This is a repository copy of Effects of N-Glycosylation Site Removal in Archaellins on the Assembly and Function of Archaella in Methanococcus maripaludis.

White Rose Research Online URL for this paper:
http://eprints.whiterose.ac.uk/83841/

Version: Published Version

**Article:**
Ding, Yan, Uchida, Kaoru, Aizawa, Shin-Ichi et al. (5 more authors) (2015) Effects of N-Glycosylation Site Removal in Archaellins on the Assembly and Function of Archaella in Methanococcus maripaludis. Plos one. e0116402. ISSN 1932-6203

https://doi.org/10.1371/journal.pone.0116402

**Reuse**
Unless indicated otherwise, fulltext items are protected by copyright with all rights reserved. The copyright exception in section 29 of the Copyright, Designs and Patents Act 1988 allows the making of a single copy solely for the purpose of non-commercial research or private study within the limits of fair dealing. The publisher or other rights-holder may allow further reproduction and re-use of this version - refer to the White Rose Research Online record for this item. Where records identify the publisher as the copyright holder, users can verify any specific terms of use on the publisher's website.

**Takedown**
If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.
Effects of N-Glycosylation Site Removal in Archaellins on the Assembly and Function of Archaella in *Methanococcus maripaludis*

Yan Ding¹, Kaoru Uchida², Shin-Ichi Aizawa², Kathleen Murphy³, Alison Berezuk³, Cezar M. Khursigara³, James P. J. Chong⁴, Ken F. Jarrell¹*

¹ Department of Biomedical and Molecular Sciences, Queen’s University, Kingston, Ontario, Canada, ² Department of Life Sciences, Prefectural University of Hiroshima, 562 Nanatsuka, Shobara, Hiroshima, Japan, ³ Department of Molecular and Cellular Biology, University of Guelph, Guelph, Ontario, Canada, ⁴ Department of Biology, University of York, Heslington, York, United Kingdom

* jarrellk@queensu.ca

Abstract

In *Methanococcus maripaludis* S2, the swimming organelle, the archaellum, is composed of three archaellins, FlaB¹S², FlaB²S² and FlaB³S². All three are modified with an N-linked tetrasaccharide at multiple sites. Disruption of the N-linked glycosylation pathway is known to cause defects in archaella assembly or function. Here, we explored the potential requirement of N-glycosylation of archaellins on archaellation by investigating the effects of eliminating the 4 N-glycosylation sites in the wildtype FlaB²S² protein in all possible combinations either by Asn to Glu (N to Q) substitution or Asn to Asp (N to D) substitutions of the N-glycosylation sequon asparagine. The ability of these mutant derivatives to complement a non-archaellated ΔflaB²S² strain was examined by electron microscopy (for archaella assembly) and swarm plates (for analysis of swimming). Western blot results showed that all mutated FlaB²S² proteins were expressed and of smaller apparent molecular mass compared to wildtype FlaB²S², consistent with the loss of glycosylation sites. In the 8 single-site mutant complements, archaella were observed on the surface of Q2, D2 and D4 (numbers after N or Q refer to the 1st to 4th glycosylation site). Of the 6 double-site mutation complements all were archaellated except D1,3. Of the 4 triple-site mutation complements, only D2,3,4 was archaellated. Elimination of all 4 N-glycosylation sites resulted in non-archaellated cells, indicating some minimum amount of archaellin glycosylation was necessary for their incorporation into stable archaella. All complementations that led to a return of archaella also resulted in motile cells with the exception of the D4 version. In addition, a series of FlaB²S² scanning deletions each missing 10 amino acids was also generated and tested for their ability to complement the ΔflaB²S² strain. While most variants were expressed, none of them restored archaellation, although FlaB²S² harbouring a smaller 3-amino acid deletion was able to partially restore archaellation.

© 2015 Ding et al. This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are within the paper.

Funding: This work is supported by grants from the Natural Sciences and Engineering Research Council of Canada (NSERC) to K.F.J. and C.M.K. and Biotechnology and Biological Sciences Research Council (BBSRC) to J.P.C. K.F.J. was the recipient of a Leverhulme Trust Visiting Professorship. Y.D. thanks the China Scholarship Council (20100622028) for financial support. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.
**Introduction**

*N*-glycosylation is a prevalent protein modification found in all three domains of life in which the attachment of the glycan is via the nitrogen atom in asparagine residues located in the acceptor glycoprotein [1–6]. General features of the *N*-glycosylation pathways are shared among the three domains [1–3,5,7,8]. The *N*-glycan precursor is first synthesized on a phosphorylated isoprene-based lipid carrier, either a dolichol derivative in Eukarya and Archaea, or an undecaprenol derivative in Bacteria, via the activities of glycosyltransferases. The assembled lipid-linked glycan is then flipped across a membrane, i.e. to face into the lumen of the endoplasmic reticulum in Eukarya or to the external face of the cytoplasmic membrane in Archaea and Bacteria. A signature enzyme of the pathway, the oligosaccharyltransferase (OST), transfers the completed glycan *en bloc* from its lipid carrier to select Asn residues in target proteins, although further sugars can still be added to the protein-bound glycan. The Asn residue to which the *N*-glycan is attached is usually located in an Asn-Xaa-Ser/Thr sequon (Xaa cannot be Pro), although in some Bacteria, i.e. *Campylobacter* spp., a negatively charged amino acid residue is also needed at the -2 position. Between the two prokaryotic domains, Bacteria and Archaea, *N*-glycosylation appears to be much more widespread among Archaea, where 166 of 168 sequenced genomes examined contained at least one copy of a gene encoding the archaean OST *AglB* [9]. A variety of archaean proteins, mainly S-layer proteins and the subunits of surface structures including both pilins and archaellins (formerly archaeal flagellins [10]), have been shown to be modified with *N*-glycans [11–18]. Recent work on archaean *N*-glycosylation systems has combined both structural and genetic methods, typically using archaellins and S-layer proteins as reporter proteins. Since the first archaean glycosylation (*agl*) genes were identified in 2006 [14,19], this combined approach has focused on a few key archaean model organisms [20]: *Methanococcus maripaludis* S2, *Methanococcus voltae* PS, *Haloferax volcanii* H53 and *Sulfolobus acidocaldarius* MW001 [1,6,11,12,15,16,21]. Interestingly, the *N*-glycosylation pathway is not essential for any of the three studied euryarchaeotes (*M. maripaludis* S2, *M. voltae* PS and *Hfx. volcanii* H53), as mutants carrying a deletion or insertional inactivation of *aglB* were readily obtained [14,19,22]. However, in the crenarchaeote *S. acidocaldarius* MW001, repeated attempts to delete or interrupt *aglB* were unsuccessful and only the integration of a second copy of *aglB* into the genome allowed for the deletion of the original *aglB* [23]. The key reported effects that result from perturbation or complete abolition of the *N*-glycosylation pathway in Archaea are on S-layer stability, growth of cells at high salinities and on archaellation and motility [18,19,22,24–30].

The archaellum (formerly archaean flagellum) is the major motility apparatus found in Archaea [10]. It is a rotating appendage unrelated to the bacterial flagellum but bearing instead similarities to the bacterial type IV pilus. These similarities include homologous ATPase and membrane platform proteins involved in assembly of the structure as well as similarities in their major structural proteins [31–35]. The structural proteins, called archaellins, are made as preproteins with class III (type IV pilin-like) signal peptides which are removed by a specific prepilin peptidase-like enzyme (termed FlaK in methanogens or PibD in Sulfolobales and halo-phil) [36–39]. Signal peptide removal is critical for incorporation of archaellins into the filament [37,39]. In addition, it appears that archaellins are commonly modified with *N*-linked glycans and interruption of the normal *N*-glycosylation pathway leads to defects in archaella assembly or function [1,40]. In species where *aglB* has been deleted or insertionally inactivated (*M. voltae* PS, *M. maripaludis* S2, *Hfx. volcanii* H53), the cells are unable to make archaella. If the pathway is interrupted at steps that lead to a truncated glycan, there is an impairment in motility, although archaella are still made unless the truncation is too great [18,19,22]. In *M. maripaludis* S2 strain, archaella contain 3 archaellins: the major archaellins FlaB1S2 and FlaB2S2 form the filament while the minor archaellin FlaB3S2 comprises the hook region [41].
The three archaellins share sequence similarities including a class III signal peptide cleaved by FlaK, conserved N-terminal and C-terminal regions, and a hypervariable region in the middle [41]. The hypervariable regions of FlaB1$_{S2}$, FlaB2$_{S2}$ and FlaB3$_{S2}$ are decorated at multiple positions with a unique tetrasaccharide [11]. Cells are archaellated if they carry deletions in agl genes that result in a N-glycan of at least two sugars, but non-archaellated if the deleted agl gene results in a glycan of only a single sugar or prevents N-glycosylation totally (as in an agl deletion) [22,42,43]. These results could mean that the archaellins must be glycosylated at some or all of its glycosylation sequons by at least a two sugar glycan for those archaellins to be assembled into a structure. However, an alternative explanation is that the necessity for the glycosylation lies at a different step in archaella assembly.

A major goal of this study was to examine the requirement of N-glycosylation of the major archaellin FlaB2$_{S2}$ for archaellation in M. maripaludis S2. FlaB2$_{S2}$ has five N-glycosylation sequons, $^{26}$NTS$^{28}$, $^{66}$NIT$^{68}$, $^{110}$NLT$^{112}$, $^{119}$NTT$^{121}$ and $^{124}$NWS$^{126}$. The first sequon $^{26}$NTS$^{28}$ located in the N-terminal conserved region was previously reported to be unoccupied with glycan while the remaining four, located in the hypervariable region, were modified with tetrasaccharide (Fig. 1, [11]). For these experiments, we eliminated the 4 occupied sequons ($^{66}$NIT$^{68}$, $^{110}$NLT$^{112}$, $^{119}$NTT$^{121}$ and $^{124}$NWS$^{126}$, designated as the 1$^\text{st}$, 2$^\text{nd}$, 3$^\text{rd}$ and 4$^\text{th}$ N-glycosylation site, respectively) in all possible combinations (creating single-, double-, triple- and quadruple-site mutations in FlaB2$_{S2}$). We also generated a series of FlaB2$_{S2}$ scanning deletions in an attempt to determine regions of the molecule that were essential for assembly of archaella.

**Materials and Methods**

**Strains and growth conditions**

*M. maripaludis S2 Δhpt* (Mm900) [44], *M. maripaludis S2 Δhpt ΔflaB2$_{S2}$ (ΔflaB2$_{S2}$ in short hereafter) [41] and all complemented strains, as well as *M. maripaludis ΔRC* (formerly *Methanothermobacter deltae* ΔRC [45,46]) were routinely cultured anaerobically in 125 mL sealed serum bottles containing 10 mL Balch medium III under an atmosphere of CO$_2$:H$_2$ (20:80) at 37°C with shaking [47]. Cells carrying a complementation plasmid were cultured in the presence of 2.5 μg/mL puromycin for plasmid selection [48]. For swarming assays, cells were inoculated onto plates of Balch medium III containing 0.25% (w/v) agar in the presence of 2.5 μg/mL puromycin [22]. *Escherichia coli* TOP 10 cells (Invitrogen Inc.), used for molecular cloning steps, were cultured at 37°C in Luria Broth (LB) medium with shaking or on LB plates (containing 1.5% w/v agar) in the presence of 100 μg/mL ampicillin for plasmid selection. Strains and plasmids used in this study are listed in Table 1.

**Construction of mutant *flaB2$_{S2}$* genes using site-directed mutagenesis (SDM)**

To generate mutant *flaB2$_{S2}$* genes, the wildtype *flaB2$_{S2}$* gene was first cloned into the pCR2.1-TOPO-TA vector (Invitrogen Inc.) to create pKJ902. This pCR2.1-TOPO-*flaB2$_{S2}$* and its derivatives were used as template to generate the mutants listed in Table 2. The wildtype *flaB2$_{S2}$* gene used in cloning was generated by PCR using the complementation primers listed in Table 3 and genomic DNA from Mm900 as template.

To generate the mutant *flaB2$_{S2}$* genes that would encode proteins in which N-glycosylation sites were eliminated, the SDM protocol was employed [49]. Forward and reverse mutagenic primer pairs were designed with nucleotide changes located in the middle of the primer that would result in a change of the N-glycosylation sequon Asn residue (Table 3). Purified PCR products were digested with *Dpn*I to remove the template plasmid, repurified and then...
transformed into *E. coli* TOP10 competent cells. Plasmids extracted from the transformants were sequenced to confirm the mutation. Using this method, 8 single-site mutant *flaB2* genes were generated that resulted in 4 N to Q single-site mutations and 4 N to D single-site mutations in their protein products. Double-site mutant *flaB2* genes were then generated using the plasmids with the single-site changes in *flaB2* as template. The same strategy was used to create the triple and quadruple glycosylation site mutant *flaB2* genes. The multi-site mutant proteins all contained N to D changes only.

The *G6* (*flaB2*ΔRC) gene was amplified by PCR using *M. maripaludis* ΔRC whole cells as template with the complementation primers in Table 3 and subsequently cloned into pCR2.1-TOPO. The *G10* gene whose protein product contains additional glycosylation sites was chemically synthesized by GenScript USA Inc. (Piscataway, NJ).

![Protein sequence alignment of FlaB2-S2, G6 (FlaB2ΔRC) and G10.](image)

**Table 1.** Strains and plasmids used in this study.

<table>
<thead>
<tr>
<th>Strains</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>M. maripaludis</strong>&lt;br&gt; <em>M. maripaludis</em> S2 Δhpt (Mm900)</td>
<td>44</td>
</tr>
<tr>
<td><em>M. maripaludis</em> S2 Δhpt ΔflaB2S2</td>
<td>41</td>
</tr>
<tr>
<td><em>M. maripaludis</em> ΔRC</td>
<td>46</td>
</tr>
<tr>
<td><strong>E. coli</strong>&lt;br&gt;<em>E. coli</em> TOP10</td>
<td>Invitrogen Inc.</td>
</tr>
</tbody>
</table>

**Plasmids**

<table>
<thead>
<tr>
<th>Plasmids</th>
<th>TA cloning vector, Amp′, Kan′&lt;br&gt;pCR2.1-TOPO</th>
<th>Invitrogen Inc.</th>
</tr>
</thead>
<tbody>
<tr>
<td>pKJ902</td>
<td>flaB2S2 in vector pCR2.1-TOPO</td>
<td>This study</td>
</tr>
<tr>
<td>pCR2.1-TOPO-flaB2S2</td>
<td>Mutant <em>flaB2</em> genes * in vector pCR2.1-TOPO derivatives</td>
<td>This study</td>
</tr>
<tr>
<td>pWLG40</td>
<td><em>hmv</em> promoter-lacZ fusion plus Pur′ cassette; Amp′</td>
<td>48</td>
</tr>
<tr>
<td>pKJ1064</td>
<td>flaB2S2 in shuttle vector pWLG40 under a <em>hmv</em> promoter</td>
<td>This study</td>
</tr>
<tr>
<td>pWLG40-flaB2S2 derivatives*</td>
<td>Mutant <em>flaB2</em> genes * in vector pWLG40</td>
<td>This study</td>
</tr>
</tbody>
</table>

* Please refer to Table 2

![DOI:10.1371/journal.pone.0116402.g001](image)

![DOI:10.1371/journal.pone.0116402.t001](image)
Table 2. Mutated FlaB2S2 derivatives generated in this study.

<table>
<thead>
<tr>
<th>Mutated FlaB2 derivatives</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control</td>
<td>WT Wildtype flaB2S2 from M. maripaludis S2</td>
</tr>
<tr>
<td>Mutated FlaB2 derivatives containing mutations at N-glycosylation sites</td>
<td></td>
</tr>
<tr>
<td>N to Q single-site mutation</td>
<td></td>
</tr>
<tr>
<td>Q1</td>
<td>FlaB2S2 N66Q</td>
</tr>
<tr>
<td>Q2</td>
<td>FlaB2S2 N110Q</td>
</tr>
<tr>
<td>Q3</td>
<td>FlaB2S2 N119Q</td>
</tr>
<tr>
<td>Q4</td>
<td>FlaB2S2 N124Q</td>
</tr>
<tr>
<td>N to D single-site mutation</td>
<td></td>
</tr>
<tr>
<td>D1</td>
<td>FlaB2S2 N66D</td>
</tr>
<tr>
<td>D2</td>
<td>FlaB2S2 N110D</td>
</tr>
<tr>
<td>D3</td>
<td>FlaB2S2 N119D</td>
</tr>
<tr>
<td>D4</td>
<td>FlaB2S2 N124D</td>
</tr>
<tr>
<td>N to D double-site mutation</td>
<td></td>
</tr>
<tr>
<td>D1,2</td>
<td>FlaB2S2 N66D N110D</td>
</tr>
<tr>
<td>D1,3</td>
<td>FlaB2S2 N66D N119D</td>
</tr>
<tr>
<td>D1,4</td>
<td>FlaB2S2 N66D N124D</td>
</tr>
<tr>
<td>D2,3</td>
<td>FlaB2S2 N110D N119D</td>
</tr>
<tr>
<td>D2,4</td>
<td>FlaB2S2 N110D N124D</td>
</tr>
<tr>
<td>D3,4</td>
<td>FlaB2S2 N119D N124D</td>
</tr>
<tr>
<td>N to D triple-site mutation</td>
<td></td>
</tr>
<tr>
<td>D1,2,3</td>
<td>FlaB2S2 N66D N110D N119D</td>
</tr>
<tr>
<td>D1,2,4</td>
<td>FlaB2S2 N66D N110D N124D</td>
</tr>
<tr>
<td>D1,3,4</td>
<td>FlaB2S2 N66D N119D N124D</td>
</tr>
<tr>
<td>D2,3,4</td>
<td>FlaB2S2 N110D N119D N124D</td>
</tr>
<tr>
<td>N to D quadruple-site mutation</td>
<td></td>
</tr>
<tr>
<td>D1,2,3,4</td>
<td>FlaB2S2 N66D N110D N119D N124D</td>
</tr>
<tr>
<td>Mutated FlaB2S2 with extra N-glycosylation sequons</td>
<td></td>
</tr>
<tr>
<td>G6</td>
<td>FlaB2S2 60AGT to 56NGS 106 DDT 106 to 104NIS 106, N110D, A129N</td>
</tr>
<tr>
<td>G10</td>
<td>FlaB2S2 60AGT to 56NGS 106 G79T, V88T, 96TTK 98 to 94NTT 98, 104DDT 106 to 104NIS 106, A129N</td>
</tr>
<tr>
<td>Ten-amino acid scanning deletions</td>
<td></td>
</tr>
<tr>
<td>Δ2–10</td>
<td>FlaB2S2 missing 2SGIGTLIVF 10</td>
</tr>
<tr>
<td>Δ4–10</td>
<td>FlaB2S2 missing 4IGTLIVF 10</td>
</tr>
<tr>
<td>Δ11–20</td>
<td>FlaB2S2 missing 11IAMVLVAAVA 30</td>
</tr>
<tr>
<td>Δ21–30</td>
<td>FlaB2S2 missing 21ASVLINTSGF 30</td>
</tr>
<tr>
<td>Δ31–40</td>
<td>FlaB2S2 missing 31LQQKASTTGK 30</td>
</tr>
<tr>
<td>Δ41–50</td>
<td>FlaB2S2 missing 41DSTEQVASGL 50</td>
</tr>
<tr>
<td>Δ51–60</td>
<td>FlaB2S2 missing 51QMGISGYQA 60</td>
</tr>
<tr>
<td>Δ61–70</td>
<td>FlaB2S2 missing 61GTANANTKL 70 a</td>
</tr>
<tr>
<td>Δ71–80</td>
<td>FlaB2S2 missing 71AITYPNAGS 60</td>
</tr>
<tr>
<td>Δ81–90</td>
<td>FlaB2S2 missing 81AADMNQVVL 90</td>
</tr>
<tr>
<td>Δ91–100</td>
<td>FlaB2S2 missing 91TLDGTTKTV 100</td>
</tr>
<tr>
<td>Δ101–110</td>
<td>FlaB2S2 missing 101TKYDTTAYTN 110 a</td>
</tr>
<tr>
<td>Δ111–120</td>
<td>FlaB2S2 missing 111LTAGGDLYNT 120 a</td>
</tr>
<tr>
<td>Δ121–130</td>
<td>FlaB2S2 missing 121TVNWSKLAND 130 a</td>
</tr>
</tbody>
</table>

(Continued)
For construction of the mutant versions of \textit{flaB}2\textsubscript{S2} whose protein products would contain scanning deletions, inverse PCR and overlapping primers was employed, again with \textit{pKJ902} as template. The forward primer was designed to contain the \textit{flaB}2\textsubscript{S2} gene sequence flanking the desired in-frame 30 bp deletion. The reverse primer was complementary to the forward primer sequence upstream of the deletion (Table 2). After \textit{DpnI} digestion, purified PCR products were transformed into \textit{E. coli} TOP10 competent cells. Recombinant plasmids were extracted from the transformants and used as template for amplifying the \textit{flaB}2\textsubscript{S2} mutant genes by PCR with the complementation primers listed in Table 2. The smaller \textit{flaB}2\textsubscript{S2} mutant genes were identified by agarose gel electrophoresis of the PCR products which were also sequenced to confirm the deletion. The same protocol was used to generate the 3 amino acid deletion, \textit{Δ}161–170 \textit{flaB}2\textsubscript{S2} missing 161\textit{IIVSGVSFDT}\textsuperscript{170}, where the deletion of the 9bp resulted in the removal of an \textit{RsaI} restriction site. This allowed for the screening of the \textit{flaB}2\textsubscript{S2} gene in plasmids carried by transformants for the small deletion by digestion of subsequent \textit{flaB}2\textsubscript{S2} PCR products with \textit{RsaI}. Lastly, a mutant \textit{flaB}2\textsubscript{S2} gene encoding a mutated \textit{FlaB}2\textsubscript{S2} protein with a 10-amino acid deletion at 91\textit{TLSDGTTKTV}\textsuperscript{100} had those amino acids replaced with a copy of the 10 amino acids at 161\textit{IIVSGVSFDT}\textsuperscript{170}, thereby generating a substitution mutant version of \textit{FlaB}2\textsubscript{S2} (SUB) that was still the same length as the wildtype version. All the mutant versions of \textit{flaB}2\textsubscript{S2} generated were sequenced to confirm the mutations.

**Table 2. (Continued)**

<table>
<thead>
<tr>
<th>Mutated \textit{FlaB2} derivatives</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Δ131–140</td>
<td>\textit{FlaB}_{2\textsubscript{S2}} missing \textsuperscript{131}\textit{TTEFGIVEI}Q\textsuperscript{140}</td>
</tr>
<tr>
<td>Δ141–150</td>
<td>\textit{FlaB}_{2\textsubscript{S2}} missing \textsuperscript{141}\textit{DADLSFT}SSA\textsuperscript{150}</td>
</tr>
<tr>
<td>Δ151–160</td>
<td>\textit{FlaB}_{2\textsubscript{S2}} missing \textsuperscript{151}\textit{PVINKGDIV}A\textsuperscript{160}</td>
</tr>
<tr>
<td>Δ161–170</td>
<td>\textit{FlaB}_{2\textsubscript{S2}} missing \textsuperscript{161}\textit{IIVSGVS}F\textsuperscript{170}SDT</td>
</tr>
<tr>
<td>Δ171–180</td>
<td>\textit{FlaB}_{2\textsubscript{S2}} missing \textsuperscript{171}\textit{RMESGT}VQPD\textsuperscript{180}</td>
</tr>
<tr>
<td>Δ181–190</td>
<td>\textit{FlaB}_{2\textsubscript{S2}} missing \textsuperscript{181}\textit{EFGAPGIS}F\textsuperscript{190}</td>
</tr>
<tr>
<td>Δ194–204</td>
<td>\textit{FlaB}_{2\textsubscript{S2}} missing \textsuperscript{194}\textit{STFTEK}VSLQ\textsuperscript{204}</td>
</tr>
</tbody>
</table>

**Three-amino acid deletion**

| Δ3AA     | \textit{FlaB}_{2\textsubscript{S2}} missing \textsuperscript{61}\textit{GTA}\textsuperscript{63} |

**Substitution mutant**

| SUB      | \textit{FlaB}_{2\textsubscript{S2}} having \textsuperscript{91}\textit{TLSDGTT}KT\textsuperscript{100}V replacing 161\textit{IIVSGVS}FSDT |

The mutant \textit{flaB}_{2\textsubscript{S2}} genes were first generated in \textit{pCR2.1-TOPO} and then cloned into the \textit{M. maripaludis} expression vector \textit{pWL}G40.

\*bold letters: \textit{N}-glycosylation sequons with the asparagine underlined


doi:10.1371/journal.pone.0116402.t002

For construction of the mutant versions of \textit{flaB}_{2\textsubscript{S2}} whose protein products would contain scanning deletions, inverse PCR and overlapping primers was employed, again with pKJ902 as template. The forward primer was designed to contain the \textit{flaB}_{2\textsubscript{S2}} gene sequence flanking the desired in-frame 30 bp deletion. The reverse primer was complementary to the forward primer sequence upstream of the deletion (Table 2). After \textit{DpnI} digestion, purified PCR products were transformed into \textit{E. coli} TOP10 competent cells. Recombinant plasmids were extracted from the transformants and used as template for amplifying the \textit{flaB}_{2\textsubscript{S2}} mutant genes by PCR with the complementation primers listed in Table 2. The smaller \textit{flaB}_{2\textsubscript{S2}} mutant genes were identified by agarose gel electrophoresis of the PCR products which were also sequenced to confirm the deletion. The same protocol was used to generate the 3 amino acid deletion, \textsuperscript{61}\textit{GTA}\textsuperscript{63}, where the deletion of the 9bp resulted in the removal of an \textit{RsaI} restriction site. This allowed for the screening of the \textit{flaB}_{2\textsubscript{S2}} gene in plasmids carried by transformants for the small deletion by digestion of subsequent \textit{flaB}_{2\textsubscript{S2}} PCR products with \textit{RsaI}. Lastly, a mutant \textit{flaB}_{2\textsubscript{S2}} gene encoding a mutated \textit{FlaB}_{2\textsubscript{S2}} protein with a 10-amino acid deletion at \textsuperscript{91}\textit{TLSDGTT}KT\textsuperscript{100}V had those amino acids replaced with a copy of the 10 amino acids at \textsuperscript{161}\textit{IIVSGVS}FSDT\textsuperscript{170}, thereby generating a substitution mutant version of \textit{FlaB}_{2\textsubscript{S2}} (SUB) that was still the same length as the wildtype version. All the mutant versions of \textit{flaB}_{2\textsubscript{S2}} generated were sequenced to confirm the mutations.

**Construction of complementation vectors**

To generate complementation plasmids for \textit{M. maripaludis} S2, mutant \textit{flaB}_{2\textsubscript{S2}} genes in the \textit{pCR2.1-TOPO} vector were PCR amplified using complementation primers with an \textit{NsI} restriction site incorporated into the forward primer and an \textit{XbaI} site into the reverse primer (Table 2). After \textit{NsI} and \textit{XbaI} digestion, the PCR product was cloned into the shuttle vector \textit{pWL}G40 where transcription of the cloned gene is under the control of the strong constitutive \textit{hmw} promoter [48]. Mutant \textit{flaB}_{2\textsubscript{S2}} genes in \textit{pWL}G40 were sequenced to confirm the insert gene sequence. As a further control, plasmids were re-isolated from the complemented cells and re-sequenced.
Table 3. Primers used in this study

<table>
<thead>
<tr>
<th>Complementation primers</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>comp-F CTAGATGCAATGAAAAATAACAGAATTTTGGAAAGCAAAAGGTGCTTC</td>
<td>NsiI</td>
</tr>
<tr>
<td>comp-R ACGTCTAGAGTTATGTAATGAAACTATTTTTTCAGTGATTTGGAAG</td>
<td>XbaI</td>
</tr>
</tbody>
</table>

**SDM primers**

Ten-amino acid deletion primers

| 2-F | ATTCATGAAAGAAGCCAAAGGGTGT | |
| 2-R | AGCACCCTTTTTGCTTTTCTAATGATC | |
| 4-F | ATTCATGAAAGAAGCCAAAGGGTGT | |
| 4-R | GCCAGAAGCAACCTTTTTGCTTTTCTAATGATC | |
| 11-F | GCCAGTCTAGGCTTTTCTATTTTTTTCTGG | |
| 11-R | AAAAAAAACTAACAGTTTTTCTGGCAACAACTGCAACAAGTAC | |
| 21-F | CAAATGAGAGTTATTTCAGG | |
| 21-R | TGCAAAGTCGCAAGCAACTAATACAATGTC | |
| 31-F | CGTCTTTTAAAAATACACAAAGGCG | |
| 31-R | GAATCCGCTTTGTTTATTTTTACAGCTTG | |
| 41-F | CAAACAAAAAGCCCTTAACACTGGAATATAAGTCATGAGCGATACC | |
| 41-R | TTTAAGCTTTGTTG | |
| 51-F | TTACCCGAGAGGCTTTTCTGGCAACAACTGCAACAAGTAC | |
| 51-R | TAAACCGCTTTGCAACACTGCTGG | |
| 61-F | GGGATTTAGACG | |
| 61-R | TGCCCTGCTTCCGATAATCCATAATTTG | |
| 71-F | GCTAACCAGCAGACTACATCATAACTCTTCATACGCAGG | |
| 71-R | TAATTGCTATTGCTGGTGTAAGC | |
| 81-F | CATACCTCCTTTACGCAGGGTTAGACACTTTTCAGCAGCAGACAAACACTGGAATTTTTCAGG | |
| 81-R | ACTTCTCGCCGTTAGTAGTTATG | |
| 91-F | GCAATAGACATGGAATCTCACCATAACTCTTCATAACGCAGG | |
| 91-R | TAAACAAAACCTTTGCTTTTCTGCTATTG | |
| 101-F | CTTCCAGCGAAGCTACAACACTTACCTTGCAGTGGAAGACCTTTTAC | |
| 101-R | AACAGTTTTTTGATGGTTTCCGCTGGAAG | |
| 111-F | ACAGAATACCCGCTACACAAACAGCAACTGTAACTGTCAAATTTTAGC | |
| 111-R | GGTGTTGTATGGGCTGTAGTGTATGTTATG | |
| 121-F | GCAAGGTTGAGACCTTTACACACTACAGAATTTTTTTTGGGAATAGTTGGAATTTTAC | |
| 121-R | AGTTTTGTTAAGGTCCCTCCACCTGCGAAG | |
| 131-F | GTAACCTGCTTTGCTAAAATTACAGTGATCGAGATCTTTTCTAGTACGAAGCTTTTAC | |
| 131-R | ATCTGTCTAAATTGGATTCAGATTACAGTACATGCAGGACCTTTTAC | |
| 141-F | GAAAATGGAATAGTGGAAATCTTACAACACGTTTACAAACAAAGGGAACCTGGAATTTTTCAGG | |
| 141-R | TTGTAGTCTCTTTACACTATTCA | |
| 151-F | GATCTTTTTATTAAAGCTTTGCAAGAGTTTTTTTGGATTTTCACTTTCCATTTTACG | |
| 151-R | TGCAAGGTTGTGAAGGAGATTCGTG | |
| 161-F | CAAACAAAAAGCTTTGCAAGAGTTTTTTTGGATTTTCACTTTCCATTG | |
| 161-R | TGCAAGGTTGTGAAGGAGATTCGTG | |
| 171-F | GTAAGGCTGTTTATTCCGATAAGATTTTTTGGTCTCCAGGAAGATTTTACG | |
| 171-R | TGCTAGGGAATGGAATTTTTCCTTACCTG | |
| 181-F | GGAATTTTTACAGTCTGTGCAAGCTTTGCTAAACCACTTTCAACATCAGT | |
| 181-R | TGGCTGACAGATCCTGAAGATTTC | |
| 194-F | CATTCACCCACCTAATTCAATACACATGAAAGT | |
| 194-R | TACCTTTTCTAGTGAAATGTTTCAAGGTTGGAATTTTACG | |

(Continued)
Complementation of a ΔflaB<sub>2S2</sub> mutant using mutant flaB<sub>2S2</sub> derivatives

To determine if the mutant FlaB<sub>2S2</sub> proteins generated above could restore archaellation and motility in the ΔflaB<sub>2S2</sub> mutant, recombinant pWLG40 plasmids carrying the various mutant flaB<sub>2S2</sub> derivatives were transformed individually into the ΔflaB<sub>2S2</sub> mutant using a PEG-based method [41,50]. Transformants were cultured in Balch medium III containing 0.25 μg/mL puromycin for plasmid selection [48].

Western blot analysis of the ΔflaB<sub>2S2</sub> mutant complemented with mutant flaB<sub>2S2</sub> derivatives

Whole cell lysates of complemented cells carrying the various mutant flaB<sub>2S2</sub> genes were separated by SDS-PAGE (15% gels) and then transferred onto an Immobilon-P membrane (Millipore Inc.) [51]. Mutant FlaB<sub>2S2</sub> proteins were detected using chicken anti-FlaB<sub>2S2</sub> specific primary antibody [41]. Horseradish peroxidase-conjugated rabbit anti-chicken immunoglobulin Y (Jackson Immuno Research Laboratories) was used as secondary antibody, and the blots were developed using Immobilon Western Chemiluminescent HRP Substrate (Millipore Inc.).
Swarming assay of the ΔflaB2S2 mutant complemented with mutant flaB2S2 derivatives

Complemented ΔflaB2S2 strains carrying plasmids with mutant flaB2S2 genes encoding proteins having mutations at the various N-glycosylation sites were examined for motility using semi-solid swarm plates [22]. Briefly, the OD600 of an overnight cell culture was measured and adjusted to 1.0. Five microliters of the adjusted cell culture were inoculated onto semi-solid Balch medium containing 0.25% (w/v) agar using a micropipette in an anaerobic chamber by stabbing the tip into the agar. Plates were incubated in an anaerobic canister at 37°C for 4 or 6 days.

Electron microscopy of the ΔflaB2S2 mutant complemented with mutant flaB2S2 derivatives

Complemented M. maripaludis ΔflaB2S2 cells carrying mutant flaB2S2 genes were collected from an overnight culture by centrifugation at 20 000 g for 1 min, washed with 2% (w/v) NaCl and resuspended in phosphate-buffered saline. Resuspended cells were loaded onto carbon-Formvar-coated copper grids and stained with 2% phosphotungstic acid, pH 7.0. Grids were examined in a Hitachi 7000 electron microscope operating at an accelerating voltage of 75 kV.

Results and Discussion

Generation of mutant flaB2S2 derivatives

While deletions in genes that affect N-glycosylation are known to cause severe defects in archaellation and motility [18,22,28,29], it is not clear if the defects are related directly to the inability of non-glycosylated archaellins or archaellins glycosylated with truncated glycans to assemble into archaella, or whether the glycosylation defect affected other steps in the assembly of archaella. For example, it may be that another protein critical for assembly of archaella, but not an archaellin, must be glycosylated in order to function properly. In Hfx. volcanii H53, changing the sequence of the major archaellin flgA at any of the 3 examined sequons so that the encoded amino acid changed from Asn to Gln led to mutant forms of the protein that could not rescue the swimming defect of an flgA deletion strain, suggesting that each glycosylation site was necessary for archaellation [18]. However, this is not the case for M. maripaludis S2. Previous work in this methanogen showed that a strain that had a spontaneous mutation in flaB2S2 which led to the loss of the 2nd N-glycosylation site of the archaellin that is normally decorated with the N-linked tetrasaccharide, was, nonetheless, still archaellated and motile [43]. To examine the possible role that each N-glycosylation site, either alone or in combination with other sites, might have on archaella formation and motility in M. maripaludis S2, various mutant flaB2S2 genes whose products were lacking single to quadruple N-glycosylation sites either by Asn to Gln (N to Q) substitution, or Asn to Asp (N to D) substitution of the N-glycosylation sequon asparagine were generated and cloned into the complementation vector pWLG40 (Table 2). For these mutant constructs we used D or Q followed by a number to indicate that the change was N to D or N to Q with the number representing the site changed, i.e. Q1 indicates mutant FlaB2S2 with an N to Q substitution at the 1st N-glycosylation site.

In addition, two other mutant genes, designated G6 and G10 (Fig. 1), whose products contain extra N-glycosylation sequons were generated and cloned into pWLG40. The G6 sequence encodes the wildtype FlaB2ARC protein (i.e. FlaB2ARC from M. maripaludis ΔRC). FlaB2ARC and FlaB2s2 share 95% identity, with the differences almost exclusively confined to several N-glycosylation sites (Fig. 1). Compared to FlaB2S2, FlaB2ARC shares 3 sites, plus it has 3 additional sequons, 60NGS62, 104NIS106, 129NDT131, but it is missing the 2nd N-glycosylation site 110NLT112 in FlaB2S2. The G10 protein has all the sites present in FlaB2S2 and FlaB2ARC, with 3
additional sites created in the hypervariable region at sites requiring only minimal amino acid changes to generate a total of 10 possible sites (Fig. 1). G6 and G10 both have the $^{26}$NIS$^{28}$ sequon in the N-terminal conserved region that is not occupied with N-glycan in FlaB$_{S2}$. After transformation of these recombinant plasmid pWLG40-flaB$_{S2}$ mutants into a ΔflaB$_{S2}$ mutant, the complemented strains were examined for expression of the mutant FlaB$_{S2}$ proteins, archaella formation and cell motility.

**Western blot analysis of the ΔflaB$_{S2}$ strain complemented with flaB$_{S2}$ derivatives containing mutations at N-glycosylation sites**

Western blots were run to detect the expression and stability of the various mutant versions of FlaB$_{S2}$ in the complemented ΔflaB$_{S2}$ mutant. As shown in Fig. 2, all mutant versions of FlaB$_{S2}$ except G10 were successfully expressed in the ΔflaB$_{S2}$ mutant. All mutant FlaB$_{S2}$ proteins were expressed in similar amounts and all appeared stable as judged by the general lack of any cross-reacting smaller molecular mass bands which could be indicative of protein degradation. The amount of the G10 version of FlaB$_{S2}$ detected in western blots was very low and could be only observed when blots were overexposed (data not shown). We have found previously that cells carrying mutations in any gene that prevents assembly of archaella (as in the ΔflaB$_{S2}$ mutant) often stop transcribing the fla operon after several sub-cultures in the laboratory. This then makes the complementation of the original gene deletion back to an archaellated state impossible [22]. For this reason, the presence of FlaE, whose gene is a downstream member of the fla operon, was also confirmed by western blot to ensure that the fla operon was still transcribed in the ΔflaB$_{S2}$ mutant during the course of the complementation experiments (data not shown) [41].

In general, mutant FlaB$_{S2}$ proteins missing N-glycosylation sites all had a smaller apparent molecular mass than that of wildtype FlaB$_{S2}$ when examined by western blotting, with the
possible exceptions of D4 and Q4 which ran at very close to wildtype size. The greater the number of N-glycosylation sites eliminated in a particular FlaB2S2 mutant, the faster the mutant proteins migrated, i.e., single-site mutants had the largest apparent molecular mass, and the quadruple-site mutant D1,2,3,4 had the smallest. However, the 8 single-site mutants did not migrate as proteins of the same apparent molecular mass. Of the 8 mutants, D1 and Q1, both of which had the 1st N-glycosylation site eliminated, had the smallest apparent molecular mass, while D4 and Q4 had the largest. Similar results were observed from the double-site and triple-site mutations. In the 6 double-site mutations, FlaB2S2 with D1,2 and D1,3 sites eliminated had the smallest apparent molecular mass, followed by FlaB2S2 with D1,4 and D2,3 sites eliminated, while the archaellin having the D2,4 and D3,4 sites eliminated migrated with the largest apparent molecular mass. In the triple-site mutants, FlaB2S2 with any of the 1st N-glycosylation site eliminated (D1,2,3, D1,2,4 and D1,3,4) migrated at the same apparent molecular mass while FlaB2S2 with the other triple combination of sites eliminated (D2,3,4) migrated more slowly.

One possible explanation for the observed different electrophoretic mobilities is that mutant FlaB2S2 proteins lacking the same number of N-glycosylation sites have the same number of N-glycans attached but the attachment of N-glycan on some sequons might have effects on the local protein structure so that the glycoprotein is not able to be totally denatured by SDS, thus resulting in an unusual migration pattern. This unusual electrophoretic mobility has been observed in other similar studies and been the suggested explanation. Human erythropoietin (Epo) has 3 N-glycosylation sites, and the 3 single-site mutants generated by N to Q SDM showed uneven migrations on western blot, although all of the 3 mutants had the same theoretical molecular mass but differed only in the position of the N-glycans [52]. Similar uneven migration was also observed in the 4 single-site mutations of hepatitis C virus envelope protein E1 each missing one N-glycosylation site [53].

While local folding effects might explain the altered electrophoretic mobility, another possible explanation for this unusual western blot result is that elimination of the 1st N-glycosylation site might interfere with the cell’s ability to N-glycosylate the remaining sites, resulting in FlaB2S2 where not all the remaining sequons are actually occupied (and so run as smaller molecular mass proteins). On the other hand, it may be that elimination of one or more glycosylation sites results in the attachment of glycan to the sequon 26NTS28 that is not glycosylated under our usual growth conditions. This could explain why, for example, the FlaB2S2 mutants that are missing the 4th N-glycosylation site (D4 and Q4) showed a larger apparent molecular mass than the other single mutants and close to wildtype size. If D4 and Q4, missing the 4th glycosylation site, now had N-glycan attached at the normally unused sequon, these mutant proteins would have an identical mass as the wildtype. There is precedent for glycosylation at one sequon influencing what happens at distant sites. For example, it has been reported for rabies virus glycoprotein that N-glycosylation at one sequon can influence processing of the N-glycans at a different site on the protein [54]. In M. voltae PS, the archaella are composed of 4 archaellins FlaA_Mv, FlaB1_Mv, FlaB2_Mv and FlaB3_Mv, which share high sequence similarity in the N-terminal conserved region with FlaB1S2 and FlaB2S2 [41,55]. Interestingly, the first 40 amino acids in the mature FlaB1_Mv and FlaB2_Mv, including the 26NTS28 sequon, are identical to those of FlaB2S2, but in the case of the M. voltae PS archaellins, the 26NTS28 sequon was found to be occupied with N-linked glycan [12]. Clearly, this region of the archaellin can be glycosylated and possibly the 26NTS28 sequon in FlaB2S2 might be able to be N-glycosylated in M. maripaludis S2 under different conditions.

The G6 mutant FlaB2S2 protein (identical to the FlaB2ARC) with additional glycosylation sequons compared to wildtype FlaB2S2 had a larger apparent molecular mass (~32 kDa) than wildtype FlaB2S2 (~27 kDa) in western blots. Since G6 has 6 N-glycosylation sites (excluding the 26NTS28 sequon), its larger apparent molecular mass suggests that at least some, and
possibly all, of the extra sequons are, in fact, occupied since the mass of the tetrasaccharide is only 1036 Da [11]. Although the expression level of G10 was extremely low, on over-exposed western blots, the apparent molecular mass (37kDa) of this “artificially designed” glycoprotein was even larger than G6, indicating that AglB can recognize and transfer glycan to at least some of the newly introduced sequons in the hypervariable region.

Electron microscopy of the ΔflaB2s2 strain complemented with flaB2s2 derivatives containing mutations at N-glycosylation sites

Complemented cells carrying FlaB2s2 proteins having mutations at N-glycosylation sites were examined by transmission electron microscopy for archaellation, and the results are listed in Table 4. The majority of the complementations either restored archaellation to essentially all cells or were unable to restore archaellation to any cells. In only a couple of cases did the complementation lead to a population which contained roughly equal number of archaellated and non-archaellated cells (D1,2 and D3,4). Fig. 3 shows electron microscopy pictures of a number of selected complements (Q2, Q4, D2, D3, D4, D1,3, D2,4, D1,2,4, D2,3,4, D1,2,3,4, G6 and G10).

The archaellation state of the 4 control strains was as expected. Wildtype Mm900 cells (900) were archaellated while the ΔflaB2s2 mutant strain was non-archaellated. Archaela were

<table>
<thead>
<tr>
<th>Complements</th>
<th>Archaellationa</th>
<th>Motility</th>
</tr>
</thead>
<tbody>
<tr>
<td>Controls</td>
<td></td>
<td></td>
</tr>
<tr>
<td>WT</td>
<td>++ (100%)</td>
<td>++</td>
</tr>
<tr>
<td>Blank</td>
<td>- (0%)</td>
<td>-</td>
</tr>
<tr>
<td>NQ single</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Q1</td>
<td>- (3%)b</td>
<td>-</td>
</tr>
<tr>
<td>Q2</td>
<td>++ (100%)</td>
<td>++</td>
</tr>
<tr>
<td>Q3</td>
<td>- (3%)b</td>
<td>-</td>
</tr>
<tr>
<td>Q4</td>
<td>- (0%)</td>
<td>-</td>
</tr>
<tr>
<td>ND single</td>
<td></td>
<td></td>
</tr>
<tr>
<td>D1</td>
<td>- (3%)b</td>
<td>-</td>
</tr>
<tr>
<td>D2</td>
<td>++ (100%)</td>
<td>++</td>
</tr>
<tr>
<td>D3</td>
<td>- (0%)</td>
<td>-</td>
</tr>
<tr>
<td>D4</td>
<td>++ (93%)</td>
<td>-</td>
</tr>
<tr>
<td>ND double</td>
<td></td>
<td></td>
</tr>
<tr>
<td>D1,2</td>
<td>+ (47%)</td>
<td>+</td>
</tr>
<tr>
<td>D1,3</td>
<td>- (0%)</td>
<td>-</td>
</tr>
<tr>
<td>D1,4</td>
<td>++ (97%)</td>
<td>++</td>
</tr>
<tr>
<td>D2,3</td>
<td>++ (100%)</td>
<td>++</td>
</tr>
<tr>
<td>D2,4</td>
<td>++ (100%)</td>
<td>+</td>
</tr>
<tr>
<td>D3,4</td>
<td>+ (60%)</td>
<td>+</td>
</tr>
<tr>
<td>ND triple</td>
<td></td>
<td></td>
</tr>
<tr>
<td>D1,2,3</td>
<td>- (0%)</td>
<td>-</td>
</tr>
<tr>
<td>D1,2,4</td>
<td>- (0%)</td>
<td>-</td>
</tr>
<tr>
<td>D1,3,4</td>
<td>- (0%)</td>
<td>-</td>
</tr>
<tr>
<td>D2,3,4</td>
<td>++ (100%)</td>
<td>+</td>
</tr>
<tr>
<td>ND quadruple</td>
<td></td>
<td></td>
</tr>
<tr>
<td>D1,2,3,4</td>
<td>- (0%)</td>
<td>-</td>
</tr>
<tr>
<td>Additional sequons</td>
<td>G6</td>
<td>++ (100%)</td>
</tr>
<tr>
<td></td>
<td>G10</td>
<td>- (0%)</td>
</tr>
</tbody>
</table>

For each strain a minimum of 30 random cells were assessed for the presence or absence of archaella. Values in parentheses describe the percentage of cells with observable archaella.

Only a rare cell was observed with archaella, typically very few in number and abnormally short.

doi:10.1371/journal.pone.0116402.t004
observed when the ΔflaB2S2 strain was complemented with the wildtype version of flaB2S2, but not when the ΔflaB2S2 strain was complemented with the empty vector pWLG40.

In the 8 single-site mutation complements, 3 mutant versions of FlaB2S2 (Q2, D2 and D4), could restore archaellation. In the 3 archaellated complements, Q2 and D2 had different amino acid substitutions at the same 2nd N-glycosylation site. The N to D amino acid change at the 2nd sequon generated in this study replicates the spontaneous mutation in flaB2S2 that we previously reported [43]. Both that spontaneous mutant and the complemented cells carrying the mutant D2 gene generated in this study showed no impairment in archaellation or swarming motility [43]. These results indicate that missing the 2nd N-glycosylation site alone does not significantly interfere with archaellation. In contrast, complementation of the ΔflaB2S2 strain with flaB2S2 lacking the 4th N-glycosylation site differed markedly depending on what amino acid the original Asn was changed to; cells complemented with the D4 version had archaella under electron microscopy while cells complemented with the Q4 version did not (Fig. 3). None of the other single-site mutation complements (Q1, Q3, Q4, D1 and D3 (Fig. 3)) were considered archaellated, although in each of the Q1, Q3 and D1 complementations a rare cell with short archaella was observed (Table 4).

Among the 6 double-site mutation complements, all but the D1,3 version could assemble archaella (electron micrographs of D1,3 and D2,4 complemented cells are shown in Fig. 3). This was surprising since many of these double-site mutants contained eliminated sites which if deleted alone resulted in non-archaellated cells. In the 4 triple-site mutation complements, 3 complements D1,2,3, D1,2,4 and D1,3,4 were non-archaellated, while archaella could be assembled in D2,3,4 (Fig. 3). However, when all glycosylation sites were eliminated in the complementing version of flaB2S2 (D1,2,3,4), the ΔflaB2S2 cells were not able to assemble archaella. These results indicate that in M. maripaludis S2, archaella can be assembled using FlaB2S2 lacking as many as 3 out of the 4 glycosylation sites, as long as the first site remained intact (D2,3,4) but not when the archaellin is entirely non-glycosylated (D1,2,3,4).

In the two complementations where new sequons were introduced into FlaB2S2, different results were observed. In the G6 complemented cells, the ΔflaB2S2 strain were now archaellated, suggesting that the FlaB2S2 protein with extra N-glycan modifications in the hypervariable region could be incorporated into the archaellar filament by the archaella assembly apparatus in M. maripaludis S2. This was not unexpected since this version of FlaB2 already exists naturally in the archaellated M. maripaludis strain ΔRC. In contrast, no archaella were observed on the ΔflaB2S2 strain complemented with the G10 version. The G10 version had extra glycosylation sequons added to the internal hypervariable region of the protein. While this protein appeared to be modified at, at least, some of these additional sequons with glycan, judging from its higher apparent molecular mass in western blots, it was very poorly expressed in the cells under our normal growth conditions and this low expression may explain the lack of archaella observed by electron microscopy (Fig. 3).

Swarming assays of complements with mutant flaB2S2 derivatives containing mutations at N-glycosylation sites

In addition to the restoration of archaellation, the complemented cells were also examined for possible restoration of motility using semi-solid agar plates (Fig. 4A). Motility assay results are summarized in Table 4, which also incorporates the archaellation status of the complemented strains for comparison.

In general, complemented cells in which archaellation was restored were also motile on swarming plates and all of the complemented cells that were non-archaellated as determined by electron microscopy (D1, D3, Q1, Q3, D1,3, D1,2,3, D1,2,4, D1,3,4, D1,2,3,4 and G10) were
also non-motile on swarm plates, even after an extra 2-day incubation (Fig. 4B). The unusual exception was the complementation with the D4 version of FlaB2s2 which was archaellated but non-motile (Fig. 3, Fig. 4A). However, among the motile complemented cells, the swarming diameter was not always returned to the wildtype level. Cells complemented with the flaB2s2 genes carrying D2, Q2, D1,4, D2,3 and D2,3,4, mutations swarmed out to a similar distance on semi-solid agar (swarming diameter of D2/ WT = 1.01±0.11, Q2/WT = 1.10±0.09, D1,4/WT =

Fig 3. Transmission electron micrographs of select ΔflaB2s2 strain complemented with flaB2s2 with mutations at various N-glycosylation sites.

Wildtype Mm900 cell (900) was archaellated, while the ΔflaB2s2 mutant was not. Wildtype FlaB2s2 protein expressed in ΔflaB2s2 restored archaellation (WT), but the empty vector could not. Archaella were observed on surface of Q2, D2, D4, D2,4, D2,3,4 and G6 complemented cells. Cells complemented with Q4, D3, D1,3, D1,2,4, D1,2,3,4 and G10 were non-archaellated. Archaella are indicated by arrows. Bar equals 500 nm.

doi:10.1371/journal.pone.0116402.g003
0.93±0.14, D2,3/WT = 0.96±0.20, D2,3,4/WT = 1.08±0.22), while complements with the flaB2S2 genes carrying D1,2, D2,4 and D3,4 had a smaller swarming diameter (swarming diameter of D1,2/WT = 0.51±0.05, D2,4/WT = 0.68±0.13, D3,4/WT = 0.52±0.07). For the D1,2 and D3,4 complementations, the smaller swarming distance may be explained by the lower percentage of cells that were observed by electron microscopy to be archaellated, but in the case of the D2,4 all cells examined were archaellated. Interestingly, G6 appeared hyper-motile as it consistently swarmed further than cells complemented with the wildtype version of flaB2S2 (swarming diameter of G6/WT = 1.24±0.10).

In Hfx. volcanii H53, none of the 3 FlgA single-site mutation complements showed motility [18]. However, in this study, M. maripaludis S2 cells were still as motile as wildtype cells when the ΔflaB2S2 strains was complemented with flaB2S2 with the D2,3,4 changes, in which archaella were assembled using FlaB2S2 lacking 3 out of the 4 N-glycosylation sites. The structural protein (flagellin) of the functionally analogous bacterial swimming organelle, the flagellum, can also be modified with glycan, especially in Gram-negative bacteria, although the linkage is O-glycosidic rather than N-glycosidic [56,57]. The O-glycan modification in bacterial flagellin can be critical for flagella assembly, stabilization, motility, and even virulence in pathogens [56,58–60]. In Pseudomonas syringae pv. tabaci, flagellin FliC has 6 O-glycosylation sites, and single-site mutations in any of these sites resulted in various impairments in motility, while a mutant carrying mutations to eliminate all 6 O-glycosylation sites in FliC was non-motile [59]. The structural protein (pilin) from bacterial type-IV pili, structures which share several significant similarities with archaella [10,34], can also be O-glycosylated [61,62]. Elimination of O-glycosylation of type IV pilin resulted in reduced twitching motility in Pseudomonas aeruginosa 1244 and P. syringae pv. tabaci but did not interfere with pilin assembly [62,63]. However, in P. aeruginosa 5196 in which a different O-glycan was attached to the type IV pilin PilA, O-glycosylation played critical roles in both type IV pilin assembly and twitching motility [2,64].

**Fig 4. Swarming plates showing the motility of the ΔflaB2S2 strain complemented with flaB2S2 with mutations at various N-glycosylation sites.** A. Plates were incubated at 37°C for 4 days. B. Complemented cells that did not show motility or showed poor motility after 4 days incubation were incubated a further 2 days.

doi:10.1371/journal.pone.0116402.g004
The results obtained with the G6 complemented cells indicate that an increase in glycosylation can lead to hyper-motility. The western blots results indicate that the G6 version of FlaB2 is hyper-glycosylated compared to the wildtype FlaB2Δ140 version. Of all the complemented strains, only the G6 complement consistently demonstrated an increased zone of swarming compared to the wildtype. Interestingly, similar results were also observed in regards the O-glycosylation of flagellin in *Helicobacter pylori* [65] where O-glycosylation of the flagellins FlmA and Flab with pseudaminic acid is essential for flagella assembly and cell motility [66,67]. A *H. pylori* mutant defective in deglycosylation of flagellins showed both hyper-O-glycosylation (3 fold more pseudominic acid) of FlmA as well as hyper-motility [65]. However, there is a limit to how many extra sequons can be added to archaellins since archaellin synthesis was very poor in the G10 complemented cells, even though the small amount of the G10 version detected was apparently modified at, at least, some of the extra sites.

The data obtained from the glycosylation site elimination mutants indicates that while no particular single site of glycosylation on FlaB2Δ21 is essential, nonetheless glycosylation at some site is necessary for archaella formation. In the case of the 4 triple-site mutants it is clear that glycosylation of only site 1 is sufficient for archaellation and motility. However, removal of the 1st site did not always lead to non-archaellated cells as witnessed by the archaellated and motile cells observed in the D1,4 complementation, suggesting that glycosylation of FlaB2Δ21 at several different combinations of sites could be sufficient for incorporation of the subunits into functional archaella. In some ways, this is reminiscent of a situation in Wzc, a tyrosine autokinase essential for capsule formation in *E. coli*. Phosphorylation of tyrosine residues in the C-terminus of Wzc are necessary for its function but no single tyrosine is essential for phosphorylation and it was suggested that the overall level of phosphorylation rather than a precise combination of tyrosine residues accessible to phosphorylation is what is important for Wzc activity [68].

**Western blot analysis of the ΔflaB2Δ21 strain complemented with flaB2Δ21 scanning deletions**

To determine which regions of the FlaB2Δ21 protein are critical for archaella formation, a series of FlaB2Δ21 scanning deletion mutants that sequentially lacked 10 amino acids were generated in the complementation vector pWLG40 and transformed into the ΔflaB2Δ21 mutant. The scanning deletions in flaB2Δ21 were identified since they migrated slightly faster than the wildtype version of flaB2Δ21 in 0.8% agarose gels due to the 30 bp deletion (Fig. 5 shows an example for screening of ΔA31–40). For the first 10 amino acids, two versions were generated. The first was deleted for amino acids 2–10 (named as Δ2–10), leaving the +1 amino acid which we thought might be important for successful cleavage of the 12 amino acid signal peptide. We also generated a 4–10 amino acid deletion (named as Δ4–10) since the +3 glycine of the mature protein is needed for signal peptide removal in archaellins of the related methanogen *M. voltae* PS [69].

Mutant FlaB2Δ21 proteins from whole cell lysates of the various complemented cells were detected using anti-FlaB2Δ21 specific antibody on western blot analysis, as shown in Fig. 6. In the 21 FlaB2Δ21 scanning deletions, 19 mutant proteins (all, except Δ4–10 and Δ11–20) were readily detected on western blot by anti-FlaB2Δ21 specific antibody, although the expression level of Δ2–10 and Δ21–31 was relatively lower compared to that of the other mutants. Evidence of some possible protein degradation was observed in the Δ2–10 and Δ161–170 FlaB2Δ21 as multiple lower molecular mass bands were detected in these two lanes. These results indicate that some of the mutant proteins were either not expressed or were unstable and degraded. Fourteen mutant proteins, including Δ21–30, Δ31–40, Δ41–50, Δ51–60, Δ71–80, Δ81–90, Δ91–100, Δ131–140, Δ141–150, Δ151–160, Δ161–170, Δ171–180, Δ181–190 and Δ194–204, had similar apparent molecular masses, which were smaller than FlaB2Δ21 from the wildtype strain Mm900 (900)
Fig 5. PCR screening of the Δ31–40 scanning deletion. Following the deletion procedure, the flaB2\textsubscript{S2} gene was amplified using flaB2\textsubscript{S2} complementation primers and the PCR product analyzed by agarose gel electrophoresis along with the amplification product obtained with the wildtype flaB2\textsubscript{S2} gene using the same primers. The scanning deletion is readily distinguished from wildtype flaB2\textsubscript{S2} by the faster migration of its 30 bp smaller PCR product M: 100 bp DNA ladder; flaB2\textsubscript{S2}: PCR products using pKJ902 as template; Δ31–40: PCR products using plasmid isolated from one colony of the Δ31–40 transformants as template.

doi:10.1371/journal.pone.0116402.g005

Fig 6. Western blot analysis of the ΔflaB2\textsubscript{S2} strain complemented with flaB2\textsubscript{S2} scanning deletions. Except for Δ4–10 and Δ11–20, all the FlaB2\textsubscript{S2} scanning deletion proteins were expressed, although the expression level of Δ2–10 and Δ21–30 was relatively low. FlaB2\textsubscript{S2} scanning deletion proteins Δ61–70, Δ101–110, Δ111–120 whose 10-amino acid deletion contains an N-glycosylation site (shown in red) had smaller apparent molecular masses due to the loss of N-glycan usually attached at this site. Δ121–130 missing the 4\textsuperscript{th} N-glycosylation site (shown in red) had unusual bigger apparent molecular mass than the other 3 mutants also missing N-glycosylation site.

doi:10.1371/journal.pone.0116402.g006
or from the ΔflaB2Δ strain complemented with the wildtype version of flaB2Δ. This is consistent with the fact that all these mutation proteins are 10 amino acids shorter than wildtype FlaB2Δ. The upper band in the mutant Δ2–10 lane appears slightly larger than the neighboring WT lane. This might be due to the lack of processing of the archaellin signal peptide in this deletion since in M. voltae PS the +3 glycine position was essential for cleavage of the signal peptide by the pre-archaellin peptidase FlaK and this is missing in the Δ2–10 version of FlaB2Δ [69]. The presence of the signal peptide does not prevent the attachment of N-glycans [37] so this version of FlaB2Δ would be expected to have a full complement of attached N-glycans as well as the extra amino acids of the signal peptide contributing to make it run slightly larger in the western blots than the processed wildtype version of FlaB2Δ.

The FlaB2Δ mutants Δ61–70, Δ101–110 and Δ111–120 had even smaller apparent molecular masses compared to the FlaB2Δ carrying other scanning deletions. This can be attributed to the fact that these three deletions result in the loss of one N-glycosylation sequon. Surprisingly, the scanning deletion that contains the remaining, 4th, glycosylation sequon, namely Δ121–131, does not follow this pattern. In this lone case, FlaB2Δ migration in western blots is slower than the other three deletion mutants missing N-glycosylation sites. This observation is consistent with the results of the N-glycosylation sequon mutation complementations. In the case of both single 4th sequon mutants (D4 and Q4), FlaB2Δ had a larger apparent molecular mass than the other single-site mutations (Fig. 2), suggesting either that the lack of glycosylation at this site has unusual effects on the migration of FlaB2Δ in western blots or that when the 4th glycosylation site is not available, the normally unoccupied Δ6NTSΔ is now decorated with glycan, adding to the molecular mass. Further studies are necessary to confirm the glycosylation status of this Δ6NTSΔ sequon in FlaB2Δ in these mutant proteins.

Electron microscopy of the ΔflaB2Δ strain complemented with FlaB2Δ scanning deletions

To examine if ΔflaB2Δ cells could assemble archaella after being complemented with any of the FlaB2Δ scanning deletions, cells from each complementation were observed by transmission electron microscopy for the presence of archaella. All of the 21 10-amino acid scanning deletion complements were found to remain non-archaellated (data not shown) even though most of them produced FlaB2Δ detected by western blot. These results suggested that either all regions of the molecule were essential for archaella formation or that a certain critical archaellin length is important for the archaella filament to be assembled.

Since none of the FlaB2Δ scanning deletion mutants could restore archaellation, we suspected that the 10-amino acid deletion in the scanning deletions was too long for FlaB2Δ to be assembled into archaella. To address this, we created a shorter 3-amino acid deletion in the hypervariable region of FlaB2Δ. This protein was detected by western blot at a similar apparent molecular mass as the wildtype FlaBΔ and this is missing in the Δ2–10 version of FlaB2Δ [69]. The presence of the signal peptide does not prevent the attachment of N-glycans [37] so this version of FlaB2Δ would be expected to have a full complement of attached N-glycans as well as the extra amino acids of the signal peptide contributing to make it run slightly larger in the western blots than the processed wildtype version of FlaB2Δ.

The FlaB2Δ mutants Δ61–70, Δ101–110 and Δ111–120 had even smaller apparent molecular masses compared to the FlaB2Δ carrying other scanning deletions. This can be attributed to the fact that these three deletions result in the loss of one N-glycosylation sequon. Surprisingly, the scanning deletion that contains the remaining, 4th, glycosylation sequon, namely Δ121–131, does not follow this pattern. In this lone case, FlaB2Δ migration in western blots is slower than the other three deletion mutants missing N-glycosylation sites. This observation is consistent with the results of the N-glycosylation sequon mutation complementations. In the case of both single 4th sequon mutants (D4 and Q4), FlaB2Δ had a larger apparent molecular mass than the other single-site mutations (Fig. 2), suggesting either that the lack of glycosylation at this site has unusual effects on the migration of FlaB2Δ in western blots or that when the 4th glycosylation site is not available, the normally unoccupied Δ6NTSΔ is now decorated with glycan, adding to the molecular mass. Further studies are necessary to confirm the glycosylation status of this Δ6NTSΔ sequon in FlaB2Δ in these mutant proteins.
sequons to minimally reduce the effects from disruption of the conserved regions that might be involved in subunit-subunit interaction or in glycosylation. The SUB protein was expressed in the complement cells and showed similar apparent molecular mass as that of WT protein, as expected (Fig. 7A). However, the SUB protein complement could not restore archaellation either (Fig. 7B), suggesting that the particular 10-amino acid sequence TLSDGTTKTV is critical for archaella assembly, despite the fact it is located in a hypervariable region of the molecule. We had anticipated that at least some of the scanning deletions covering the hypervariable region may have been tolerated and allow for formation of archaella while those located in the conserved N-terminus believed to critical for subunit-subunit interactions in the filament would not be tolerated [70,71]. It is known that in the case of bacterial flagellins large internal deletions can be accepted; for instance in *E. coli*, the 493 amino acid flagellin can be reduced by internal deletions so that only the N-terminal 193 residues and the 117 C-terminal amino acids are required for filament formation [72]. In addition, sequences in the internal hypervariable region of bacterial flagellin can be replaced with completely unrelated sequences [73]. This is also true for archaellins in *Halobacterium salinarum* where both FLAG (8 amino acid peptide) and a gold-binding 12 amino acid peptide have been inserted into variable regions of different archaellins and these mutant proteins were still able to be assembled into archaella [74]. However, for type IV pilins, it has been shown in a number of studies that very small changes at key amino acids in the major pilins can result in instability of the pilins and ones that cannot assemble into pili [75–77].

Archaella are unique swimming organelles that are thought to be assembled like bacterial type IV pili but function like bacterial flagella by filament rotation [10,33,34,78]. So far little is known about details of the incorporation of individual archaellins into the archaella filament. N-glycosylation seems to be a common modification of archaellins, but the relationship between N-glycosylation and archaella assembly is unclear [1,40]. In this study, we investigated the effects of eliminating potential N-glycosylation sites as well as scanning deletions of the archaellin FlaB2s2 on archaella assembly and function in *M. maripaludis* S2. In *M. maripaludis* S2, functional archaella can be assembled using FlaB2s2 lacking as many as 3 out of 4 glycosylation sites (D2,3,4), but not when the archaellin is entirely non-glycosylated (D1,2,3,4). A hyper-N-glycosylated version of FlaB2s2 (G6) resulted in hyper-motile *M. maripaludis* S2 cells. Attempts to define essential and nonessential domains of the archaellin by scanning deletion analysis revealed that no contiguous 10 amino acid stretch could be deleted and still have the archaellin complement a ΔflaB2s2 strain back to an archaellated phenotype.
Author Contributions
Conceived and designed the experiments: YD SIA JPJC KFJ. Performed the experiments: YD KU SIA KM AB CK KFJ. Analyzed the data: YD KU SIA KM AB CK KFJ. Contributed reagents/materials/analysis tools: SIA CK JPJC KFJ. Wrote the paper: YD KFJ.

References


