not act at the same hierarchical level, since ArnA/B are most probably activated as response to the ArnR/R1 induced archaella biosynthesis. However it is still unknown how both the archaella regulator pairs are coupled together. Moreover it appears that the so far discovered four Arn factors represent only a small part of the entire archaellum regulatory pathway of *Sulfolobus*, what makes it to an extremely interesting and promising field for future study.

Conclusions:

In only few years we were able to make a significant progress in the understanding of Archaeal motility. In that time, we showed that the archaeal movement is mediated by a unique molecular machine, the Archaellum. We confirmed the functional analogies to the bacterial flagellum, like a similar rotary mode of action as also the structural homologies to the bacterial type IV pili including the pre-archaellins processing or similar motor proteins. We were able to elucidate the composition of the crenarchaeal archaellum basal body and made first successful attempts to show the universal role of FlaHIJ as archaella core structure in Archaea in general. We demonstrated the dual role of the motor protein FlaI and delivered structural evidence of its involvements in assembly and torque generation. Thereby we pointed the basic differences to other members of the secretion ATPase superfamily. We characterized FlaH and presented first indication on its role.

Nevertheless the topic of the Archaeallum is still very far from being exhausted and despite the recently obtained results allowed us to understand the principles of this motility system, its investigation process raised more questions than gave answers. A detailed description of the molecular architecture of the archaellum, including defined stoichiometry and a CryoEM tomography based three dimensional visualization, is still missing. Also the conserved components FlaF and FlaG are still very enigmatic. Preliminary results suggest a role in the archaellum elongation; however the available data is still not enough to conclusively link these proteins with any function. An interesting a highly important aspect is the differences between the crenarchaeal and euryarchaeal archaellum system. Thus, the nearest future attempts should be performed to elucidate the role of FlaCDE and investigate their position in the archaella architecture. Also the regulatory network seems to be an interesting subject, since it appears to be much differentiated and subkingdom specific.

In general, understanding how Archaea moves appears to be an intriguing and challenging biological puzzle, which could bring light on the evolution and the adaptation of T4P related systems. It seems highly fascinating how Archaea managed to compress the function of the highly complex flagellum, which requires in bacteria dozens of components into a strictly compressed form resembling the type IV pilus. Hence the archaella of *S.acidocaldarius* contains only seven different components and it represents most probably the simplest biological nanomachine, which can be in the future a useful target for applied Nanotechnology.

5. Summary

The archaeal motility structure, the archaellum is an intriguing hybrid of the function and architecture of two distinct motility organelles, the bacterial flagellum and the T4P, respectively. This rotating T4P is an astonishing example of evolutionary adaptation and represents indeed a unique, third way to move. This microbial structure was however for long time ignored and while many bacterial structures have been already well characterized, the knowledge about the archaellum remains still scare. The so far performed studies were restricted to motility in Euryarchaeota and included physiological and genetic analyses of few species. Here we present a detailed systematic and structural analysis of the crenarchaeal archaellum using the thermoacidophile *Sulfolobus acidocaldarius* as model organism.

S. acidocaldarius has the most minimalistic known archaellum system, composed of only seven Fla proteins. In-frame deletion strain analysis revealed all seven *fla* genes to be essential for proper archaellum assembly. All these mutants were non-motile, conclusively linking the archaellum of Crenarchaeota with their swimming motility. Moreover, using immunoblot analysis we found the archaella biosynthesis to be induced under nutrient depleting conditions. We could also demonstrate that despite that all the seven *fla* genes are clustered in one genomic locus, they are expressed in two different transcriptional units. Thus the archaellin FlaB and the remaining structural components FlaXHGFHIJ encoding genes are expressed separately.

The main focus of this work was however the structural aspect of the *S. acidocaldarius* archaellum. Thus we are presenting here a detailed biochemical and structural characterization of the two cytosolic components of the archaellum: the RecA family protein FlaH and the ATPase FlaI. By elucidating the interaction network of FlaH and FlaI within the archaellum, we could place them together with FlaX and FlaJ as structural components of the basal body

The ATPase FlaI was successfully crystallized in hexameric form and we could solve this structure at 2.0 Å resolution. FlaI hexamer forms a crown-like structure with subunits at three different conformational states, assembled together in a rare cross-subunit interacting fashion. Further analysis revealed also that the enzymatic activity and system specificity of FlaI are structurally separated, since the ATPase is restricted to the C-terminal domain, while the functional part is represented by the N-terminal domain. We demonstrated moreover that FlaI has a dual role and is involved in generating the energy necessary for both, the archaellum assembly and its rotation. The functions of FlaI could be uncoupled by deleting the first 29 amino acids of the N-terminus, resulting in archaellated, but not motile phenotype.

FlaH was characterized as an ATP-binding protein, since no ATPase activity could be detected. It has a well conserved Walker A, but an incomplete Walker B motif and as we could show with *in vivo* and *in vitro* analysis both motifs are important for ATP binding and also were essential for archaella assembly and motility. The structure of FlaH was solved at 2.3 Å resolution, revealing the presence of a bound ATP molecule, supporting the hypothesis that FlaH does not hydrolyze ATP. Structural similarities to the CII domain of KaiC and a proved auto-phosphorylation activity, suggest that FlaH plays a regulatory role and controls the archaellum assembly/function in a phosphorylation dependent manner.

Taking together, all the presented here data provide insights into the role of the archaellum of *S. acidocaldarius*, its genomic organization and unique molecular architecture. Furthermore our structural analysis revealed differences between the motor proteins within the archaellum and the related bacterial systems, elucidating the phenomenon of the rotating type IV pilus. However, many questions regarding the archaellum remain still open and present a challenge for further motility studies in Archaea.

5. Zusammenfassung

Die archaeale Fortbewegungstruktur, das Archaellum ist ein faszinierendes Hybrid, welches die Funktion einer bakteriellen Flagelle aber die Architektur eines Type IV Pilus (T4P) in sich vereint. Aufgrund seiner rotierenden Eigenschaft ist dieser T4P ein erstaunliches Beispiel der evolutionären Anpassung und stellt eine neue bisher unbekannte Art sich zu bewegen dar. Trotz dem Wissen über die Beweglichkeit einzelner Archaeen von über 20 Jahren, konnte durch das Fehlen genetischer Methoden in Archaeen die genaue Struktur dieser Fortbewegungseinheit sowie die zugehörigen Gene nicht identifiziert werden.. Alle bisher durchgeführten Studien zur physiologischen und genetischen Analyse des Archaellums sind auf wenigen Euryarchaeota Arten beschränkt. In dieser Studie präsentieren wir eine detaillierte, systematische sowie strukturelle Analyse des crenarchaealen Archaellums aus dem thermoacidophilen *Sulfolobus acidocaldarius*.

Das Archaellum von *S. acidocaldarius* besteht aus nur sieben elementaren Untereinheiten und repräsentiert damit das minimalistischste der bekannten Archaella Systeme. Deletionsanalysen ergaben, dass alle der sieben *fla*-Gene essentiell für die ordnungsgemäße Assemblierung des Archaellums sind. Ferner waren alle diese *fla*-Mutanten unbeweglich, wodurch die Funktion der Fla- Proteine für die Motilität bewiesen wurde. Darüber hinaus konnte mit Immunoblot Analyse gezeigt werden, dass die Biosynthese des Archaellum unter nährstofflimitierten Bedingungen induziert wird. Obwohl sich alle sieben *fla*-Gene hintereinander in einem genomischen Locus befinden, konnte gezeigt werden, dass zwei verschiedenen Transkriptionseinheiten exprimiert werden. So werden das Archaellin FlaB und die restlichen Komponenten FlaBXHGFHIJ separat synthetisiert.

Der Schwerpunkt dieser Arbeit fokussierte sich auf die strukturelle Analyse des Archaellums von *S. acidocaldarius*. So wurden detaillierte biochemische und strukturelle Charakterisierungen der beiden cytoplasmatischen Komponenten des Archaellums, die RecA Proteine FlaH und der ATPase FlaI durchgeführt. Durch die Aufklärung des Interaktionsnetzwerks von FlaH, FlaXund FlaI, konnten wir diese Proteine als strukturelle Komponenten des basalen Körper zusammen mit FlaJ identifizieren. Aufgrund dieser biochemischen Analysen konnte ein Model der Archaellum Assemblierung entwickelt werden, in dem das Transmembranprotein FlaJ als Montageplattform dient, welches die ATPase FlaI zur Membran rekrutiert. Dieser FlaI/FalJ Komplex wird von FlaX in Form einer riesigen Ringstrukturen umgeben, welche direkt mit FlaI interagiert. Aufgrund von Affinitätsstudien von FlaH mit FlaX und FlaI konnte die Bindung zu den beiden Strukturproteinen nachgewiesen werden..

Aufgereinigte ATPase FlaI konnte erfolgreich kristallisiert werden, wodurch seine hexamere Struktur mit einer Auflösung von 2,0 Å aufgeklärt werden konnte. Die Hexamere von FlaI bilden eine kronenartige Struktur, dessen Untereinheiten sich in drei verschiedenen Konformationen befinden können und in einer seltenen Queruntereinheit Interactions Weise zusammengebaut sind. Durch Deletionsstudien konnte weiter gezeigt werden, dass FlaI die erzeugte Energie sowohl für die Assemblierung des Archaellums so wie auch für dessen Rotation bereitstellt. Durch die Deletion der ersten 29 Aminosäuren des N-Terminus von FlaI konnte das Filament noch zusammengebaut, allerdings nicht mehr rotiert werden.

In vivo und *in vitro* Analyse von FlaH zeigten, dass FlaH ein gut konserviertes Walker A allerdings nur ein unvollständiges Walker-B-Motiv hat, beide Motive sind wichtig für die ATP-Bindung und Hydrolyse. Da keine ATPase-Aktivität festgestellt werden konnte, wurde FlaH als ein ATP-bindendes Protein gekennzeichnet

Die Kristallstruktur von FlaH konnte mit 2,3 Å auflöst werden. Diese Kristallstruktur enthielt ein gebundenes ATP-Molekül, wodurch zusätzlich bestätigt wurde, dass FlaH ATP nicht hydrolysieren kann. Strukturelle Ähnlichkeiten mit der CII-Domäne von KaiC suggerierten, dass

FlaH ein Phosphorylierung abhängiger Regulator für die Assemblierung und Funktion des Archaellums sein kann. Dieser Verdacht wird weiter durch die Auto-Phosphorylierungs Aktivität von FlaH verstärkt.

Zusammenfassend, geben die in dieser Arbeit präsentierten Ergebnisse neue Einblicke in die genomische Organisation und einzigartige molekulare Architektur des Archaellums von *S. acidocaldarius*. Außerdem konnten durch die hier dargestellte Strukturanalysen, Unterschiede zwischen den Motorproteinen des Archaellums und des bakteriellen Flagellums, aufgezeigt werden.

6. References

Abendroth, J., D.D. Mitchell, K.V. Korotkov, T.L. Johnson, A. Kreger, M. Sandkvist & W.G. Hol, (2009) The three-dimensional structure of the cytoplasmic domains of EpsF from the type 2 secretion system of Vibrio cholerae. *J Struct Biol* **166**: 303-315.

Adams, P.D., P.V. Afonine, G. Bunkoczi, V.B. Chen, I.W. Davis, N. Echols, J.J. Headd, L.W. Hung, G.J. Kapral, R.W. Grosse-Kunstleve, A.J. McCoy, N.W. Moriarty, R. Oeffner, R.J. Read, D.C. Richardson, J.S. Richardson, T.C. Terwilliger & P.H. Zwart, (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D* **66**: 213-221.

Ajon, M., S. Frols, M. van Wolferen, K. Stoecker, D. Teichmann, A.J. Driessen, D.W. Grogan, S.V. Albers & C. Schleper, (2011) UV-inducible DNA exchange in hyperthermophilic archaea mediated by type IV pili. *Molecular microbiology* **82**: 807-817.

Albers, S.V. & A.J. Driessen, (2008) Conditions for gene disruption by homologous recombination of exogenous DNA into the Sulfolobus solfataricus genome. *Archaea* **2**: 145-149.

Albers, S.V. & B.H. Meyer, (2011) The archaeal cell envelope. *Nature Reviews Microbiology* **9**: 414-426.

Albers, S.V., Z. Szabo & A.J. Driessen, (2003) Archaeal homolog of bacterial type IV prepilin signal peptidases with broad substrate specificity. *J Bacteriol* **185**: 3918-3925.

Aldridge, P. & K.T. Hughes, (2002) Regulation of flagellar assembly. *Current opinion in microbiology* **5**: 160-165.

Allers, T. & M. Mevarech, (2005) Archaeal genetics - The third way. *Nat Rev Genet* **6**: 58-73.

Arbing, M.A., S. Chan, A. Shin, T. Phan, C.J. Ahn, L. Rohlin & R.P. Gunsalus, (2012) Structure of the surface layer of the methanogenic archaean Methanosarcina acetivorans. *Proceedings of the National Academy of Sciences of the United States of America* **109**: 11812-11817.

Arts, J., A. de Groot, G. Ball, E. Durand, M. El Khattabi, A. Filloux, J. Tommassen & M. Koster, (2007) Interaction domains in the Pseudomonas aeruginosa type II secretory apparatus component XcpS (GspF). *Microbiol-Sgm* **153**: 1582-1592.

Banerjee, A., A. Ghosh, D.J. Mills, J. Kahnt, J. Vonck & S.V. Albers, (2012a) FlaX, a unique component of the crenarchaeal archaellum, forms oligomeric ring-shaped structures and interacts with the motor ATPase FlaI. *The Journal of biological chemistry* **287**: 43322-43330.

Banerjee, A., A. Ghosh, D.J. Mills, J. Kahnt, J. Vonck & S.V. Albers, (2012b) FlaX, a unique component of the crenarchaeal archaellum, forms oligomeric ring-shaped structures and interacts with the motor ATPase FlaI. *J Biol Chem*.

Banerjee, A., T. Neiner, P. Tripp & S.V. Albers, (2013) Insights into subunit interactions in the Sulfolobus acidocaldarius archaellum cytoplasmic complex. *The FEBS journal* **280**: 6141-6149.

Bardy, S.L. & K.F. Jarrell, (2002) FlaK of the archaeon Methanococcus maripaludis possesses preflagellin peptidase activity. *FEMS microbiology letters* **208**: 53-59.

Bardy, S.L. & K.F. Jarrell, (2003) Cleavage of preflagellins by an aspartic acid signal peptidase is essential for flagellation in the archaeon Methanococcus voltae. *Molecular microbiology* **50**: 1339-1347.

Bardy, S.L., T. Mori, K. Komoriya, S.I. Aizawa & K.F. Jarrell, (2002) Identification and localization of flagellins FlaA and FlaB3 within flagella of Methanococcus voltae. *Journal of bacteriology* **184**: 5223-5233.

Bardy, S.L., S.Y.M. Ng & K.F. Jarrell, (2003) Prokaryotic motility structures. *Microbiol-Sgm* **149**: 295-304.

Bell, S.D. & S.P. Jackson, (1998) Transcription and translation in Archaea: A mosaic of eukaryal and bacterial features. *Trends in microbiology* **6**: 222-228.

Bell, S.D. & S.P. Jackson, (2001) Mechanism and regulation of transcription in archaea. *Current opinion in microbiology* **4**: 208-213.

Berg, H.C. & R.A. Anderson, (1973) Bacteria swim by rotating their flagellar filaments. *Nature* **245**: 380-382.

Berkner, S., D. Grogan, S.V. Albers & G. Lipps, (2007) Small multicopy, non-integrative shuttle vectors based on the plasmid pRN1 for Sulfolobus acidocaldarius and Sulfolobus solfataricus, model organisms of the (cren-)archaea. *Nucleic acids research* **35**: e88.

Berkner, S., A. Wlodkowski, S.V. Albers & G. Lipps, (2010) Inducible and constitutive promoters for genetic systems in Sulfolobus acidocaldarius. *Extremophiles* **14**: 249-259.

Blair, D.F., (2006) Fine structure of a fine machine. *Journal of bacteriology* **188**: 7033-7035.

Bren, A. & M. Eisenbach, (1998) The N terminus of the flagellar switch protein, FliM, is the binding domain for the chemotactic response regulator, CheY. *Journal of molecular biology* **278**: 507-514.

Brochier-Armanet, C., B. Boussau, S. Gribaldo & P. Forterre, (2008) Mesophilic crenarchaeota: proposal for a third archaeal phylum, the Thaumarchaeota. *Nature Reviews Microbiology* **6**: 245-252.

Brock, T.D., K.M. Brock, R.T. Belly & R.L. Weiss, (1972) Sulfolobus: a new genus of sulfur-oxidizing bacteria living at low pH and high temperature. *Archiv fur Mikrobiologie* **84**: 54-68.

Brugger, K., E. Torarinsson, P. Redder, L. Chen & R.A. Garrett, (2004) Shuffling of Sulfolobus genomes by autonomous and non-autonomous mobile elements. *Biochemical Society transactions* **32**: 179-183.

Buchan, D.W.A., S.M. Ward, A.E. Lobley, T.C.O. Nugent, K. Bryson & D.T. Jones, (2010) Protein annotation and modelling servers at University College London. *Nucleic Acids Research* **38**: W563-W568.

Bulyha, I., E. Hot, S. Huntley & L. Sogaard-Andersen, (2011) GTPases in bacterial cell polarity and signalling. *Current opinion in microbiology* **14**: 726-733.

Chaban, B., S.Y. Ng, M. Kanbe, I. Saltzman, G. Nimmo, S. Aizawa & K.F. Jarrell, (2007) Systematic deletion analyses of the fla genes in the flagella operon identify several genes essential for proper assembly and function of flagella in the archaeon, Methanococcus maripaludis. *Molecular microbiology* **66**: 596-609.

Chaban, B., S.Y.M. Ng & K.F. Jarrell, (2006a) Archaeal habitats - from the extreme to the ordinary. *Can J Microbiol* **52**: 73-116.

Chaban, B., S. Voisin, J. Kelly, S.M. Logan & K.F. Jarrell, (2006b) Identification of genes involved in the biosynthesis and attachment of Methanococcus voltae N-linked glycans: insight into N-linked glycosylation pathways in Archaea. *Molecular microbiology* **61**: 259-268.

Chen, I. & D. Dubnau, (2004) DNA uptake during bacterial transformation. *Nature reviews. Microbiology* **2**: 241-249.

Chen, L.M., K. Brugger, M. Skovgaard, P. Redder, Q.X. She, E. Torarinsson, B. Greve, M. Awayez, A. Zibat, H.P. Klenk & R.A. Garrett, (2005) The genome of Sulfolobus acidocaldarius, a model organism of the Crenarchaeota. *J Bacteriol* **187**: 4992-4999.

Chevance, F.F. & K.T. Hughes, (2008) Coordinating assembly of a bacterial macromolecular machine. *Nature reviews. Microbiology* **6**: 455-465.

Chubb, J.R., A. Wilkins, D.J. Wessels, D.R. Soll & R.H. Insall, (2002) Pseudopodium dynamics and rapid cell movement in Dictyostelium Ras pathway mutants. *Cell motility and the cytoskeleton* **53**: 150-162.

Classen, S., I. Rodic, J. Holton, G.L. Hura, M. Hammel & J.A. Tainer, (2010) Software for the high-throughput collection of SAXS data using an enhanced Blu-Ice/DCS control system. *Journal of synchrotron radiation* **17**: 774-781.

Clausen, M., V. Jakovljevic, L. Sogaard-Andersen & B. Maier, (2009) High-force generation is a conserved property of type IV pilus systems. *Journal of bacteriology* **191**: 4633-4638.

Cohen-Krausz, S. & S. Trachtenberg, (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. *Journal of molecular biology* **321**: 383-395.

Cohen-Krausz, S. & S. Trachtenberg, (2008) The flagellar filament structure of the extreme acidothermophile Sulfolobus shibatae B12 suggests that archaeabacterial flagella have a unique and common symmetry and design. *Journal of molecular biology* **375**: 1113-1124.

Cohen, E.J. & K.T. Hughes, (2014) Rod-to-Hook Transition for Extracellular Flagellum Assembly Is Catalyzed by the L-Ring-Dependent Rod Scaffold Removal. *Journal of bacteriology* **196**: 2387-2395.

Cole, C., J.D. Barber & G.J. Barton, (2008) The Jpred 3 secondary structure prediction server. *Nucleic Acids Research* **36**: W197-W201.

Collins, R.F., S.A. Frye, S. Balasingham, R.C. Ford, T. Tonjum & J.P. Derrick, (2005) Interaction with type IV pili induces structural changes in the bacterial outer membrane secretin PilQ. *The Journal of biological chemistry* **280**: 18923-18930.

Consortium, U., (2012) Reorganizing the protein space at the Universal Protein Resource (UniProt). *Nucleic Acids Research* **40**: D71-D75.

Craig, L., M.E. Pique & J.A. Tainer, (2004) Type IV pilus structure and bacterial pathogenicity. *Nature reviews. Microbiology* **2**: 363-378.

Craig, L., N. Volkmann, A.S. Arvai, M.E. Pique, M. Yeager, E.H. Egelman & J.A. Tainer, (2006) Type IV pilus structure by cryo-electron microscopy and crystallography: implications for pilus assembly and functions. *Molecular cell* **23**: 651-662.

Crowther, L.J., R.P. Anantha & M.S. Donnenberg, (2004) The inner membrane subassembly of the enteropathogenic Escherichia coli bundle-forming pilus machine. *Molecular microbiology* **52**: 67-79.

Dean, G.E., R.M. Macnab, J. Stader, P. Matsumura & C. Burks, (1984) Gene sequence and predicted amino acid sequence of the motA protein, a membrane-associated protein required for flagellar rotation in Escherichia coli. *J Bacteriol* **159**: 991-999.

Dennis, P.P., (1997) Ancient ciphers: Translation in Archaea. *Cell* **89**: 1007-1010.

Desmond, E., C. Brochier-Armanet & S. Gribaldo, (2007) Phylogenomics of the archaeal flagellum: rare horizontal gene transfer in a unique motility structure. *Bmc Evol Biol* **7**.

Duan, X. & Z.G. He, (2011) Characterization of the specific interaction between archaeal FHA domain-containing protein and the promoter of a flagellar-like gene-cluster and its regulation by phosphorylation. *Biochem Bioph Res Co* **407**: 242-247.

Dvornyk, V., O. Vinogradova & E. Nevo, (2003) Origin and evolution of circadian clock genes in prokaryotes. *Proceedings of the National Academy of Sciences of the United States of America* **100**: 2495-2500.

Edgar, R.S., E.W. Green, Y. Zhao, G. van Ooijen, M. Olmedo, X. Qin, Y. Xu, M. Pan, U.K. Valekunja, K.A. Feeney, E.S. Maywood, M.H. Hastings, N.S. Baliga, M. Merrow, A.J. Millar, C.H. Johnson, C.P. Kyriacou, J.S. O'Neill & A.B. Reddy, (2012) Peroxiredoxins are conserved markers of circadian rhythms. *Nature* **485**: 459-464.

Egli, M., T. Mori, R. Pattanayek, Y. Xu, X. Qin & C.H. Johnson, (2012) Dephosphorylation of the core clock protein KaiC in the cyanobacterial KaiABC circadian oscillator proceeds via an ATP synthase mechanism. *Biochemistry* **51**: 1547-1558.

Egli, M., R. Pattanayek, J.H. Sheehan, Y. Xu, T. Mori, J.A. Smith & C.H. Johnson, (2013) Loop-loop interactions regulate KaiA-stimulated KaiC phosphorylation in the cyanobacterial KaiABC circadian clock. *Biochemistry* **52**: 1208-1220.

Elkins, J.G., M. Podar, D.E. Graham, K.S. Makarova, Y. Wolf, L. Randau, B.P. Hedlund, C. Brochier-Armanet, V. Kunin, I. Anderson, A. Lapidus, E. Goltsman, K. Barry, E.V. Koonin, P. Hugenholtz, N. Kyrpides, G. Wanner, P. Richardson, M. Keller & K.O. Stetter, (2008) A korarchaeal genome reveals insights into the evolution of the Archaea. *Proceedings of the National Academy of Sciences of the United States of America* **105**: 8102-8107.

Emsley, P. & K. Cowtan, (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr D* **60**: 2126-2132.

Evans, L.D., C. Hughes & G.M. Fraser, (2014) Building a flagellum outside the bacterial cell. *Trends in microbiology*.

Evans, L.D., S. Poulter, E.M. Terentjev, C. Hughes & G.M. Fraser, (2013) A chain mechanism for flagellum growth. *Nature* **504**: 287-290.

Faguy, D.M., S.F. Koval & K.F. Jarrell, (1994) Physical Characterization of the Flagella and Flagellins from Methanospirillum-Hungatei. *Journal of bacteriology* **176**: 7491-7498.

Fiala, G. & K.O. Stetter, (1986) Pyrococcus-Furiosus Sp-Nov Represents a Novel Genus of Marine Heterotrophic Archaebacteria Growing Optimally at 100-Degrees C. *Archives of microbiology* **145**: 56-61.

Fox, G.E., L.J. Magrum, W.E. Balch, R.S. Wolfe & C.R. Woese, (1977) Classification of methanogenic bacteria by 16S ribosomal RNA characterization. *Proceedings of the National Academy of Sciences of the United States of America* **74**: 4537-4541.

Frols, S., M. Ajon, M. Wagner, D. Teichmann, B. Zolghadr, M. Folea, E.J. Boekema, A.J. Driessen, C. Schleper & S.V. Albers, (2008) UV-inducible cellular aggregation of the hyperthermophilic archaeon Sulfolobus solfataricus is mediated by pili formation. *Molecular microbiology* **70**: 938-952.

Fuchs, T., H. Huber, S. Burggraf & K.O. Stetter, (1996) 16S rDNA-based phylogeny of the archaeal order Sulfolobales and reclassification of Desulfurolobus ambivalens as Acidianus ambivalens comb nov. *Syst Appl Microbiol* **19**: 56-60.

Ghosh, A. & S.V. Albers, (2011a) Assembly and function of the archaeal flagellum. *Biochemical Society transactions* **39**: 64-69.

Ghosh, A. & S.V. Albers, (2011b) Assembly and function of the archaeal flagellum. *Biochemical Society Transactions* **39**: 64-69.

Ghosh, A., S. Hartung, C. van der Does, J.A. Tainer & S.V. Albers, (2011a) Archaeal flagellar ATPase motor shows ATP-dependent hexameric assembly and activity stimulation by specific lipid binding. *The Biochemical journal* **437**: 43-52.

Ghosh, A., S. Hartung, C. van der Does, J.A. Tainer & S.V. Albers, (2011b) Archaeal flagellar ATPase motor shows ATP-dependent hexameric assembly and activity stimulation by specific lipid binding. *Biochem J* **437**: 43-52.

Giltner, C.L., Y. Nguyen & L.L. Burrows, (2012) Type IV pilin proteins: versatile molecular modules. *Microbiology and molecular biology reviews : MMBR* **76**: 740-772.

Graham, D.E., R. Overbeek, G.J. Olsen & C.R. Woese, (2000) An archaeal genomic signature. *Proceedings of the National Academy of Sciences of the United States of America* **97**: 3304-3308.

Grogan, D.W., (1989) Phenotypic characterization of the archaebacterial genus Sulfolobus: comparison of five wild-type strains. *J Bacteriol* **171**: 6710-6719.

Grogan, D.W., G.T. Carver & J.W. Drake, (2001) Genetic fidelity under harsh conditions: analysis of spontaneous mutation in the thermoacidophilic archaeon Sulfolobus acidocaldarius. *Proceedings of the National Academy of Sciences of the United States of America* **98**: 7928-7933.

Grogan, D.W. & J.E. Hansen, (2003) Molecular characteristics of spontaneous deletions in the hyperthermophilic archaeon Sulfolobus acidocaldarius. *J Bacteriol* **185**: 1266-1272.

Guttenplan, S.B., S. Shaw & D.B. Kearns, (2013) The cell biology of peritrichous flagella in Bacillus subtilis. *Molecular microbiology* **87**: 211-229.

Hagman, K.E., S.F. Porcella, T.G. Popova & M.V. Norgard, (1997) Evidence for a methyl-accepting chemotaxis protein gene (mcp1) that encodes a putative sensory transducer in virulent Treponema pallidum. *Infection and immunity* **65**: 1701-1709.

Haldenby, S., M.F. White & T. Allers, (2009) RecA family proteins in archaea: RadA and its cousins. *Biochemical Society transactions* **37**: 102-107.

Hanson, P.I. & S.W. Whiteheart, (2005) AAA+ proteins: Have engine, will work. *Nat Rev Mol Cell Bio* **6**: 519-529.

Hayashi, F., N. Itoh, T. Uzumaki, R. Iwase, Y. Tsuchiya, H. Yamakawa, M. Morishita, K. Onai, S. Itoh & M. Ishiura, (2004) Roles of two ATPase-motif-containing domains in cyanobacterial circadian clock protein KaiC. *The Journal of biological chemistry* **279**: 52331-52337.

Hayashi, F., H. Suzuki, R. Iwase, T. Uzumaki, A. Miyake, J.R. Shen, K. Imada, Y. Furukawa, K. Yonekura, K. Namba & M. Ishiura, (2003) ATP-induced hexameric ring structure of the cyanobacterial circadian clock protein KaiC. *Genes to cells : devoted to molecular & cellular mechanisms* **8**: 287-296.

Henche, A.L., A. Ghosh, X. Yu, T. Jeske, E. Egelman & S.V. Albers, (2012a) Structure and function of the adhesive type IV pilus of Sulfolobus acidocaldarius. *Environ Microbiol* **14**: 3188-3202.

Henche, A.L., A. Koerdt, A. Ghosh & S.V. Albers, (2012b) Influence of cell surface structures on crenarchaeal biofilm formation using a thermostable green fluorescent protein. *Environ Microbiol* **14**: 779-793.

Henche, A.L., M. van Wolferen, A. Ghosh & S.V. Albers, (2014) Dissection of key determinants of cleavage activity in signal peptidase III (SPaseIII) PibD. *Extremophiles* **18**: 905-913.

Hendrickson, E.L., Y. Liu, G. Rosas-Sandoval, I. Porat, D. Soll, W.B. Whitman & J.A. Leigh, (2008) Global responses of Methanococcus maripaludis to specific nutrient limitations and growth rate. *Journal of bacteriology* **190**: 2198-2205.

Holm, L. & P. Rosenstrom, (2010) Dali server: conservation mapping in 3D. *Nucleic acids research* **38**: W545-549.

Huber, H., M.J. Hohn, R. Rachel, T. Fuchs, V.C. Wimmer & K.O. Stetter, (2002) A new phylum of Archaea represented by a nanosized hyperthermophilic symbiont. *Nature* **417**: 63-67.

Huet, J., R. Schnabel, A. Sentenac & W. Zillig, (1983) Archaebacteria and Eukaryotes Possess DNA-Dependent Rna-Polymerases of a Common Type. *Embo Journal* **2**: 1291-1294.

Hura, G.L., A.L. Menon, M. Hammel, R.P. Rambo, F.L. Poole, S.E. Tsutakawa, F.E. Jenney, S. Classen, K.A. Frankel, R.C. Hopkins, S.J. Yang, J.W. Scott, B.D. Dillard, M.W.W. Adams & J.A. Tainer, (2009) Robust, high-throughput solution structural analyses by small angle X-ray scattering (SAXS). *Nat Methods* **6**: 606-U683.

Ito, H., H. Kageyama, M. Mutsuda, M. Nakajima, T. Oyama & T. Kondo, (2007) Autonomous synchronization of the circadian KaiC phosphorylation rhythm. *Nature structural & molecular biology* **14**: 1084-1088.

Ito, H., M. Mutsuda, Y. Murayama, J. Tomita, N. Hosokawa, K. Terauchi, C. Sugita, M. Sugita, T. Kondo & H. Iwasaki, (2009) Cyanobacterial daily life with Kai-based circadian and diurnal genome-wide transcriptional control in Synechococcus elongatus. *Proceedings of the National Academy of Sciences of the United States of America* **106**: 14168-14173.

Jakovljevic, V., S. Leonardy, M. Hoppert & L. Sogaard-Andersen, (2008) PilB and PilT are ATPases acting antagonistically in type IV pilus function in Myxococcus xanthus. *Journal of bacteriology* **190**: 2411-2421.

Jarrell, K.F. & S.V. Albers, (2012) The archaellum: an old motility structure with a new name. *Trends in microbiology* **20**: 307-312.

Jarrell, K.F., G.M. Jones & D.B. Nair, (2010) Biosynthesis and role of N-linked glycosylation in cell surface structures of archaea with a focus on flagella and s layers. *International journal of microbiology* **2010**: 470138.

Jarrell, K.F. & M.J. McBride, (2008a) The surprisingly diverse ways that prokaryotes move. *Nature Reviews Microbiology* **6**: 466-476.

Jarrell, K.F. & M.J. McBride, (2008b) The surprisingly diverse ways that prokaryotes move. *Nature reviews. Microbiology* **6**: 466-476.

Jonuscheit, M., E. Martusewitsch, K.M. Stedman & C. Schleper, (2003) A reporter gene system for the hyperthermophilic archaeon Sulfolobus solfataricus based on a selectable and integrative shuttle vector. *Mol Microbiol* **48**: 1241-1252.

Kageyama, H., T. Nishiwaki, M. Nakajima, H. Iwasaki, T. Oyama & T. Kondo, (2006) Cyanobacterial circadian pacemaker: Kai protein complex dynamics in the KaiC phosphorylation cycle in vitro. *Molecular cell* **23**: 161-171.

Kaiser, D., (2007) Bacterial swarming: a re-examination of cell-movement patterns. *Curr Biol* **17**: R561-570.

Kalmokoff, M.L., K.F. Jarrell & S.F. Koval, (1988) Isolation of flagella from the archaebacterium Methanococcus voltae by phase separation with Triton X-114. *Journal of bacteriology* **170**: 1752-1758.

Kandler, O. & H. Hippe, (1977) Lack of Peptidoglycan in Cell-Walls of Methanosarcina-Barkeri. *Archives of microbiology* **113**: 57-60.

Kandler, O. & H. Konig, (1978) Chemical Composition of Peptidoglycan-Free Cell-Walls of Methanogenic Bacteria. *Archives of microbiology* **118**: 141-152.

Kandler, O. & H. Konig, (1998) Cell wall polymers in Archaea (Archaebacteria). *Cell Mol Life Sci* **54**: 305-308.

Karuppiah, V., D. Hassan, M. Saleem & J.P. Derrick, (2010) Structure and oligomerization of the PilC type IV pilus biogenesis protein from Thermus thermophilus. *Proteins* **78**: 2049-2057.

Kates, M., (1992) Archaebacterial lipids: structure, biosynthesis and function. *Biochemical Society symposium* **58**: 51-72.

Kazmierczak, B.I. & D.R. Hendrixson, (2013) Spatial and numerical regulation of flagellar biosynthesis in polarly flagellated bacteria. *Molecular microbiology* **88**: 655-663.

Kearns, D.B., (2010) A field guide to bacterial swarming motility. *Nature reviews. Microbiology* **8**: 634-644.

Ken F. Jarrell, D.J.V.a.J.W., (2009) *Archaeal Flagella and Pili*. Caister Academic Press, UK.

Khan, S., M. Dapice & T.S. Reese, (1988) Effects of mot gene expression on the structure of the flagellar motor. *J Mol Biol* **202**: 575-584.

Kim, E.A., M. Price-Carter, W.C. Carlquist & D.F. Blair, (2008) Membrane segment organization in the stator complex of the flagellar motor: implications for proton flow and proton-induced conformational change. *Biochemistry* **47**: 11332-11339.

Kletzin, A., (2007) General Characteristics and Important Model Organisms. *Archaea: Molecular and Cellular Biology*: 14-+.

Kojima, S. & D.F. Blair, (2001) Conformational change in the stator of the bacterial flagellar motor. *Biochemistry* **40**: 13041-13050.

Kojima, S., K. Imada, M. Sakuma, Y. Sudo, C. Kojima, T. Minamino, M. Homma & K. Namba, (2009) Stator assembly and activation mechanism of the flagellar motor by the periplasmic region of MotB. *Molecular microbiology* **73**: 710-718.

Konarev, P.V., V.V. Volkov, A.V. Sokolova, M.H.J. Koch & D.I. Svergun, (2003) PRIMUS: a Windows PC-based system for small-angle scattering data analysis. *Journal of applied crystallography* **36**: 1277-1282.

Kupper, J., W. Marwan, D. Typke, H. Grunberg, U. Uwer, M. Gluch & D. Oesterhelt, (1994) The Flagellar Bundle of Halobacterium-Salinarium Is Inserted into a Distinct Polar-Cap Structure. *Journal of bacteriology* **176**: 5184-5187.

Larkin, M.A., G. Blackshields, N.P. Brown, R. Chenna, P.A. McGettigan, H. McWilliam, F. Valentin, I.M. Wallace, A. Wilm, R. Lopez, J.D. Thompson, T.J. Gibson & D.G. Higgins, (2007) Clustal W and clustal X version 2.0. *Bioinformatics* **23**: 2947-2948.

Lassak, K., A. Ghosh & S.V. Albers, (2012a) Diversity, assembly and regulation of archaeal type IV pili-like and non-type-IV pili-like surface structures. *Research in microbiology* **163**: 630-644.

Lassak, K., T. Neiner, A. Ghosh, A. Klingl, R. Wirth & S.V. Albers, (2012b) Molecular analysis of the crenarchaeal flagellum. *Molecular microbiology* **83**: 110-124.

Lassak, K., E. Peeters, S. Wrobel & S.V. Albers, (2013) The one-component system ArnR: a membrane-bound activator of the crenarchaeal archaellum. *Molecular microbiology* **88**: 125-139.

Leipe, D.D., L. Aravind, N.V. Grishin & E.V. Koonin, (2000) The bacterial replicative helicase DnaB evolved from a RecA duplication. *Genome research* **10**: 5-16.

Logan, S.M., (2006) Flagellar glycosylation - a new component of the motility repertoire? *Microbiol-Sgm* **152**: 1249-1262.

Macalady, J.L., M.M. Vestling, D. Baumler, N. Boekelheide, C.W. Kaspar & J.F. Banfield, (2004) Tetraether-linked membrane monolayers in Ferroplasma spp: a key to survival in acid. *Extremophiles* **8**: 411-419.

Macnab, R.M., (2003) How bacteria assemble flagella. *Annual review of microbiology* **57**: 77-100.

Magariyama, Y., S. Sugiyama, K. Muramoto, Y. Maekawa, I. Kawagishi, Y. Imae & S. Kudo, (1994) Very fast flagellar rotation. *Nature* **371**: 752.

Maniscalco, M., J. Nannen, V. Sodi, G. Silver, P.L. Lowrey & K.A. Bidle, (2014) Light-dependent expression of four cryptic archaeal circadian gene homologs. *Frontiers in microbiology* **5**.

Margolis, R.L. & L. Wilson, (1981) Microtubule Treadmills - Possible Molecular Machinery. *Nature* **293**: 705-711.
Mattick, J.S., (2002) Type IV pili and twitching motility. *Annual review of microbiology* **56**: 289-314.
McCarter, L.L., (2006) Regulation of flagella. *Current opinion in microbiology* **9**: 180-186.
Mcclure, M.A. & R.W.G. Wyckoff, (1982) Ultrastructural Characteristics of Sulfolobus-Acidocaldarius. *J Gen Microbiol* **128**: 433-437.
McCoy, A.J., R.W. Grosse-Kunstleve, P.D. Adams, M.D. Winn, L.C. Storoni & R.J. Read, (2007) Phaser crystallographic software. *Journal of applied crystallography* **40**: 658-674.
Melville, S. & L. Craig, (2013) Type IV pili in Gram-positive bacteria. *Microbiology and molecular biology reviews : MMBR* **77**: 323-341.
Merino, S., J.G. Shaw & J.M. Tomas, (2006) Bacterial lateral flagella: an inducible flagella system. *FEMS microbiology letters* **263**: 127-135.
Meyer, B.H., B. Zolghadr, E. Peyfoon, M. Pabst, M. Panico, H.R. Morris, S.M. Haslam, P. Messner, C. Schaffer, A. Dell & S.V. Albers, (2011) Sulfoquinovose synthase - an important enzyme in the N-glycosylation pathway of Sulfolobus acidocaldarius. *Molecular microbiology* **82**: 1150-1163.
Miyata, M., (2008) Centipede and inchworm models to explain Mycoplasma gliding. *Trends in microbiology* **16**: 6-12.
Moissl, C., R. Rachel, A. Briegel, H. Engelhardt & R. Huber, (2005) The unique structure of archaeal 'hami', highly complex cell appendages with nano-grappling hooks. *Molecular microbiology* **56**: 361-370.
Moissl, C., C. Rudolph & R. Huber, (2002) Natural communities of novel archaea and bacteria with a string-of-pearls-like morphology: molecular analysis of the bacterial partners. *Applied and environmental microbiology* **68**: 933-937.
Mukhopadhyay, B., E.F. Johnson & R.S. Wolfe, (2000) A novel p(H2) control on the expression of flagella in the hyperthermophilic strictly hydrogenotrophic methanarchaeaon Methanococcus jannaschii. *Proceedings of the National Academy of Sciences of the United States of America* **97**: 11522-11527.
Mullakhanbhai, M.F. & H. Larsen, (1975) Halobacterium volcanii spec. nov., a Dead Sea halobacterium with a moderate salt requirement. *Archives of microbiology* **104**: 207-214.
Nakajima, M., K. Imai, H. Ito, T. Nishiwaki, Y. Murayama, H. Iwasaki, T. Oyama & T. Kondo, (2005) Reconstitution of circadian oscillation of cyanobacterial KaiC phosphorylation in vitro. *Science* **308**: 414-415.
Nan, B. & D.R. Zusman, (2011) Uncovering the mystery of gliding motility in the myxobacteria. *Annual review of genetics* **45**: 21-39.
Ng, S.Y., B. Zolghadr, A.J. Driessen, S.V. Albers & K.F. Jarrell, (2008) Cell surface structures of archaea. *Journal of bacteriology* **190**: 6039-6047.
Ng, S.Y.M., B. Chaban & K.F. Jarrell, (2006) Archaeal flagella, bacterial flagella and type IV pili: A comparison of genes and posttranslational modifications. *J Mol Microb Biotech* **11**: 167-191.
Ng, S.Y.M., D.J. VanDyke, B. Chaban, J. Wu, Y. Nosaka, S.I. Aizawa & K.F. Jarrell, (2009) Different Minimal Signal Peptide Lengths Recognized by the Archaeal Prepilin-Like Peptidases FlaK and PibD. *Journal of bacteriology* **191**: 6732-6740.
Nickell, S., R. Hegerl, W. Baumeister & R. Rachel, (2003) Pyrodictium cannulae enter the periplasmic space but do not enter the cytoplasm, as revealed by cryo-electron tomography. *J Struct Biol* **141**: 34-42.
Nishiwaki, T., Y. Satomi, Y. Kitayama, K. Terauchi, R. Kiyohara, T. Takao & T. Kondo, (2007) A sequential program of dual phosphorylation of KaiC as a basis for circadian rhythm in cyanobacteria. *The EMBO journal* **26**: 4029-4037.
Nivaskumar, M., G. Bouvier, M. Campos, N. Nadeau, X. Yu, E.H. Egelman, M. Nilges & O. Francetic, (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.

O'Toole, G.A. & R. Kolter, (1998) Flagellar and twitching motility are necessary for Pseudomonas aeruginosa biofilm development. *Molecular microbiology* **30**: 295-304.
Olsen, G.J. & C.R. Woese, (1997) Archaeal genomics: An overview. *Cell* **89**: 991-994.
Otwinowski, Z. & W. Minor, (1997) Processing of X-ray diffraction data collected in oscillation mode. *Method Enzymol* **276**: 307-326.
Patenge, N., A. Berendes, H. Engelhardt, S.C. Schuster & D. Oesterhelt, (2001) The fla gene cluster is involved in the biogenesis of flagella in Halobacterium salinarum. *Molecular microbiology* **41**: 653-663.
Pattanayek, R., T. Mori, Y. Xu, S. Pattanayek, C.H. Johnson & M. Egli, (2009) Structures of KaiC circadian clock mutant proteins: a new phosphorylation site at T426 and mechanisms of kinase, ATPase and phosphatase. *PloS one* **4**: e7529.
Pattanayek, R., J. Wang, T. Mori, Y. Xu, C.H. Johnson & M. Egli, (2004) Visualizing a circadian clock protein: crystal structure of KaiC and functional insights. *Molecular cell* **15**: 375-388.
Pattanayek, R., D.R. Williams, S. Pattanayek, Y. Xu, T. Mori, C.H. Johnson, P.L. Stewart & M. Egli, (2006) Analysis of KaiA-KaiC protein interactions in the cyano-bacterial circadian clock using hybrid structural methods. *The EMBO journal* **25**: 2017-2028.
Pattanayek, R., D.R. Williams, G. Rossi, S. Weigand, T. Mori, C.H. Johnson, P.L. Stewart & M. Egli, (2011) Combined SAXS/EM based models of the S. elongatus post-translational circadian oscillator and its interactions with the output His-kinase SasA. *PloS one* **6**: e23697.
Pattanayek, R., K.K. Yadagiri, M.D. Ohi & M. Egli, (2013) Nature of KaiB-KaiC binding in the cyanobacterial circadian oscillator. *Cell cycle* **12**: 810-817.
Peabody, C.R., Y.J. Chung, M.R. Yen, D. Vidal-Ingigliardi, A.P. Pugsley & M.H. Saier, (2003a) Type II protein secretion and its relationship to bacterial type IV pili and archaeal flagella. *Microbiol-Sgm* **149**: 3051-3072.
Peabody, C.R., Y.J. Chung, M.R. Yen, D. Vidal-Ingigliardi, A.P. Pugsley & M.H. Saier, Jr., (2003b) Type II protein secretion and its relationship to bacterial type IV pili and archaeal flagella. *Microbiology* **149**: 3051-3072.
Pechman, K.J. & C.R. Woese, (1972) Characterization of the primary structural homology between the 16s ribosomal RNAs of Escherichia coli and Bacillus megaterium by oligomer cataloging. *Journal of molecular evolution* **1**: 230-240.
Pohlschroder, M., A. Ghosh, M. Tripepi & S.V. Albers, (2011) Archaeal type IV pilus-like structures--evolutionarily conserved prokaryotic surface organelles. *Current opinion in microbiology* **14**: 357-363.
Porter, J.R., (1976) Antony van Leeuwenhoek: tercentenary of his discovery of bacteria. *Bacteriological reviews* **40**: 260-269.
Qian, X.G., Y.J. He & Y. Luo, (2007) Binding of a second magnesium is required for ATPase activity of RadA from Methanococcus voltae. *Biochemistry* **46**: 5855-5863.
Reimann, J., K. Lassak, S. Khadouma, T.J.G. Ettema, N. Yang, A.J.M. Driessen, A. Klingl & S.V. Albers, (2012) Regulation of archaella expression by the FHA and von Willebrand domain-containing proteins ArnA and ArnB in Sulfolobus acidocaldarius. *Mol Microbiol* **86**: 24-36.
Reindl, S., A. Ghosh, G.J. Williams, K. Lassak, T. Neiner, A.L. Henche, S.V. Albers & J.A. Tainer, (2013) Insights into FlaI functions in archaeal motor assembly and motility from structures, conformations, and genetics. *Molecular cell* **49**: 1069-1082.
Rieger, G., R. Rachel, R. Hermann & K.O. Stetter, (1995) Ultrastructure of the Hyperthermophilic Archaeon Pyrodictium-Abyssi. *J Struct Biol* **115**: 78-87.
Rosbash, M., (2009) The implications of multiple circadian clock origins. *PLoS biology* **7**: e62.
Rudolph, C., G. Wanner & R. Huber, (2001) Natural communities of novel archaea and bacteria growing in cold sulfurous springs with a string-of-pearls-like morphology. *Applied and environmental microbiology* **67**: 2336-2344.
Rust, M.J., J.S. Markson, W.S. Lane, D.S. Fisher & E.K. O'Shea, (2007) Ordered phosphorylation governs oscillation of a three-protein circadian clock. *Science* **318**: 809-812.

Samatey, F.A., H. Matsunami, K. Imada, S. Nagashima, T.R. Shaikh, D.R. Thomas, J.Z. Chen, D.J. Derosier, A. Kitao & K. Namba, (2004) Structure of the bacterial flagellar hook and implication for the molecular universal joint mechanism. *Nature* **431**: 1062-1068.

Savvides, S.N., (2007) Secretion superfamily ATPases swing big. *Structure* **15**: 255-257.

Schlesner, M., A. Miller, S. Streif, W.F. Staudinger, J. Muller, B. Scheffer, F. Siedler & D. Oesterhelt, (2009) Identification of Archaea-specific chemotaxis proteins which interact with the flagellar apparatus. *Bmc Microbiol* **9**.

Shahapure, R., R.P. Driessen, M.F. Haurat, S.V. Albers & R.T. Dame, (2014a) The archaellum: a rotating type IV pilus. *Molecular microbiology* **91**: 716-723.

Shahapure, R., R.P.C. Driessen, M.F. Haurat, S.V. Albers & R.T. Dame, (2014b) The archaellum: a rotating type IV pilus. *Mol Microbiol* **91**: 716-723.

Sheng, D.H., S.S. Zhu, T. Wei, J.F. Ni & Y.L. Shen, (2008) The in vitro activity of a Rad55 homologue from Sulfolobus tokodaii, a candidate mediator in RadA-catalyzed homologous recombination. *Extremophiles* **12**: 147-157.

Shivvers, D.W. & T.D. Brock, (1973) Oxidation of elemental sulfur by Sulfolobus acidocaldarius. *J Bacteriol* **114**: 706-710.

Siewering, K., S. Jain, C. Friedrich, M.T. Webber-Birungi, D.A. Semchonok, I. Binzen, A. Wagner, S. Huntley, J. Kahnt, A. Klingl, E.J. Boekema, L. Sogaard-Andersen & C. van der Does, (2014) Peptidoglycan-binding protein TsaP functions in surface assembly of type IV pili. *Proceedings of the National Academy of Sciences of the United States of America* **111**: E953-961.

Silverman, M. & M. Simon, (1974) Flagellar rotation and the mechanism of bacterial motility. *Nature* **249**: 73-74.

Skerker, J.M. & H.C. Berg, (2001) Direct observation of extension and retraction of type IV pili. *Proceedings of the National Academy of Sciences of the United States of America* **98**: 6901-6904.

Southam, G., M.L. Kalmokoff, K.F. Jarrell, S.F. Koval & T.J. Beveridge, (1990a) Isolation, Characterization, and Cellular Insertion of the Flagella from 2 Strains of the Archaebacterium Methanospirillum-Hungatei. *Journal of bacteriology* **172**: 3221-3228.

Southam, G., M.L. Kalmokoff, K.F. Jarrell, S.F. Koval & T.J. Beveridge, (1990b) Isolation, characterization, and cellular insertion of the flagella from two strains of the archaebacterium Methanospirillum hungatei. *Journal of bacteriology* **172**: 3221-3228.

Spang, A., A. Poehlein, P. Offre, S. Zumbragel, S. Haider, N. Rychlik, B. Nowka, C. Schmeisser, E.V. Lebedeva, T. Rattei, C. Bohm, M. Schmid, A. Galushko, R. Hatzenpichler, T. Weinmaier, R. Daniel, C. Schleper, E. Spieck, W. Streit & M. Wagner, (2012) The genome of the ammonia-oxidizing Candidatus Nitrososphaera gargensis: insights into metabolic versatility and environmental adaptations. *Environ Microbiol* **14**: 3122-3145.

Sprott, G.D., M. Meloche & J.C. Richards, (1991) Proportions of Diether, Macrocyclic Diether, and Tetraether Lipids in Methanococcus-Jannaschii Grown at Different Temperatures. *Journal of bacteriology* **173**: 3907-3910.

Stetter, K.O., H. Konig & E. Stackebrandt, (1983) Pyrodictium Gen-Nov, a New Genus of Submarine Disc-Shaped Sulfur Reducing Archaebacteria Growing Optimally at 105-Degrees-C. *Systematic and applied microbiology* **4**: 535-551.

Streif, S., W.F. Staudinger, W. Marwan & D. Oesterhelt, (2008) Flagellar Rotation in the Archaeon Halobacterium salinarum Depends on ATP. *J Mol Biol* **384**: 1-8.

Szabo, Z., S.V. Albers & A.J.M. Driessen, (2006) Active-site residues in the type IV prepilin peptidase homologue PibD from the archaeon Sulfolobus solfataricus. *J Bacteriol* **188**: 1437-1443.

Szabo, Z., A.O. Stahl, S.V. Albers, J.C. Kissinger, A.J. Driessen & M. Pohlschroder, (2007a) Identification of diverse archaeal proteins with class III signal peptides cleaved by distinct archaeal prepilin peptidases. *Journal of bacteriology* **189**: 772-778.

Szabo, Z., A.O. Stahl, S.V. Albers, J.C. Kissinger, A.J.M. Driessen & M. Pohlschroder, (2007b) Identification of diverse archaeal proteins with class III signal peptides cleaved by distinct archaeal prepilin peptidases. *J Bacteriol* **189**: 772-778.

Takhar, H.K., K. Kemp, M. Kim, P.L. Howell & L.L. Burrows, (2013) The platform protein is essential for type IV pilus biogenesis. *The Journal of biological chemistry* **288**: 9721-9728.

Terauchi, K., Y. Kitayama, T. Nishiwaki, K. Miwa, Y. Murayama, T. Oyama & T. Kondo, (2007) ATPase activity of KaiC determines the basic timing for circadian clock of cyanobacteria. *Proceedings of the National Academy of Sciences of the United States of America* **104**: 16377-16381.

Thomas, D.R., N.R. Francis, C. Xu & D.J. DeRosier, (2006) The three-dimensional structure of the flagellar rotor from a clockwise-locked mutant of Salmonella enterica serovar Typhimurium. *J Bacteriol* **188**: 7039-7048.

Thomas, N.A., S.L. Bardy & K.F. Jarrell, (2001a) The archaeal flagellum: a different kind of prokaryotic motility structure. *FEMS microbiology reviews* **25**: 147-174.

Thomas, N.A., S. Mueller, A. Klein & K.F. Jarrell, (2002) Mutants in flaI and flaJ of the archaeon Methanococcus voltae are deficient in flagellum assembly. *Molecular microbiology* **46**: 879-887.

Thomas, N.A., C.T. Pawson & K.F. Jarrell, (2001b) Insertional inactivation of the flaH gene in the archaeon Methanococcus voltae results in non-flagellated cells. *Mol Genet Genomics* **265**: 596-603.

Toutain, C.M., M.E. Zegans & G.A. O'Toole, (2005) Evidence for two flagellar stators and their role in the motility of Pseudomonas aeruginosa. *Journal of bacteriology* **187**: 771-777.

Trachtenberg, S. & S. Cohen-Krausz, (2006) The archaeabacterial flagellar filament: a bacterial propeller with a pilus-like structure. *J Mol Microbiol Biotechnol* **11**: 208-220.

Trachtenberg, S., V.E. Galkin & E.H. Egelman, (2005) Refining the structure of the Halobacterium salinarum flagellar filament using the iterative helical real space reconstruction method: insights into polymorphism. *Journal of molecular biology* **346**: 665-676.

Tripathi, S.A. & R.K. Taylor, (2007) Membrane association and multimerization of TcpT, the cognate ATPase ortholog of the Vibrio cholerae toxin-coregulated-pilus biogenesis apparatus. *Journal of bacteriology* **189**: 4401-4409.

Tripepi, M., R.N. Esquivel, R. Wirth & M. Pohlschroder, (2013) Haloferax volcanii cells lacking the flagellin FlgA2 are hypermotile. *Microbiol-Sgm* **159**: 2249-2258.

Tripepi, M., S. Imam & M. Pohlschroder, (2010) Haloferax volcanii Flagella Are Required for Motility but Are Not Involved in PibD-Dependent Surface Adhesion. *Journal of bacteriology* **192**: 3093-3102.

Tripepi, M., J. You, S. Temel, O. Onder, D. Brisson & M. Pohlschroder, (2012) N-glycosylation of Haloferax volcanii flagellins requires known Agl proteins and is essential for biosynthesis of stable flagella. *Journal of bacteriology* **194**: 4876-4887.

Vakonakis, I. & A.C. LiWang, (2004) Structure of the C-terminal domain of the clock protein KaiA in complex with a KaiC-derived peptide: implications for KaiC regulation. *Proceedings of the National Academy of Sciences of the United States of America* **101**: 10925-10930.

van Wolferen, M., M. Ajon, A.J. Driessen & S.V. Albers, (2013) Molecular analysis of the UV-inducible pili operon from Sulfolobus acidocaldarius. *MicrobiologyOpen* **2**: 928-937.

VanDyke, D.J., J. Wu, S.M. Logan, J.F. Kelly, S. Mizuno, S. Aizawa & K.F. Jarrell, (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. *Molecular microbiology* **72**: 633-644.

Varga, J.J., V. Nguyen, D.K. O'Brien, K. Rodgers, R.A. Walker & S.B. Melville, (2006) Type IV pili-dependent gliding motility in the Gram-positive pathogen Clostridium perfringens and other Clostridia. *Molecular microbiology* **62**: 680-694.

Voisin, S., R.S. Houliston, J. Kelly, J.R. Brisson, D. Watson, S.L. Bardy, K.F. Jarrell & S.M. Logan, (2005) Identification and characterization of the unique N-linked glycan common to the flagellins and S-layer glycoprotein of Methanococcus voltae. *Journal of Biological Chemistry* **280**: 16586-16593.

Wagner, M., S. Berkner, M. Ajon, A.J. Driessen, G. Lipps & S.V. Albers, (2009) Expanding and understanding the genetic toolbox of the hyperthermophilic genus Sulfolobus. *Biochemical Society transactions* **37**: 97-101.

Wagner, M., M. van Wolferen, A. Wagner, K. Lassak, B.H. Meyer, J. Reimann & S.V. Albers, (2012) Versatile Genetic Tool Box for the Crenarchaeote Sulfolobus acidocaldarius. *Frontiers in microbiology* **3**: 214.

Wang, B., S.F. Yang, L. Zhang & Z.G. He, (2010) Archaeal Eukaryote-Like Serine/Threonine Protein Kinase Interacts with and Phosphorylates a Forkhead-Associated-Domain-Containing Protein. *Journal of bacteriology* **192**: 1956-1964.

Weijers, J.W.H., S. Schouten, E.C. Hopmans, J.A.J. Geenevasen, O.R.P. David, J.M. Coleman, R.D. Pancost & J.S.S. Damste, (2006) Membrane lipids of mesophilic anaerobic bacteria thriving in peats have typical archaeal traits. *Environ Microbiol* **8**: 648-657.

Whitehead, K., M. Pan, K. Masumura, R. Bonneau & N.S. Baliga, (2009) Diurnally Entrained Anticipatory Behavior in Archaea. *PloS one* **4**.

Woese, C.R., R. Gupta, C.M. Hahn, W. Zillig & J. Tu, (1984) The phylogenetic relationships of three sulfur dependent archaebacteria. *Systematic and applied microbiology* **5**: 97-105.

Woese, C.R., O. Kandler & M.L. Wheelis, (1990) Towards a Natural System of Organisms - Proposal for the Domains Archaea, Bacteria, and Eucarya. *Proceedings of the National Academy of Sciences of the United States of America* **87**: 4576-4579.

Woese, C.R., L.J. Magrum & G.E. Fox, (1978) Archaebacteria. *Journal of molecular evolution* **11**: 245-251.

Xia, Q., T.S. Wang, E.L. Hendrickson, T.J. Lie, M. Hackett & J.A. Leigh, (2009) Quantitative proteomics of nutrient limitation in the hydrogenotrophic methanogen Methanococcus maripaludis. *Bmc Microbiol* **9**.

Xu, Y., T. Mori, R. Pattanayek, S. Pattanayek, M. Egli & C.H. Johnson, (2004) Identification of key phosphorylation sites in the circadian clock protein KaiC by crystallographic and mutagenetic analyses. *Proceedings of the National Academy of Sciences of the United States of America* **101**: 13933-13938.

Xu, Y., T. Mori, X. Qin, H. Yan, M. Egli & C.H. Johnson, (2009) Intramolecular regulation of phosphorylation status of the circadian clock protein KaiC. *PloS one* **4**: e7509.

Yamagata, A. & J.A. Tainer, (2007) Hexameric structures of the archaeal secretion ATPase GspE and implications for a universal secretion mechanism. *The EMBO journal* **26**: 878-890.

Yamaguchi, S., H. Fujita, A. Ishihara, S. Aizawa & R.M. Macnab, (1986) Subdivision of flagellar genes of Salmonella typhimurium into regions responsible for assembly, rotation, and switching. *J Bacteriol* **166**: 187-193.

Yamauchi, K., K. Doi, Y. Yoshida & M. Kinoshita, (1993) Archaebacterial Lipids - Highly Proton-Impermeable Membranes from 1,2-Diphytanyl-Sn-Glycero-3-Phosphocholine. *Biochim Biophys Acta* **1146**: 178-182.

Yonekura, K., S. Maki, D.G. Morgan, D.J. DeRosier, F. Vonderviszt, K. Imada & K. Namba, (2000) The bacterial flagellar cap as the rotary promoter of flagellin self-assembly. *Science* **290**: 2148-2152.

Appendix

It seems that FlaX is a highly important component of the crenarchaeal archaellum, which determines the size and shape of the basal body. We demonstrated the strong interaction between FlaX, FlaH, FlaI and most probably also FlaJ, which form together the cytoplasmic core of the archaellum of *S. acidocaldarius* (Lassak et al., 2012b, Banerjee et al., 2012a, Banerjee et al., 2013). The euryarchaeal archaellum lacks FlaX, however it contains a set of Fla proteins, which are absent in crenarchaeota: FlaC, FlaD, FlaE. So far it could been demonstrated that FlaCDE play a role in chemotaxis, but nothing is known about their structures and spatial organization within the euryarchaeal archaellum (Schlesner et al., 2009). However our first interests were the homologues of proteins we already characterized in the *S. acidocaldarius* archaella system, namely FlaI and FlaH. We started wondering how preserved their functions in Archaea are in general.

We started with the construction of heterologous expression vectors, where we selected three FlaH donor species: *Pyrococcus furiosus*, *Methanococcus maripaludis* and *Methanocaldococcus jannaschii*. We obtained a high expression of FlaH from *P. furiosus*, which could be also successfully purified with a high yield and purity (Fig. 1). Because the expression of the other FlaH homologues was at that point more problematic, we decided to focus on FlaH from *P. furiosus*. It seemed to be a reasonable choice for our interspecies comparison, thus *P. furiosus* is similarly to *S. acidocaldarius* an extreme thermophile and both inhabit environments with temperature above 70^{o}C.

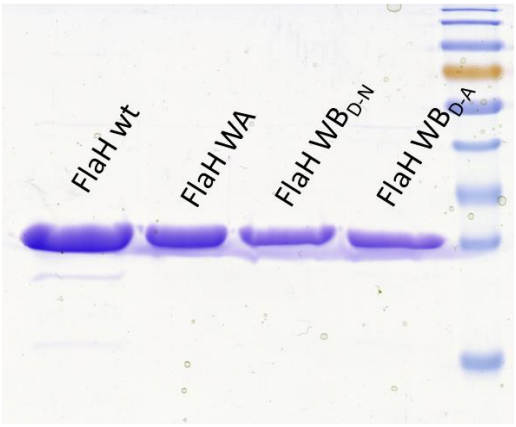


Fig.1. SDS-PAGE of *P. furiosus* FlaH final purification products: (from left to right) FlaH^{wt}; FlaH^{K39A}; FlaH^{D126N}; FlaH^{D126A};

P. furiosus FlaH shows similar Walker motifs composition as its *S. acidocaldarius* homologue, carrying the same non canonical Walker B. Similarly like for *S. acidocaldarius* FlaH we could not demonstrate any ATPase activity, but using FRET we could confirm strong ATP binding. In order to confirm if FlaH of *P. furiosus* also utilized both its Walker boxes for nucleotide binding, we introduced the same mutations into the Walker motifs, as we did for FlaH of *S. acidocaldarius*. Subsequently the FlaH WA$_{K-A}$, WB$_{D-N}$ and WB$_{D-A}$ mutants were heterologously expressed in *E. coli* and successfully purified (Fig.1). Next we performed FRET and analyzed the FlaH WA$_{K-A}$ and WB$_{D-N}$ mutants for ATP binding and the obtained results were in a line with what we could show for *S. acidocaldarius* FlaH (Fig.2). Thus both Walker mutants

showed strongly reduced ATP-binding capability. This data shows that FlaH is in general an ATP binding protein, which utilizes both Walker motifs for the nucleotide binding, but does not possess any ATPase activity.

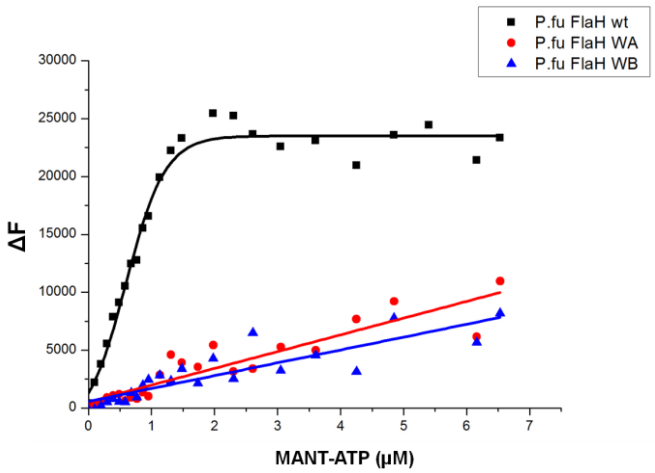


Fig.2. *P. furiosus* FlaH requires both WA and WB motifs for ATP binding. FRET analysis of the MANT-ATP binding to *P.furiosus* FlaH and *P.furiosus* FlaH Walker motif mutants. Change in fluorescence (ΔF) upon MANT-ATP binding is shown as a function of the MANT-ATP concentration. Black square, $FlaH^{WT}$; red circle, $FlaH^{K39A}$; blue triangle, $FlaH^{D126N}$. Fluorescence spectra were corrected for the fluorescence of unbound MANT-ATP.

In the next step we analyzed whether *P. furiosus* FlaH can autophosphorylate like it was demonstrated already for FlaH of *S. acidocaldarius.* For this purpose we performed an *in vitro* assay, where we incubated *P. furiosus* FlaH with radioactively labeled γ-P^{32}-ATP at 55^0C and collected samples at increasing time intervals, like it was done for FlaH of *S. acidocaldarius* [Chapter 3.4]. Also the obtained here results ware highly consistent with what was shown in case of the crenarchaeal FlaH. A clear phosphorylation signal could be detected already after 15 minutes of incubation and its intensity was increasing over time (Fig. 3.). The negative control as also the five minutes incubation sample didn't show any phosphorylation pattern, indicating that the later signals appeared due to covalently bound radioactive labeled phosphate group, not because of an unspecific ATP binding. Summarizing these results it seems that FlaH is an ATP-binding protein which most probably regulates the archaellum assembly or function in a time dependent phosphorylation manner.

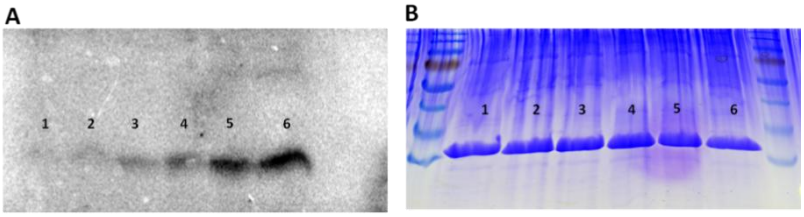


Fig.3. *P. furiosus* FlaH shows autophosphorylation activity. (A) Autophosphorylation assay showing that FlaH auto-phosphorylates. FlaH was incubated with γ-P_{32}-ATP at 55^oC for 0, 5, 15, 30, 60 and 120 minutes corresponding to the samples 1-6 respectively. (B) The corresponding SDS-PAGE.

Erklärung

Ich versichere, dass ich meine Dissertation:

„*Decipering the subunit interaction in the crenarchaeal archaellum*“

selbstständig, ohne unerlaubte Hilfe angefertigt und mich dabei keiner anderen als der von mir ausdrücklich bezeichneten Quellen und Hilfen bedient habe. Die Dissertation wurde in der jetzigen oder einer ähnlichen Form noch bei keiner anderen Hochschule eingereicht und hat noch keinen sonstigen Prüfungszwecken gedient.

Marburg, den ……

_____	_____
Tomasz Neiner	Prof. Dr. Sonja-Verena Albers
(Verfasser)	(Betreuerin)