nature microbiology

SUPPLEMENTARY INFORMATION

https://doi.org/10.1038/s41564-019-0622-3

In the format provided by the authors and unedited.

The structure of the periplasmic FlaG–FlaF complex and its essential role for archaellar swimming motility

Chi-Lin Tsai[1,10], Patrick Tripp[2,3,10], Shamphavi Sivabalasarma[2], Changyi Zhang[4,5], Marta Rodriguez-Franco[6], Rebecca L. Wipfler[4], Paushali Chaudhury[2], Ankan Banerjee[7], Morgan Beeby[8], Rachel J. Whitaker[4,5], John A. Tainer[1,9]* and Sonja-Verena Albers[2,3]*

[1]Department of Molecular and Cellular Oncology, The University of Texas MD Anderson Cancer Center, Houston, TX, USA. [2]Molecular Biology of Archaea, Institute of Biology II, Faculty of Biology, University of Freiburg, Freiburg, Germany. [3]Spemann Graduate School of Biology and Medicine, University of Freiburg, Freiburg, Germany. [4]Carl R. Woese Institute for Genomic Biology, University of Illinois at Urbana-Champaign, Urbana, IL, USA. [5]Department of Microbiology, University of Illinois at Urbana-Champaign, Urbana, IL, USA. [6]Cell Biology, Institute of Biology II, Faculty of Biology, University of Freiburg, Freiburg, Germany. [7]Molecular Biology Program, Sloan-Kettering Institute, New York, NY, USA. [8]Department of Life Sciences, Imperial College London, London, UK. [9]Molecular Biophysics and Integrated Bioimaging Division, Lawrence Berkeley National Laboratory, Berkeley, CA, USA. [10]These authors contributed equally: C.-L. Tsai, P. Tripp. *e-mail: jtainer@mdanderson.org; sonja.albers@biologie.uni-freiburg.de

Supplementary Results

Structural analysis of the tetrameric FlaF/G complex. The solved complex structure exhibited four molecules per asymmetric unit and unusually high solvent content ~76%. Analysis of the complex assemblies by PISA[1] showed the sFlaG/F tetrameric complex to be a stable complex with the calculated free energy of the sFlaG/F tetrameric complex $\Delta G^{assembly}$ = -12.0 kcal/mol and $\Delta G^{dissociation}$ = 18.8 kcal/mol; and of the sFlaG/F dimer $\Delta G^{assembly}$ = -10.3 kcal/mol and $\Delta G^{dissociation}$ = 1.5 kcal/mol. Importantly, the previously shown disordered N-terminal helices of sFlaF (PDB: 4P94) and sFlaG (PDB: 5TUH) became ordered in the sFlaG/F tetrameric complex: they show well-defined connectivity and occupy a distinct region of the electron density map (Supplementary Fig. 2). Supporting the crystal structure complex, the SAXS data of sFlaG/F complex showed a calculated MW ~57 kDa corresponding to a tetrameric complex (Supplementary Table 1), and the sFlaG/F tetramer assemble model with disordered N-terminal His-tag of sFlaG/F generated from MODELLER[2] and MultiFoXS[3] software fits the measured SAXS data with χ^2 = 1.73 Å (Supplementary Fig. 3a). The SAXS data furthermore showed the tetrameric complex to be a rigid and well-folded protein shown in a dimensionless Kratky plot (Supplementary Fig. 3b and Supplementary Table 1), which increases confidence in the assemble[4].

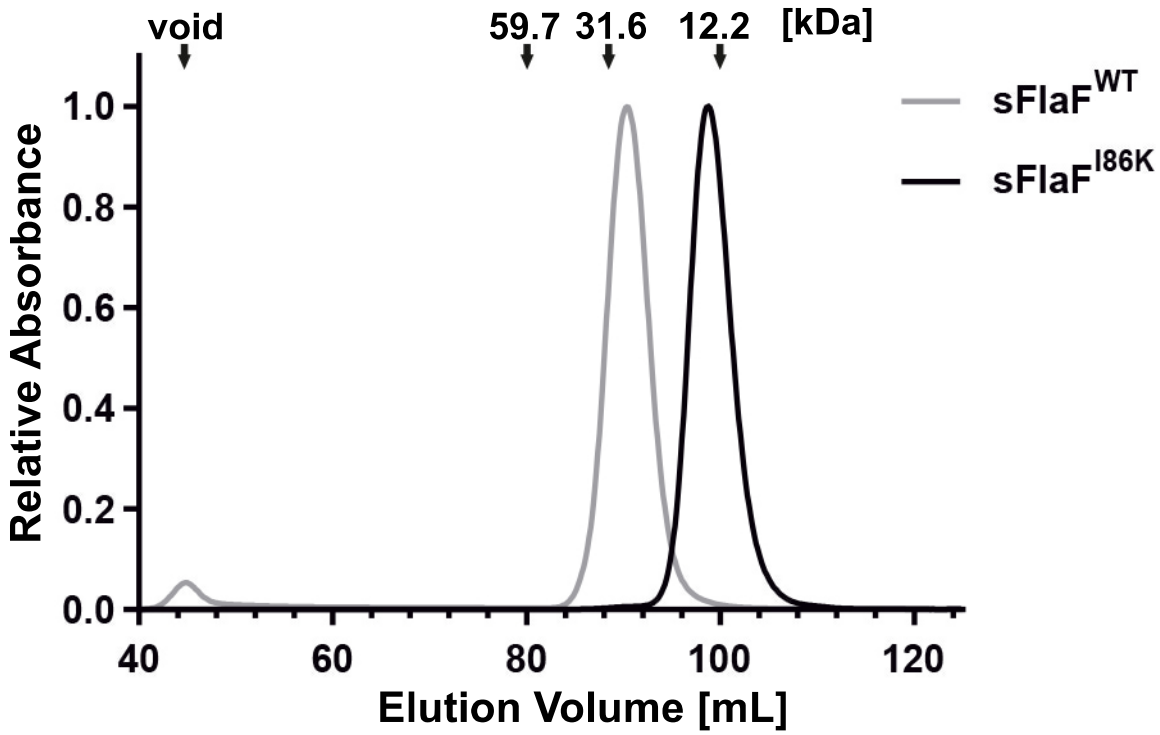
The orientation of the FlaG and FlaF N-termini was pointed in opposite directions, suggesting an alternative oligomeric complex assembly conformation although N-termini of FlaF and FlaG contain the transmembrane sequence (Fig. 2b). However, we do not exclude the possibility of the bipolar direction of FlaG/F N-termini was due to the crystallographic effect. We also observed the annotation of secondary structure of sFlaG N-terminal strands (β1 and β2) (Supplementary Fig. 4c-d) is missing from

PDBsum database. We reasoned that N-terminal strands (β1 and β2) of sFlaG, which interacts with adjacent sFlaG and sFlaF subunits, only appear in the context of sFlaG/F tetrameric complex while PDBsum uses individual sFlaG structure, which appears to be unstructured on its N-terminal region. In sum, N-terminal strands of sFlaG stabilize the tetramer by intertwining with another sFlaG, which is further stabilized by two adjacent sFlaF proteins via a long α-helix in addition to hydrophobic interaction in dimer interface.

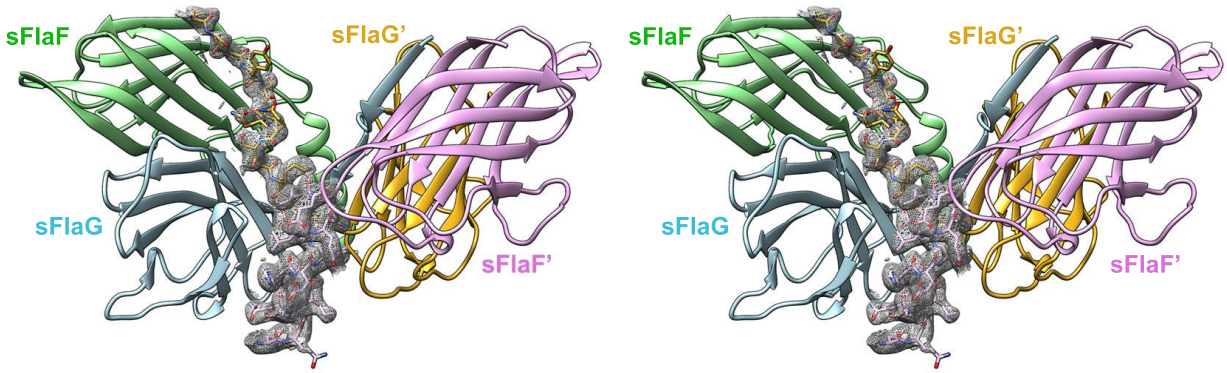
Analysis of FlaF N-glycosylation sites

Since glycosylation is a common modification to extracellular proteins in Archaea, we reasoned that FlaF might undergo N-glycosylation. Therefore, we identified several Asn residues as good candidates for potential glycosylation sites using NetNGlyc server 1.0[5]. Indeed, the N92A variant of FlaF shows a loss of the higher MW species (upper band), but not the lower band, while the band pattern of the N105A variant did not change (Supplementary Fig. 10a), suggesting that $FlaF^{N92}$ is N-glycosylated. The $FlaF^{N92A}$-HA variant lacking glycosylation was still able to complement motility in a Δ*aapF*Δ*flaF* background strain (Supplementary Fig. 10b), suggesting that residue N92 glycosylation of FlaF is dispensable for cell motility.

Supplementary Figures

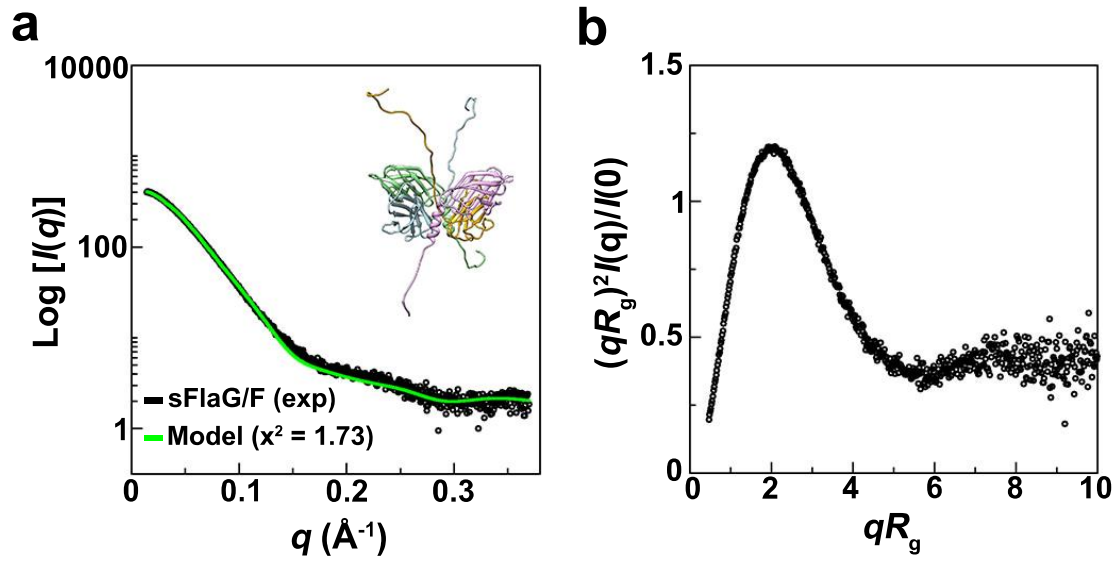


Supplementary Figure 1: SEC profile of sFlaF variants at pH 3 citric acid buffer. The oligomeric states of sFlaF^{WT} and sFlaF^{I86K} were assessed at pH 3 citric acid buffer by SEC and shown to be consistent with the previously reported dimer and monomer, respectively, at pH 8 buffer[6].

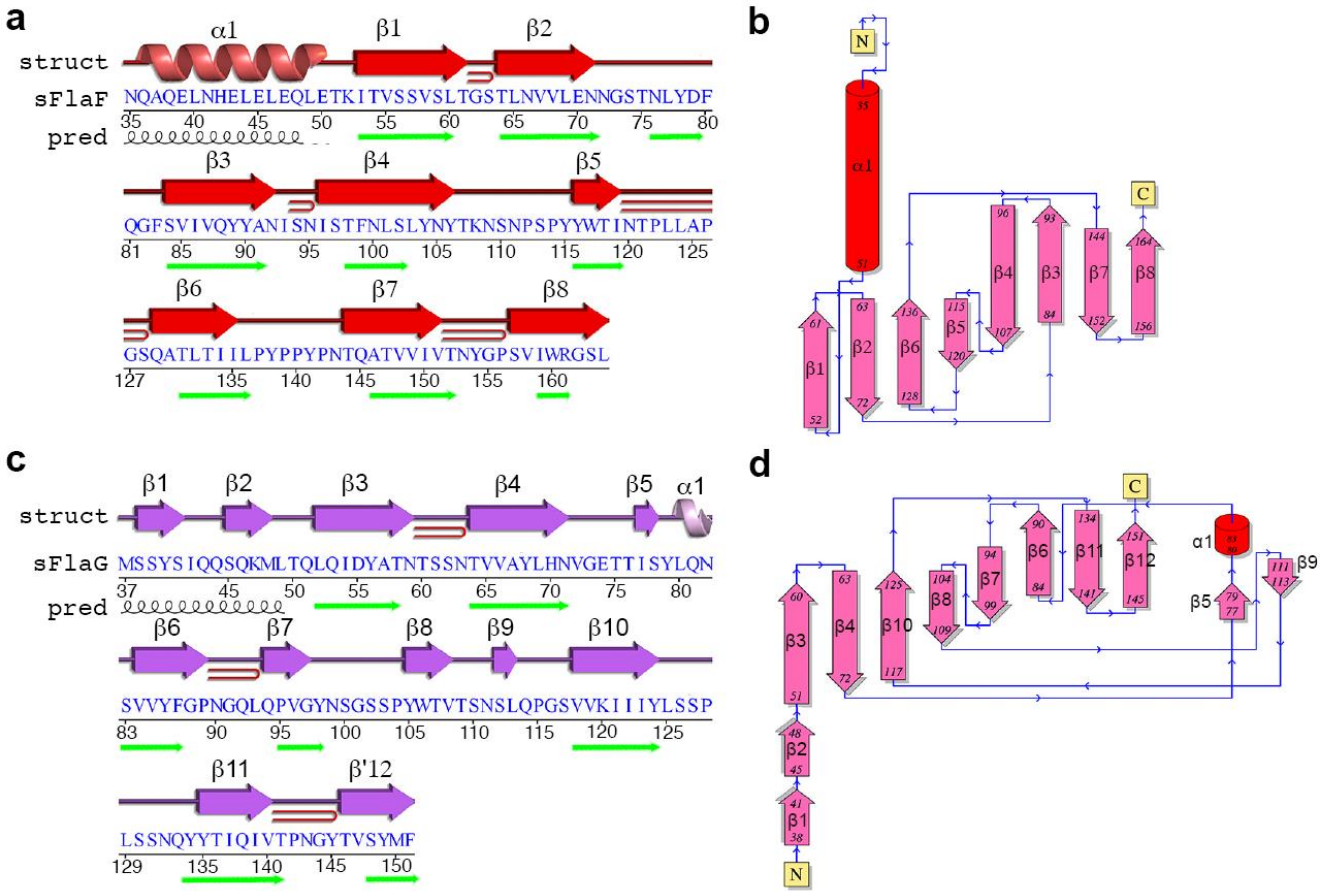


Supplementary Figure 2: Electron density map of the sFlaG/F complex.

Stereoview (wall-eyed) of electron density map (2Fo-Fc, gray mesh) generated from x-ray data of sFlaG/F tetrameric complex is mapped on the N-terminal α-helix of sFlaF (pink) and β-strands of sFlaG (gold) at 1σ contour level.

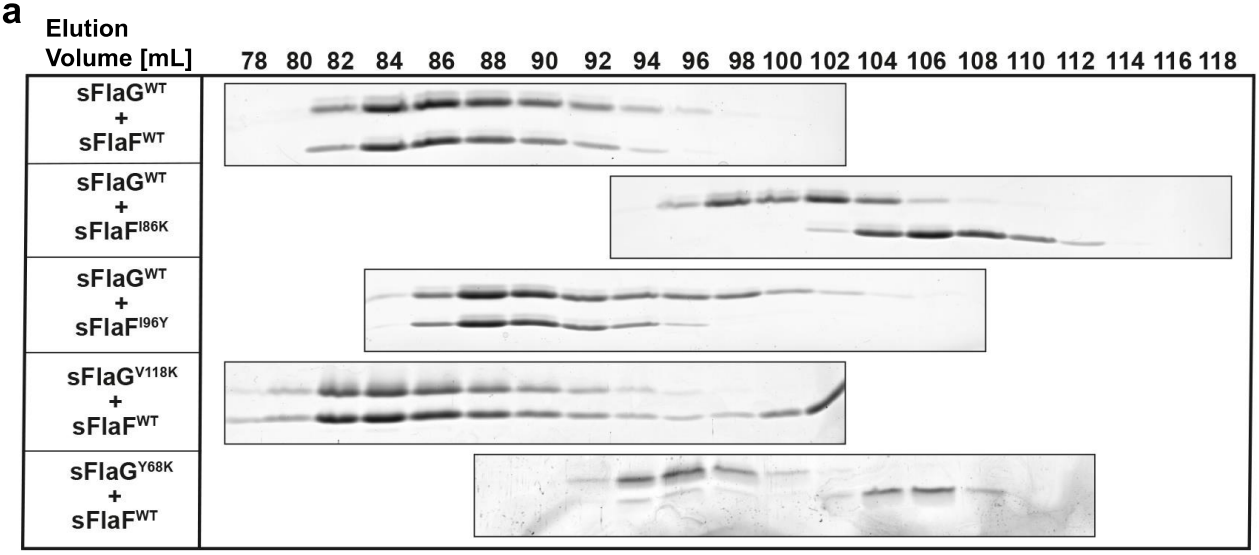


Supplementary Figure 3: SAXS analysis of the heterotetrameric sFlaG/F complex. a, The sFlaG/F tetrameric complex model fit to SAXS experimental data shows that the N-terminal regions of sFlaG and sFlaF are flexible in solution. **b,** The dimensionless Kratky plot of the sFlaG/F complex from SAXS data exhibits a well-folded protein.



Supplementary Figure 4: The secondary structure layout, prediction, and topology from crystal structure of the sFlaG/sFlaF complex. a, sFlaF secondary structure. **b,** sFlaF fold topology. **c,** sFlaG secondary structure. **d,** sFlaG fold topology. The layout and topology were generated from the PDBsum database except the β1 and β1 strands of sFlaG are annotated based on the tetrameric complex , and secondary structure prediction is generated from the Jpred4 server[7].

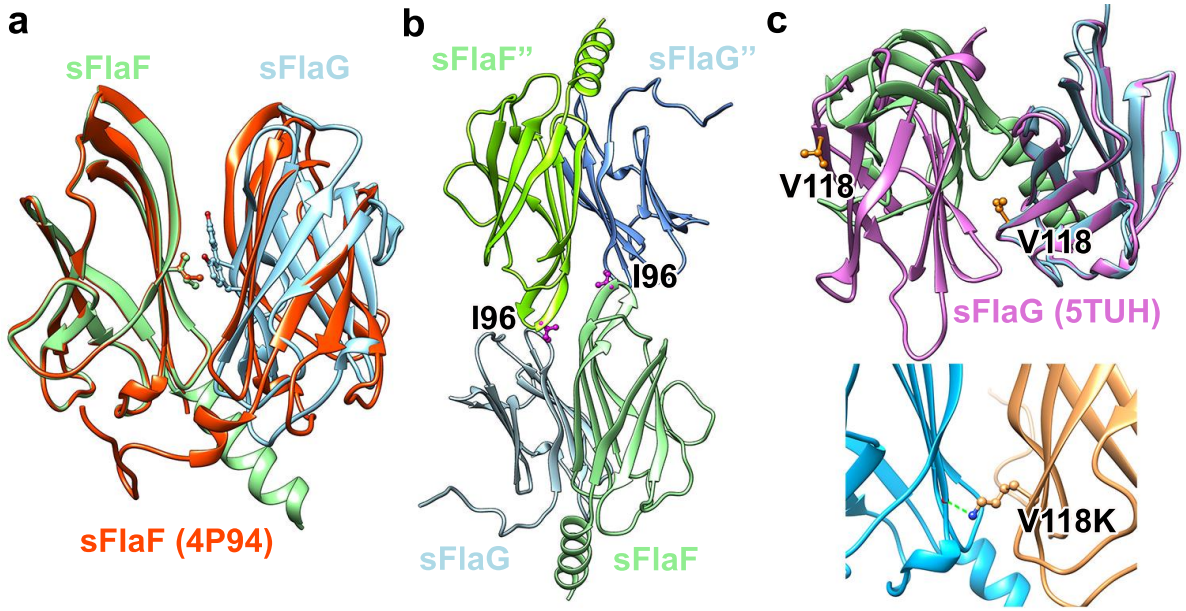
a



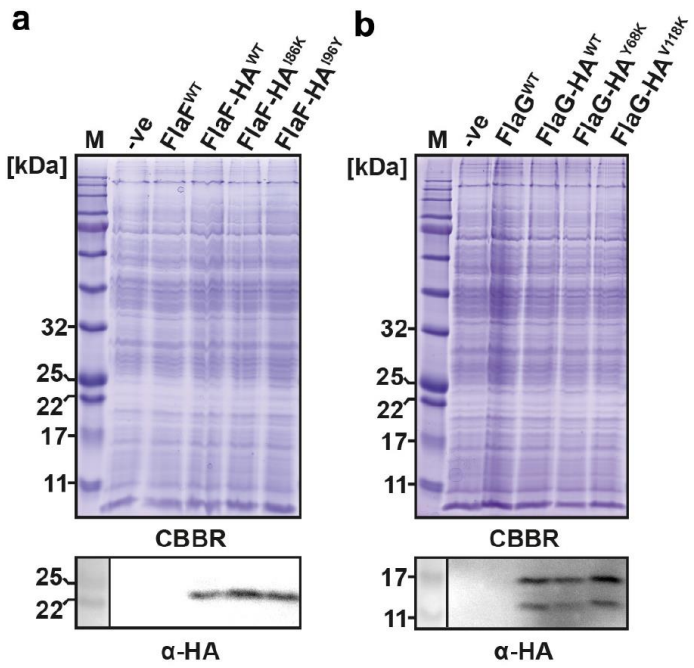
b

Protein	$sFlaG^{WT}$	$sFlaF^{WT}$	$sFlaG^{Y68K}$	$sFlaG^{V118K}$	$sFlaF^{I86K}$	$sFlaF^{I96Y}$	$sFlaG^{WT}$/ $sFlaF^{WT}$	$sFlaG^{WT}$/ $sFlaF^{I96Y}$	$sFlaG^{V118K}$ /$sFlaF^{WT}$
MW (kDa)	12.5	26.2	12.5	11.3	13.5	17.5	47	34.5	53

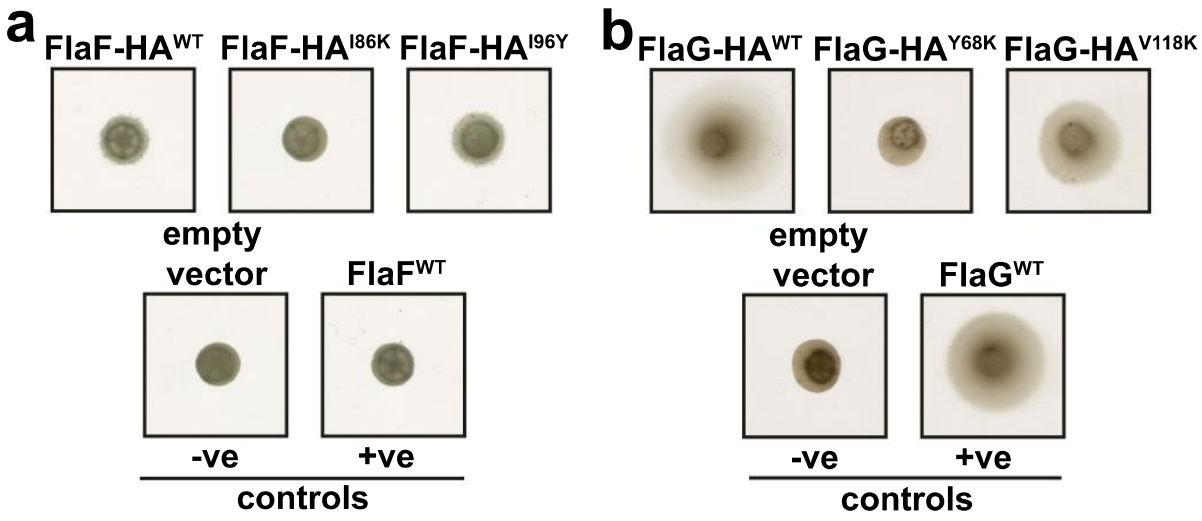
Supplementary Figure 5: SDS-PAGE analysis and MW calculations of the reconstituted sFlaG/sFlaF-complex after SEC. a. The corresponding SEC fractions from figures 3a-d were analyzed on SDS-PAGE. Gels were arranged according to the respective elution volume of the fractions. **b.** The MW of each SEC peak in Fig. 3a-d was calculated from SEC elution volume according to the molecular weight standards. The theoretical MW for sFlaF and sFlaG is 16 and 15 kDa, respectively.



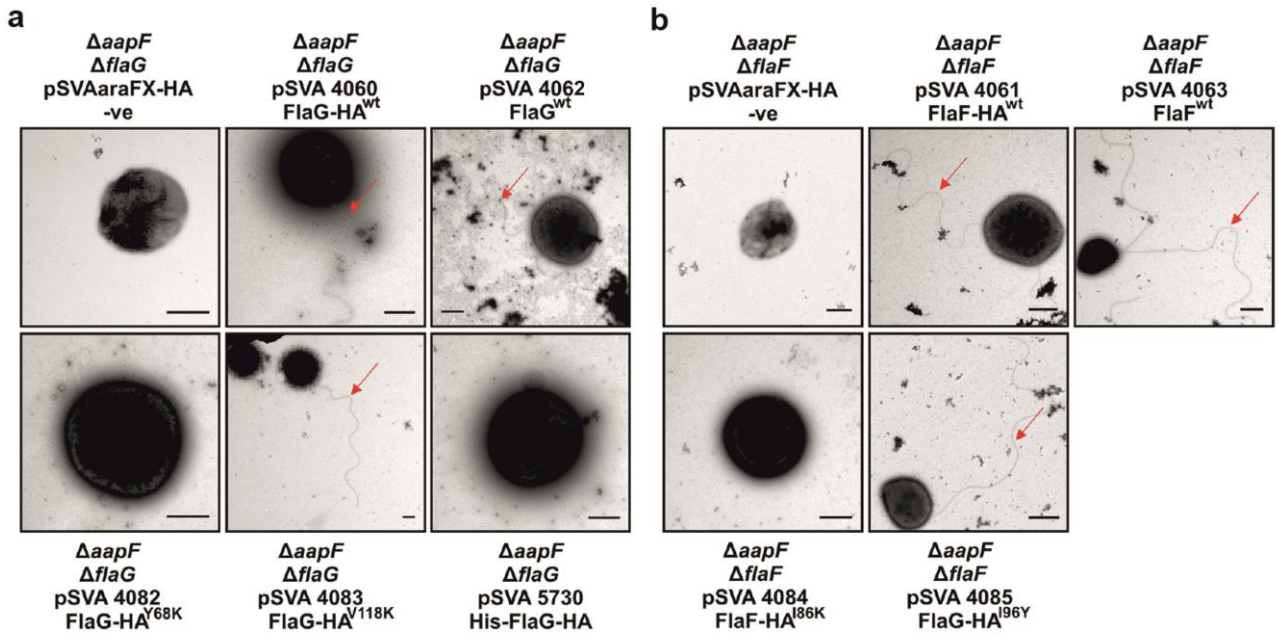
Supplementary Figure 6: Separation-of-function mutations show that sFlaG and sFlaF share a common dimer interface. a, Overlay of residue I86 on sFlaG/sFlaF dimer (PDB: 5TUG) and sFlaF-sFlaF dimer (PDB: 4P94) reveals that the same hydrophobic dimer interface was used. **b,** Residue I96 of sFlaF located at different crystallographic tetramer interfaces was substituted to tyrosine (Y) to disrupt the possible interaction. **c,** (top panel) Residue V118 located at the sFlaG (PDB: 5TUH) dimer interface was substituted to lysine (K) to disrupt the possible sFlaG dimer. (bottom panel) The crystal structure of FlaG^{V118K}/FlaF complex showed V118K substitution does not disrupt complex formation and is hydrogen-bonding with carbonyl group of F83 on sFlaF (shown in blue).



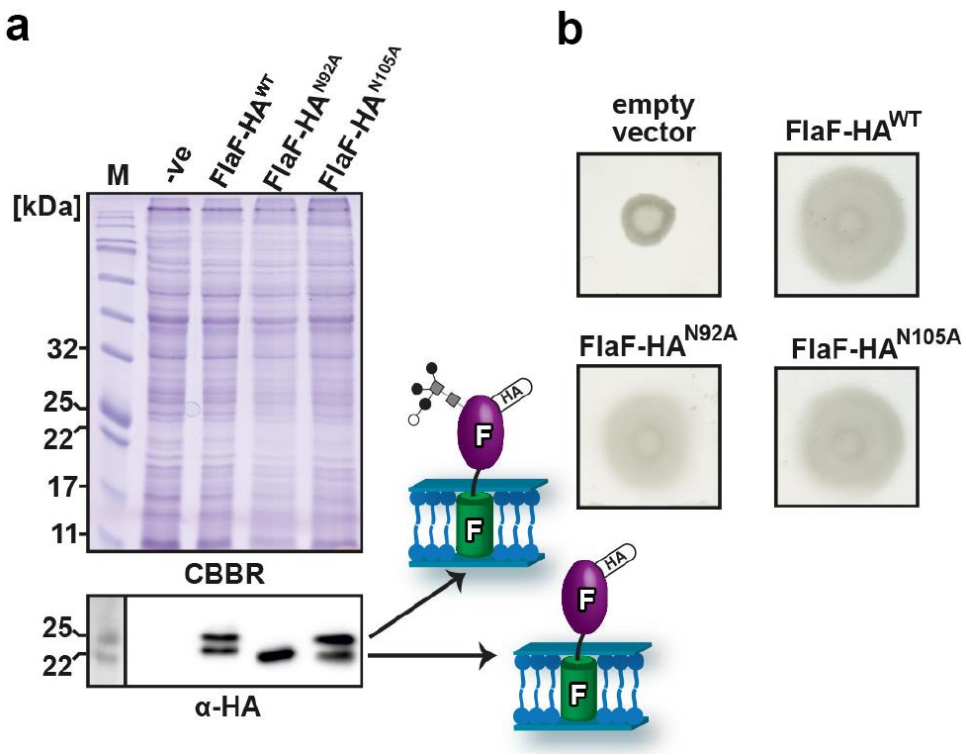
Supplementary Figure 7: Expression level assessment of FlaF and FlaG variants for *in trans* complementation experiments. Coomassie Brilliant Blue stained SDS-PAGE (CBBR) and anti-HA western blot (α-HA) analysis of *in trans* complemented Δ*aapF*Δ*flaF* (**a**) and Δ*aapF*Δ*flaG* (**b**) strains. Empty plasmid (pSVAaraFX-HA, -ve) and untagged wild-type proteins served as negative controls, and stained SDS-PAGE served as loading control. Protein expression was induced by induction/starvation with 0.4% L-arabinose in medium lacking NZ-Amine. **a,** HA-tagged FlaF^{WT} and variants were produced in equal amounts and detected as a single band. **b,** HA-tagged FlaG^{WT} and variants were produced in equal amounts and detected as an upper, full-length band and a lower, degraded band.



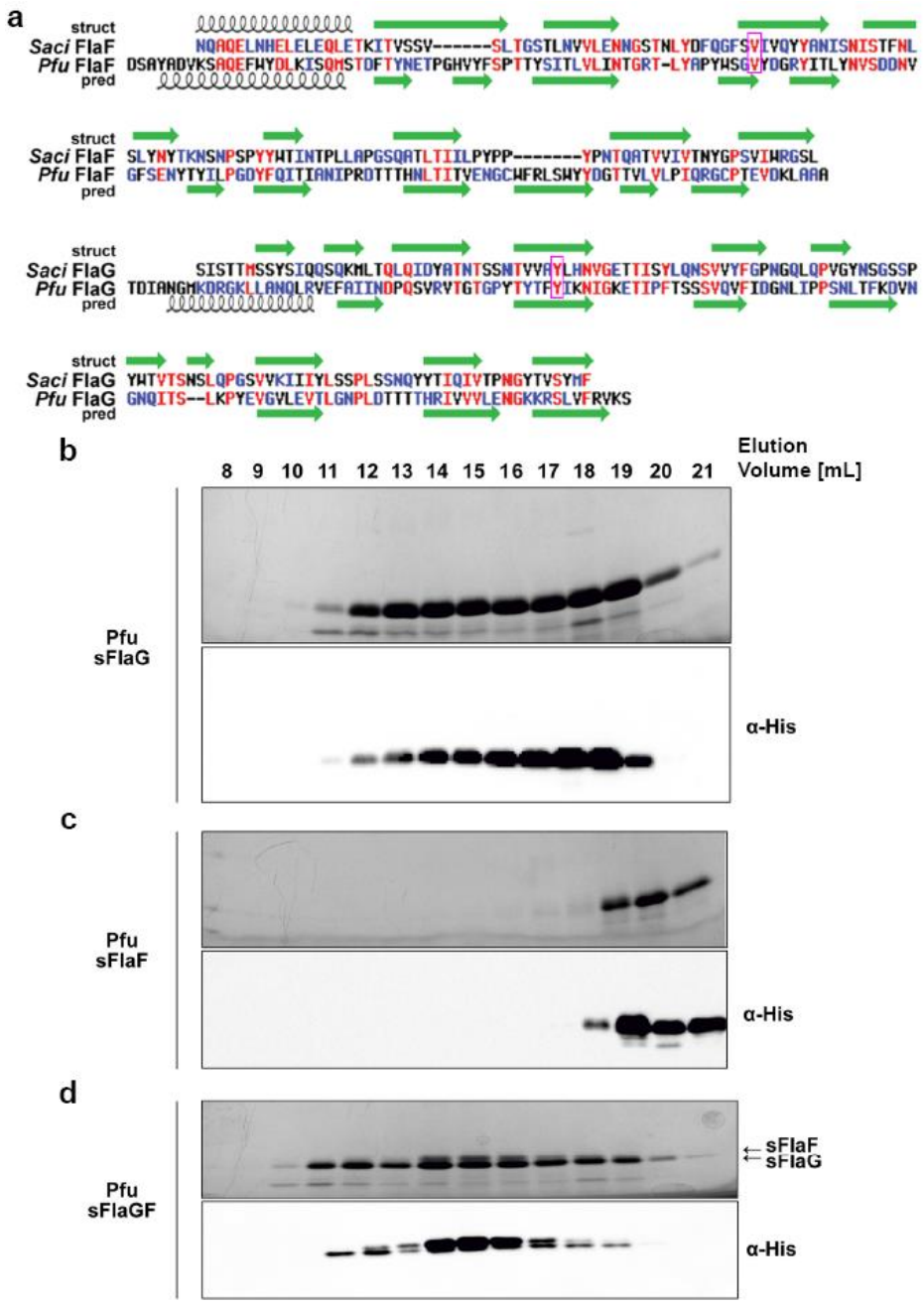
Supplementary Figure 8: *In trans* complementation of Δ*aapF*Δ*flaF* and Δ*aapF*Δ*flaG* mutants with expression vectors carrying different FlaF or FlaG species under control of an arabinose-inducible promoter. a, Upon addition of 0.4% L-arabinose, motility was abolished for FlaF-overexpressing variants. Even the wild-type complementation was rendered non-motile, showing that FlaF overexpression impairs motility. **b**, Upon addition of 0.4% L-arabinose, FlaG overexpression did not impair cell motility.



Supplementary Figure 9: Visualization of the archaellum in FlaF and FlaG variants under transmission electron microscopy. a, Δ*aapF*Δ*flaG* strain with expressing FlaG variants. **b**, Δ*aapF*Δ*flaF* strain with expressing FlaF variants. Notably, FlaG-HA^{Y68K}, His-FlaG-HA, and FlaF-HA^{I86K} variants have no archaellum formation. Red arrows point to archaella. Scale bars = 500 nm.



Supplementary Figure 10: FlaF glycosylation is not essential for swimming motility. a, Membrane fractions of Δ*aapF*Δ*flaF* cells expressing HA-tagged wild-type FlaF and Asn-variants analyzed by Coomassie stain and anti-HA western blot. **b,** Effect of Asn-variants on motility in *S. acidocaldarius*.



Supplementary Figure 11: Secondary structure prediction, sequence alignment, and purification of Pfu sFlaG and sFlaF. **a**, Comparison of secondary structure between Saci and Pfu sFlaG and sFlaF. Prediction was performed in JPred 4 server[7]. The FlaG/F complex disrupting residues FlaF^{I86} and FlaG^{Y68} are shown in the pink boxes. **b-d**, SDS-PAGE and western-blot analyses of the Pfu sFlaG (b), sFlaF (c), and sFlaG/F (d) SEC elution fractions were analyzed by Coomassie-stain (top panel) and anti-His western-blot (α-His).

Supplementary Tables

Supplementary Table 1: Summary of x-ray scattering data analysis

Protein Variants	sFlaG	sFlaG/F complex WT	sFlaG/F complex I96Y-sFlaF	sFlaG/F complex V118K-sFlaG
Data collection				
Beamline	ALS BL12.3.1	ALS BL12.3.1	ALS BL12.3.1	ALS BL12.3.1
Beam energy, keV	11	11	11	11
Sample-detector distance, m	1.5	1.5	1.5	1.5
Detector	Pilatus	Pilatus	Pilatus	Pilatus
Exposure time (s)	Every 0.3s. Total: 10.2 s.	Every 0.3s. Total: 10.2 s.	Every 0.3s. Total: 10.2 s.	Every 0.3s. Total: 10.2 s.
Images No.	Total: 33 images	Total: 33 images	Total: 33 images	Total: 33 images
Sample cell thickness, mm	1.5	1.5	1.5	1.5
Temperature, K	283	283	283	283
Final q range, $Å^{-1}$	0.01 to 0.5	0.01 to 0.5	0.01 to 0.5	0.01 to 0.5
Data analysis				
Programs	SCÅTTER 3.1R	SCÅTTER 3.1R	SCÅTTER 3.1R	SCÅTTER 3.1R
Buffer	25 mM citric acid/sodium citrate (pH 3), 150mM NaCl, 3% glycerol	25 mM citric acid/sodium citrate (pH 3), 150mM NaCl, 3% glycerol	25 mM citric acid/sodium citrate (pH 3), 150mM NaCl, 3% glycerol	25 mM citric acid/sodium citrate (pH 3), 150mM NaCl, 3% glycerol
Protein concentration, mg/mL	1.5, 3, 5	2, 4, 6	1.5, 2, 3	0.95, 1.4, 1.9
Points used for Guinier analysis	1-38	1-53	1-62	1-51
Guinier qR_g limits	1.3	1.3	1.3	1.3
$I(0)$, cm^{-1}	45.030 ± 0.498	442.400 ± 1.037	214.700 ± 1.081	16.810 ± 0.050
Guinier R_g, Å	36.31 ± 0.22	32.15 ± 0.38	28.10 ± 0.80	32.29 ± 0.50
D_{max}, Å	180	125	82	125
R_g (real), Å	43.91 ± 2.74	33.97 ± 0.16	26.74 ± 0.04	34.26 ± 0.32
R_g (reciprocal), Å	43.55	33.82	26.75	34.05
MW estimation, kDa*	24..49	57.06	50.85	61.27
Porod volume, $Å^3$ (q-range, $Å^{-1}$)	186874 (0.085 – 0.140)	129162 (0.093 – 0.130)	104487 (0.090 – 0.108)	128994 (0.090 – 0.106)
Porod Exponent	1.61 ± 0.05	3.71 ± 0.02	3.92 ± 0.02	3.72 ± 0.05
SASBDB ID	SASDEU7	SASDES7	SASDEV7	SASDET7

*MW is estimated based on $MW = \frac{(V_c)2/R_g}{0.1231}$ using SCÅTTER program 3.1R (http://www.bioisis.net/scatter)

Supplementary Table 2. X-ray diffraction data collection and refinement statistics for sFlaG and the sFlaG/sFlaF complexes

	sFlaG/sFlaF Complex	sFlaG^{V118K}/ sFlaF Complex	sFlaG Native		sFlaG (Pt)	
Data collection						
Space group	$P6_1$	$P6_1$	$C222_1$		$C222_1$	
Cell dimensions						
a, *b*, *c* (Å)	119.95 119.95 152.44	120.038 120.038 152.138	83.02, 100.32, 62.82		83.52, 99.32, 62.21	
α, β, γ (°)	90.00 90.00 120.00	90.00 90.00 120.00	90.00 90.00 90.00		90.00 90.00 90.00	
				Peak	*Inflection*	*Remote*
Wavelength	0.9918	0.9795	1.0714	1.0714	1.0718	1.0448
Resolution (Å)	49.16 – 2.47 (2.56 – 2.47)*	49.19 – 2.80 (2.96 – 2.80)*	44.82 − 1.93 (1.97 − 1.93)*	44.58 – 2.50 (2.60 – 2.50)*	44.61 – 2.50 (2.60 – 2.50)*	44.59 – 2.50 (2.60 – 2.50)*
R_{sym} or R_{merge}	0.10 (1.23)	0.08 (1.17)	0.09 (0.97)	0.10 (0.54)	0.11 (0.67)	0.10 (0.58)
I / σI	19.2 (2.4)	16.4 (1.5)	14.9 (2.1)	21.5 (5.1)	21.1 (4.1)	21.4 (4.9)
Completeness (%)	100 (100)	99.9 (99.7)	100 (99.6)	99.7 (99.0)	99.7 (99.4)	99.7 (99.2)
Redundancy	11.4 (11.5)	6.8 (6.8)	7.2 (7)	14.3 (13.9)	14.4 (13.9)	14.4 (14.5)
Refinement						
Resolution (Å)	49.15 – 2.47	39.29 – 2.81	44.82 – 1.93			
No. reflections	44511	30378	20093			
R_{work} / R_{free}	16.3/19.4	17.1/20.2	18.1/21.3			
No. atoms						
Protein	3836	3858	1644			
Water	184	147	135			

B-factors	51.44	75.88	33.86
Protein	51.48	75.60	33.12
Ligand/ion			
Water	51.81	76.67	42.27
R.m.s deviations			
Bond lengths (Å)	0.008	0.008	0.007
Bond angles (°)	0.985	1.070	0.815

Supplementary Table 3: Strains, plasmids and primers

Escherichia coli strains

Strain	Genotype	Source
NEB 10 beta	Δ(ara-leu) 7697 araD139 fhuA ΔlacX74 galK16 galE15 e14-ϕ80dlacZΔM15 recA1 relA1 endA1 nupG rpsL (StrR) rph spoT1 Δ (mrr-hsdRMS-mcrBC)	New England Biolabs
Rosetta(DE3)pLysS	F– ompT gal dcm lon? hsdSB(rB–mB–) λ(DE3 [lacI lacUV5-T7p07 ind1 sam7 nin5]) [malB+]K-12(λS) pLysSRARE [T7p20 ileX argU thrU tyrU glyT thrT argW metT leuW proL orip15A](CmR)	Novagen
ER1821	F- glnV44 e14-(McrA-) rfbD1 relA1 endA1 spoT1	New England Biolabs

Sulfolobus acidocaldarius strains

Strain	Genotype	Source
MW001	DSM 639 Δ*pyrEF* (91-412 bp)	Ref. [8]
MW453	Δ*pyrEF* (91-412 bp) Δ*aapF* Δ*flaG*	Ref. [9]
MW454	Δ*pyrEF* (91-412 bp) Δ*aapF* Δ*flaF*	Ref. [9]
MW502	Δ*pyrEF* (91-412 bp) Δ*aapF* Δ*pibD*	This study

Sulfolobus islandicus strains

Strain	Genotype	Source
RJW004	Δ*pyrEF* Δ*lacS* Δ*argD*; Derived from *S. islandicus* M.16.4	Ref. [10]
Δ*slaA*	Δ*pyrEF* Δ*lacS* Δ*argD*Δ*slaA*	Ref. [11]

Δ*slaB*	Δ*pyrEF* Δ*lacS* Δ*argD* Δ*slaB*	Ref. [11]
Δ*slaAB*	Δ*pyrEF* Δ*lacS* Δ*argD* Δ*slaAB*	Ref. [11]

Plasmids

Plasmid	Relevant characteristics	Source
pET-Duet 1	Amp^r expression plasmid containing replicon ColE1 (pBR322) and two MCS (MCS1 and MCS2)	Novagen
pSVAaraFX-HA	Plasmid for expression in *S. acidocaldarius* based on pSVAmZ-SH10[83] with *ara*-promoter; containing MCS compatible with FX-cloning primers[84] adding C-terminal HA-tag to gene of interest	Wagner et al., unpublished
pSVAaraFX-stop	Plasmid for expression in *S. acidocaldarius* based on pSVAmZ-SH10[83] with *ara*-promoter; containing MCS compatible with FX-cloning primers[84]	Wagner et al., unpublished
pSVA 1921	Amp^r, pET-Duet 1 based expression vector, carrying N-terminal His$_6$-tagged Saci *sflaF* in MCS 1, using EcoRI/HindIII restriction sites	Ref. [6]
pSVA 2816	Amp^r, pSVA 1921 based expression vector, carrying Saci *sflaF*^{I86K}, using site-directed mutagenesis PCR	Ref. [6]
pSVA 2822	Amp^r, pET-Duet 1 based expression vector, carrying N-terminal His$_6$-tagged Pfu *sflaF* in MCS 1, using EcoRI/HindIII restriction sites	This study
pSVA 2826	Amp^r, pET-Duet 1 based expression vector, carrying N-terminal His$_6$-tagged Saci *sflaG* in MCS 1, using EcoRI/HindIII restriction sites	This study
pSVA 4060	Amp^r, pSVAaraFx-HA expression vector, carrying C-terminal hemagglutinin-tagged Saci *flaG*, using NcoI/XhoI restriction sites	This study
pSVA 4061	Amp^r, pSVAaraFX-HA expression vector, carrying C-terminal hemagglutinin-tagged Saci *flaF*, using NcoI/XhoI restriction sites	This study

pSVA 4062	Amp^r, pSVAaraFX-stop expression vector, carrying Saci *flaG*, using NcoI/XhoI restriction sites	This study
pSVA 4063	Amp^r, pSVAaraFX-stop expression vector, carrying Saci *flaF*, using NcoI/XhoI restriction sites	This study
pSVA 4014	Amp^r, pET-Duet 1 based expression vector, carrying N-terminal His_6-tagged Pfu *sflaG* in MCS 1, using EcoRI/HindIII restriction sites	This study
pSVA 4015	Amp^r, pSVA 2826 based expression vector, carrying Saci *sflaG^{V118K}*, using site-directed mutagenesis PCR	This study
pSVA 4016	Amp^r, pSVA 2826 based expression vector, carrying Saci *sflaG^{Y68K}*, using site-directed mutagenesis PCR	This study
pSVA 4019	Amp^r, pSVA 1921 based expression vector, carrying Saci *sflaF*^{I96Y}, using site-directed mutagenesis PCR	This study
pSVA 4082	Amp^r, pSVA 4060 based expression vector, carrying Saci *flaG*^{Y68K}, using site-directed mutagenesis PCR	This study
pSVA 4083	Amp^r, pSVA 4060 based expression vector, carrying Saci *flaG*^{V118K}, using site-directed mutagenesis PCR	This study
pSVA 4084	Amp^r, pSVA 4061 based expression vector, carrying Saci *flaF*^{I86K}, using site-directed mutagenesis PCR	This study
pSVA 4085	Amp^r, pSVA 4061 based expression vector, carrying Saci *flaF*^{I96Y}, using site-directed mutagenesis PCR	This study
pSVA 4087	Amp^r, pSVA 4061 based expression vector, carrying Saci *flaF*^{N92A}, using site-directed mutagenesis PCR	This study
pSVA 4088	Amp^r, pSVA 4061 based expression vector, carrying Saci *flaF*^{N105A}, using site-directed mutagenesis PCR	This study
pSVA 5730	Amp^r, pSVA 4060 based expression vector, carrying Saci *His_6-flaG*, using using overlap PCR to insert the additional tag	This study

Primers

Primer	Sequence and Characteristics	Source
3749	CCCGAATTCGATCTCTACGACAATGTC, forward primer for Saci *sflaG* containing EcoRI restriction site	This study
2143	GGGGAAGCTTTCAAAACATGTAACTAACAG; reverse primer for Saci *sflaG* containing HindIII restriction site	This study
3760	CAAGGTTTTTCCGTTAAAGTACAATACTATGC forward primer for Saci FlaF I86K mutation	Ref. [6]
3761	GCTAATGTTAGCATAGTATTGTACTTTAACGG reverse primer for Saci FlaF I86K mutation	Ref. [6]
3741	GGGGGAATTCGTGGGACAGTGCTTACGCTGACG, forward primer for Pfu *sflaF* containing EcoRI restriction site	This study
3742	GGGGAAGCTTGTCCACCTCCGTTGGGCAGCCTCTCTGGATG, reverse primer for Pfu *sFlaF* containing HindIII restriction site	This study
6657	GCGGAATTCGACAGATATAGCCAATGGCATG, forward primer for Pfu *sflaG* containing EcoRI restricition site	This study
6658	CGCAAGCTTTTAACTTTTCACTCTAAATACGAGAGATC, reverse primer for Pfu *sflaG* containing HindIII restriction site	This study
6667	CTATGCTAACATTAGCAATTATTCAACTTTCAATC forward primer for Saci FlaF I96Y mutation	This study

6668	CGAGAGATTGAAAGTTGAATAATTGCTAATGTTAG reverse primer for Saci FlaF I96Y mutation	This study
6659	CACTACAGCCAGGGTCAAAAGTAAAAATCATTATC forward primer for Saci FlaG V118K mutation	This study
6660	GACAAATAGATAATGATTTTTACTTTTGACCCTGGCTG reverse primer for Saci FlaG V118K mutation	This study
6661	CTCAAACACTGTAGTTGCAAAATTACATAATGTGG forward primer for Saci FlaG Y68K mutation	This study
6662	GTTTCCCCCACATTATGTAATTTTGCAACTACAGTG reverse primer for Saci FlaG Y68K mutation	This study
8236	GCGCCATGGTTAGTGAGGTTATAAGTGAGAC; forward primer for Saci *flaG* containing NcoI restriction site	This study
8215	CGCCTCGAGGCTTACCTTAAACATGTAACTAAC reverse primer for Saci *flaG* containing XhoI restriction site	This study
8237	GCGCCATGGGAGTGTCACAAACTTTG; forward primer for Saci *flaF* containing NcoI restriction site	This study
8216	CGCCTCGAGTAGGCTTCCCCTCCATATTAC; reverse primer for Saci *flaF* containing XhoI restriction site	This study
8246	GTTATAGTACAATACTATGCTGCCATTAGCAATATCTC AACTTTC forward primer for Saci FlaF N92A mutation	This study
8247	GTTGAGATATTGCTAATGGCAGCATAGTATTGTACTA TAACGG reverse primer for Saci FlaF N92A mutation	This study

8248	CTTTCAATCTCTCGTTATATGCCTACACAAAGAACTCTAACC forward primer for Saci FlaF N105A mutation	This study
8249	GAGTTCTTTGTGTAGGCATATAACGAGAGATTGAAAGTTGAG reverse primer for Saci FlaF N105A mutation	This study
9268	CATCACCATCACCATCACGTTAGTGAGGTTATAAGTGAGACCATTATG; forward primer to add His_6-tag to Saci FlaG N-terminus	This study
9269	GTGATGGTGATGGTGATGCATGGTATGATAAGTAAGACGCTTATC; reverse primer to add His_6-tag to Saci FlaG N-terminus	This study

Supplementary References

1. Krissinel, E. & Henrick, K. Inference of Macromolecular Assemblies from Crystalline State. *J. Mol. Biol.* **372,** 774–797 (2007).

2. Webb, B. & Sali, A. Comparative protein structure modeling using MODELLER. *Curr. Protoc. Bioinforma.* **2016,** 5.6.1-5.6.37 (2016).

3. Schneidman-Duhovny, D., Hammel, M., Tainer, J. A. & Sali, A. FoXS, FoXSDock and MultiFoXS: Single-state and multi-state structural modeling of proteins and their complexes based on SAXS profiles. *Nucleic Acids Res.* **44,** W424–W429 (2016).

4. Rambo, R. P. & Tainer, J. A. Super-Resolution in Solution X-Ray Scattering and Its Applications to Structural Systems Biology. *Annu. Rev. Biophys.* **42,** 415–441 (2013).

5. R. Gupta, E. J. and S. B. Prediction of N-glycosylation sites in human proteins. Unpublished

6. Banerjee, A. *et al.* FlaF is a β-sandwich protein that anchors the archaellum in the archaeal cell envelope by binding the S-layer protein. *Structure* **23,** 863–872 (2015).

7. Drozdetskiy, A., Cole, C., Procter, J. & Barton, G. J. JPred4: A protein secondary structure prediction server. *Nucleic Acids Res.* **43,** W389–W394 (2015).

8. Wagner, M. *et al.* Versatile genetic tool box for the crenarchaeote Sulfolobus acidocaldarius. *Front. Microbiol.* **3,** (2012).

9. Lassak, K. *et al.* Molecular analysis of the crenarchaeal flagellum. *Mol. Microbiol.* **83,** 110–124 (2011).

10. Zhang, C., Cooper, T. E., Krause, D. J. & Whitaker, R. J. Augmenting the genetic toolbox for Sulfolobus islandicus with a stringent positive selectable marker for agmatine prototrophy. *Appl. Environ. Microbiol.* **79,** 5539–5549 (2013).

11. Zhang, C., Phillips, A. P. R., Wipfler, R. L., Olsen, G. J. & Whitaker, R. J. The essential genome of the crenarchaeal model Sulfolobus islandicus. *Nat.*

Commun. **9,** 4908 (2018).