Phylogenetic distribution of the euryarchaeal archaellum regulator EarA and complementation of a *Methanococcus maripaludis* ΔearA mutant with heterologous earA homologues

Yan Ding,¹ Alison Berezuk,² Cezar M. Khursigara² and Ken F. Jarrell¹,*

**Abstract**

Archaella are the swimming organelles in the Archaea. Recently, the first archaellum regulator in the Euryarchaeota, EarA<sub>Mma</sub>, was identified in *Methanococcus maripaludis*, one of the model organisms used for archaellum studies. EarA<sub>Mma</sub> binds to 6 bp consensus sequences upstream of the fla promoter to activate the transcription of the fla operon, which encodes most of the proteins required for archaella synthesis. In this study, synteny analysis showed that earA homologues are widely distributed in the phylum of Euryarchaeota, with the notable exception of extreme halophiles. We classified Euryarchaeota species containing earA homologues into five classes based on the genomic location of the earA genes relative to fla and chemotaxis operons. EarA homologues from *Methanococcus vannielii*, *Methanothermococcus thermolithothrophicus* and *Methanocaldococcus jannaschii* successfully complemented the function of EarA<sub>Mma</sub> in a ΔearA<sub>Mma</sub> mutant, demonstrated by the restoration of FlaB2 expression in Western blot analysis and the appearance of archaella on the cell surface in complemented cells. Furthermore, the 6 bp consensus sequence was also found in the fla promoter region in these methanogens, indicating that the EarA homologues use a similar mechanism to activate transcription of the fla operons in their own hosts. Attempts to demonstrate complementation of the function of EarA<sub>Mma</sub> in a ΔearA<sub>Mma</sub> mutant by the EarA homologue of *Pyrococcus furiosus* were unsuccessful, despite the presence of a copy of the 6 bp consensus EarA-binding sequence upstream of the fla promoter in the *P. furiosus* genome.

**INTRODUCTION**

The swimming organelle now designated as the archaellum is the best-studied surface appendage in the third domain of life, the Archaea [1–3]. Genes encoding most of the archaellum structural components are often located in a single fla operon [4]. The fla operon is widespread in the Archaea, having been identified in three major phyla, namely Euryarchaeota, Crenarchaeota and Thaumarchaeota [3, 5, 6]. The fla operon typically begins with archaellin genes encoding structural proteins composing the archaellum filament, followed by the fla-associated genes flaC-J in Euryarchaeota or flaX and flaG-J in Crenarchaeota [2]. Early work identified FlaI and Flai as bacterial type IV pilus system homologues, with Flai being a polytopic membrane protein (homologue of PilC in the *Pseudomonas aeruginosa* type IV pilus system) and Flai being an ATPase (homologue of PilB/PilT in *P. aeruginosa*) [7, 8]. Subsequent insights into the functions of Flai and many of the other Fla-associated proteins have been provided, mainly in *Sulfolobus acidocaldarius*, by the Sonja Albers group. Flai plays a unique role in providing the energy for both the assembly of the archaellum (analogous to the role of PilB in *P. aeruginosa* type IV pilus systems) and for archaellum rotation [9]. FlaH has ATP-binding, but not ATP-hydrolysing, properties and may regulate the activity of Flai [2, 10, 11]. *In vitro*, FlaH forms a ring within the ring formed by FlaX [10]. FlaX, FlaH and Flai interact *in vitro* and are likely to be the core of the cytoplasmic motor complex [12]. Flai is the stator for the archaellum, anchoring it to the S-layer [13]. The pre-archaellin peptidase Flak/PibD [14–17] cleaves the signal peptide from archaellin subunits in a step that is critical for archaellum assembly [14, 18]. This gene is typically, though not always, located at a site distinct from the fla operon [4].

Current knowledge about the regulation of the fla operon at the molecular level in the Archaea is extremely limited. In the crenarchaeon *S. acidocaldarius*, several regulators, some...
specifically controlling the fla operon with others working as global regulators, have been identified. In this thermoacidophile, a series of both activators and repressors has been identified in the archaellum regulatory network (Arn proteins) and the key involvement of phosphorylation in motility regulation has been demonstrated [2]. Archaea formation is highly upregulated under starvation conditions [19]. If growth conditions are favourable, ArnA (containing a forkhead-associated domain) and ArnB (containing a von Willebrand domain) form a phosphorylation-dependent complex that binds to the flaB promoter, leading to repression [20]. Eukaryotic protein kinase (ePK) ArnC phosphorylates both ArnA and ArnB while another ePK, ArnD, only phosphorylates ArnB [20, 21]. Further evidence for the importance of phosphorylation in archaellum regulation was demonstrated when deletion of the gene saci_PP2A encoding a serine/threonine phosphatase was found to result in hypermotility, identifying the phosphatase as a repressor for archaella production [22]. As cells encounter starvation conditions, however, a one-component regulator, ArnR, is expressed which acts as a positive regulator for the archaellum operon, binding to specific motifs within the flaB promoter. Unique to S. acidocaldarius, a gene duplication of arnR occurs downstream of flaI resulting in the arnR paralogue arnR1 encoding a gene product that appears to be involved in fine-tuning the expression of flaB. The deletion of arnR leads to complete inhibition of expression of FlaB, while deletion of arnR1 has a much less severe effect [23]. Recently, another serine/threonine kinase involved in the regulation of archaella in S. acidocaldarius was identified and designated ArnS, located downstream of arnR [24]. Interestingly, deletion of arnS leads to reduced motility, even though archaella are apparently fully assembled. Data indicate that ArnS paradoxically inhibits arnR transcription yet promotes ArnR translation. In addition to these components, yet another regulator was identified. Archael biofilm regulator 1 (AbfR1) was also shown to be involved in archaellum regulation as well as biofilm formation [25, 26]. Deletion of abfR1 resulted in impaired synthesis of Fla proteins and decreased motility but increased biofilm production, indicating that it is a positive regulator for archaellation while also suggesting that it is involved in the transition between the planktonic and biofilm states [25, 26]. AbfR1 exhibits high affinity binding to the flaB promoter.

Recently, we identified the first euryarchaeal archaellum regulator, EarA from Methanococcus maripaludis, and demonstrated that it acts as a transcriptional activator essential for the transcription of the fla operon [27]. EarA is 110 aa in length and contains a helix-turn-helix DNA binding domain as the only identified protein domain. It is not homologous to any of the known crenarchaeal Arn proteins. Deletion of earA in M. maripaludis prevented transcription of the fla operon, as monitored using flaB as the reporter gene; flaB2 is the second gene in the fla operon and encodes one of the major archaeplasts [28]. Furthermore, four 6 bp consensus sequences [TACATA located immediately upstream of the transcription factor B-recognition element (BRE) and TATA box] were found and confirmed in electrophoretic mobility shift assay (EMSA) experiments to be EarA binding sites [27]. Elimination of the four EarA binding sites in vivo abolished the transcription of the fla operon, indicating that at least one of the four binding sites is critical for EarA-dependent transcription [27].

In this contribution we have further investigated the EarA regulator with synten analysis and show a wide distribution of this transcriptional regulator in the Euryarchaeota phylum. Four EarA homologues, three from methanogens Methanococcus vannielii, Methanothermococcus thermolithotrophicus and Methanocaldococcus jannaschii, and one from the hyperthermophile Pyrococcus furiosus, were tested for heterologous complementation of EarA activity in a ΔearA mutant of M. maripaludis. We found that the EarA homologues from the three methanogens, but not EarA from M. jannaschii, could functionally complement EarA in a ΔearA deletion strain, resulting in the complemented cells producing FlaB and assembling archaella. At least one copy of the 6 bp consensus sequence to which EarA binds was found upstream of the putative BRE and TATA box of the fla operon in the three methanogens, as well as in P. furiosus. These data suggest EarA to be a widely distributed transcription regulator of the fla operon within the Euryarchaeota that might share the same activation mechanism, especially within methanogens.

**METHODS**

**Strains and culture conditions**

*M. maripaludis* S2 Δhpt (Mm900) and the *M. maripaludis* ΔearA (ΔearA, hereafter) mutant [27] derived from it were routinely cultured in Balch medium III under a head-space of H2:CO2 (80:20) with shaking at 35 °C [29, 30]. *M. maripaludis* strains carrying complementation vectors were cultured in a nitrogen-free medium supplemented with either 10 mM l-alanine or 10 mM NH4Cl as sole nitrogen source in the presence of 2.5 µg ml−1 puromycin for plasmid selection [31]. *M. vannielii* SB was also cultured anaerobically in Balch medium III at 35 °C while *M. thermolithotrophicus* DSM 2095 was grown in the same medium at 60 °C. *M. jannaschii* JAL-1 was grown at 80 °C in the minimal medium described by [32]. E. coli TOP10 cells, used for cloning, were cultured in Luria–Bertani (LB) broth or on LB agar plates in the presence of 100 µg ml−1 ampicillin for plasmid selection. Strains and plasmids used in this study are listed in Table 1.

**Phylogenetic analysis of EarA in the Archaea**

The protein sequence of EarA from *M. maripaludis* S2 (NCBI version CAF31274.1) was used as a query for synteny analysis among 237 archaeal genomes using the Prokaryotic Synteny and Taxonomy Explorer SyntTax [33]. Phylogenetic tree construction was performed using the Phylogeny.fr program (www.phylogeny.fr/) [34].
Table 1. Strains and plasmids used in this study

<table>
<thead>
<tr>
<th>Strain or plasmid</th>
<th>Description</th>
<th>Reference/source</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>M. maripaludis strains</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mm900</td>
<td><em>M. maripaludis</em> S2 Δhpt, wild-type strain in this study</td>
<td>[29]</td>
</tr>
<tr>
<td>ΔearAmma</td>
<td>Mm900 ΔearAmma</td>
<td>[27]</td>
</tr>
<tr>
<td><strong>Other methanogen strains</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mcc. jannaschii IAL-1</td>
<td>Wild-type</td>
<td></td>
</tr>
<tr>
<td>Mtc. thermolithotrophicus DSM 2095</td>
<td>Wild-type</td>
<td></td>
</tr>
<tr>
<td>M. vannielii SB</td>
<td>Wild-type</td>
<td></td>
</tr>
<tr>
<td><strong>E. coli strain</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TOP10</td>
<td>F' nifA1 Δ(nrd-hsdRMS-mcrBC) p80lacZΔM15 ΔlacX74 nupG recA1 araD139 Δ(ara-leu)7697 galE15 galK16 rpsL(StrR) endA1 λ−</td>
<td>Invitrogen</td>
</tr>
<tr>
<td><strong>Plasmids</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pHW40</td>
<td>nif promoter-lacZ fusion plus Pur' cassette; Amp'</td>
<td>[31]</td>
</tr>
<tr>
<td>pKJ1217</td>
<td>pHW40 harbouring earAmma, with a C-terminal hexahistidine tag under the control of a nif promoter</td>
<td>[27]</td>
</tr>
<tr>
<td>pKJ1280</td>
<td>pHW40 harbouring <em>M. vannielii</em> earA under the control of a nif promoter</td>
<td>This study</td>
</tr>
<tr>
<td>pKJ1281</td>
<td>pHW40 harbouring <em>M. jannaschii</em> earA under the control of a nif promoter</td>
<td>This study</td>
</tr>
<tr>
<td>pKJ1283</td>
<td>pHW40 harbouring <em>P. furiosus</em> earA under the control of a nif promoter</td>
<td>This study</td>
</tr>
<tr>
<td>pKJ1284</td>
<td>pHW40 harbouring <em>M. thermolithotrophicus</em> earA under the control of a nif promoter</td>
<td>This study</td>
</tr>
</tbody>
</table>

Construction of plasmids carrying genes encoding EarA homologues from selected archaea

Based on the SyntTax results, EarA homologues from *Mcc. jannaschii*, *M. vannielii* and *P. furiosus*, designated as EarA<sub>Mj</sub> (Mj_0905, NCBI version AAH89151.1), EarA<sub>Mv</sub> (Mvan_1020, NCBI version ABR54923.1) and EarA<sub>Pf</sub> (Pf_0340, NCBI version WP_011011455.1), respectively, were selected for complementation studies for EarAmma activity in the ΔearAmma strain. In addition, an EarA homologue from *Mtc. thermolithotrophicus*, designated as EarA<sub>Mt</sub> (NCBI version WP_026182974.1), was found via a BLASTP search using EarAmma as a query against a whole-genome shotgun sequence of *Mtc. thermolithotrophicus* DSM 2095. This allowed us to compare complementation of the *M. maripaludis* ΔearA mutant with earA homologues from a mesophilic (*M. vannielii*), thermophilic (*Mtc. thermolithotrophicus*) and hyperthermophilic (*Mcc. jannaschii*) member of the Methanococcales.

Genes encoding EarA<sub>Mj</sub> and EarA<sub>Mv</sub> were amplified by PCR with the primers listed in Table 2 and using washed whole cells of the corresponding strain as template. Several attempts to amplify earA<sub>Mt</sub> using washed whole cells of *Mtc. thermolithotrophicus* as template, however, were not successful. Thus, genomic DNA of *Mtc. thermolithotrophicus* was isolated using a DNeasy Blood and Tissue Kit (Qiagen) and used as template in the PCR to obtain earA<sub>Mt</sub>. Genomic DNA of *P. furiosus*, kindly provided by Dr Frank Robb (Institute of Marine and Environmental Technology, Baltimore MD), was used as a template with the corresponding primer pair listed in Table 2 to obtain earA<sub>Pf</sub>.

These four PCR products were digested with NsiI and MluI and cloned into an NsiI/MluI-digested *M. maripaludis* shuttle vector pHW40 to generate the recombinant plasmids listed in Table 1 that were used in the complementation studies [31]. The cloning into pHW40 results in the transcription of the cloned gene being dependent on an inducible nif promoter.

Complementation of the ΔearAmma mutant with earA homologues

Plasmids pKJ1217, pKJ1280, pKJ1281, pKJ1283 and pKJ1284 were transformed into the ΔearAmma deletion mutant [27] using a PEG-based method as described previously [35]. Transformants were subsequently cultured in a nitrogen-free medium supplemented with either L-alanine (nif promoter on) or NH₄Cl (nif promoter off) as the sole nitrogen source in the presence of puromycin.

Western blotting analysis of the ΔearAmma deletion mutant complemented with earA homologues from selected Archaea

The presence of archaellin FlaB2 in the ΔearAmma deletion mutant complemented with earA homologues was determined by Western blotting with anti-FlaB2 antibody, as described previously [28].

Electron microscopy

Cells of the ΔearAmma deletion mutant complemented with earA homologues from selected archaea were briefly washed with 2 % NaCl and resuspended in PBS. Cells were loaded on carbon-Formvar-coated copper grids and stained with 2 % phosphotungstic acid, pH 7. Grids were examined under a Phillips CM-10 transmission electron microscope operating at an accelerating voltage of 80 kV.
# RESULTS

## Phylogenetic distribution of EarA homologues across the Archaea

The first euryarchaeal transcriptional activator of the *fla* operon, EarA<sub>MMa</sub>, was recently identified in *Mc. maripaludis* [27]. Since the presence of archaella is widespread in various phyla of the Archaea [3, 5, 36] and the transcription or translation of the archaellum operon is known to be regulated in several model archaea [5], it was of interest to determine the phylogenetic distribution of *earA* as a potential universal transcriptional regulator of the *fla* operon in the Archaea.

Using EarA<sub>MMa</sub> as a query in protein BLAST [37], homologues of this protein were found strictly in members of the Euryarchaeota, although not in any member of the Halobacteria, which do contain archaellated species including *Halobacterium salinarum, Haloarcula marismortui* and *Halofex volcanii*. No significant similarity was found within members of other major phyla of the Archaea: Crenarchaeota, Thaumarchaeota, Korarchaeota or Nanoarchaeota. The genomic context of EarA homologues was then analysed with the synteny analysis tool SyntTax [33] (http://archaea.u-psud.fr/synttax/) against 237 archaeal genomes. Based on the genomic neighbourhood of *earA* and its proximity to the *fla* operon and chemotaxis operons, we grouped species containing *earA* into five classes, as shown in Fig. 1. Class I is represented by *Mc. maripaludis*, the organism in which *earA* was originally identified (Fig. 1a). In the genome of *Mc. maripaludis*, earA<sub>MMa</sub> is the second of six co-transcribed genes [27]. The operon begins with *mmp1719* encoding a conserved hypothetical protein, followed by earA<sub>MMa</sub> *mmp1717* encoding a methyltransferase, *mmp1716* encoding a methylenetetrahydromethanopterin dehydrogenase, *mmp1715* encoding a putative GTPase and *mmp1714* encoding a predicted regulator of Ras-like GTPase activity. In addition to *Mc. maripaludis*, this six-gene operon was found in several methanogens in the family Methanococcaceae, including *M. vanielli* and *Mtic. thermolithotrophicus*. In *Methanococcus voltae*, which is a close relative of *Mc. maripaludis*, only genes homologous to *mmp1719*, *earA* and *mmp1717* were likely to be co-transcribed. Homologues of *mmp1714* and *mmp1715* were not found in the *Mc. voltae* genome. *mmp1716*, encoding a second copy of the methylenetetrahydromethanopterin dehydrogenase in *Mc. maripaludis*, is homologous to the only copy of the methylenetetrahydromethanopterin dehydrogenase gene in *Mc. voltae*. The location of *earA* is not in the immediate neighbourhood of *fla* or *che* genes in the class I organisms. The *fla* operon genes in *Mc. maripaludis* (*flaB1 to flaJ; mmp1666-1676*), for example, are over 35 kb upstream of *earA*.

In contrast to the situation found in the class I organisms, in the genomes of archaellar species belonging to classes II–V, only the *earA* and the methyltransferase gene (*mmp1717*) homologues were typically located adjacent to each other and likely to be co-transcribed (Fig. 1). Unlike in *Mc. maripaludis*, where earA<sub>MMa</sub> is located at a distance from the *fla* operon, in classes II–V, *earA* and the methyltransferase gene are in the same immediate genomic neighbourhood as the *fla* operon and, sometimes, chemotaxis operons (Figs. 1b–e). In rare cases in classes II–V, such as *Methanocaldococcus bathoaredescens* and *Mcc. jannaschii* DSM 2661, a homologue of *mmp1719* was located adjacent or near the *earA* and the methyltransferase homologues (Fig. 1b).

In class II organisms, *earA* and the methyltransferase gene are located downstream of the *fla* operon. This class includes *Mcc. jannaschii*, *M. bathoaredescens*, *Geoglobus ahangari* and *Ferroglobus placidus*. The first two species are hyperthermophilic methanogens and the last two belong to the family Archaeoglobaceae. In the genomes of the two Archaeoglobaceae, *earA* and the methyltransferase gene are located immediately downstream of the last gene of the *fla* operon, *flaJ*. In the genomes of the two *Methanocaldococcus* spp., the pre-archaellin peptidase gene *flaK* is located next to the *fla* operon, followed by earA<sub>MMa</sub> and the methyltransferase gene homologue, gapped by one or two genes, including an *mmp1719* homologue.

Archaea classified in classes III–V are members of the Thermococcales. In class III, *earA* and the methyltransferase gene are located immediately upstream of the *fla* operon. In classes IV and V, *earA* and the methyltransferase gene are located between the *fla* operon and chemotaxis genes. In class IV organisms, *earA* and the methyltransferase gene are located immediately upstream of the *fla* operon and transcribed in the same orientation, while in class V they are

### Table 2. Primers used in this study

<table>
<thead>
<tr>
<th>Primer</th>
<th>Sequence</th>
<th>Restriction site incorporated (underlined)</th>
</tr>
</thead>
<tbody>
<tr>
<td>earA&lt;sub&gt;Mca&lt;/sub&gt;-F</td>
<td>GACATGCATGGGATCAGAAGACATGATATTT</td>
<td>NsiI</td>
</tr>
<tr>
<td>earA&lt;sub&gt;Mca&lt;/sub&gt;-R</td>
<td>ACTACGGGTTCACCAAATTATAGTGCGGAAATC</td>
<td>MluI</td>
</tr>
<tr>
<td>earA&lt;sub&gt;Mmc&lt;/sub&gt;-F</td>
<td>GACATGCATGGGATCAGAAGACATGATATTT</td>
<td>NsiI</td>
</tr>
<tr>
<td>earA&lt;sub&gt;Mmc&lt;/sub&gt;-R</td>
<td>ACTACGGGTTCACCAAATTATAGTGCGGAAATC</td>
<td>MluI</td>
</tr>
<tr>
<td>earA&lt;sub&gt;Msp&lt;/sub&gt;-F</td>
<td>ACTACGGGTTCACCAAATTATAGTGCGGAAATC</td>
<td>NsiI</td>
</tr>
<tr>
<td>earA&lt;sub&gt;Msp&lt;/sub&gt;-R</td>
<td>ACTACGGGTTCACCAAATTATAGTGCGGAAATC</td>
<td>NsiI</td>
</tr>
<tr>
<td>earA&lt;sub&gt;Mva&lt;/sub&gt;-F</td>
<td>GACATGCATGGGATCAGAAGACATGATATTT</td>
<td>NsiI</td>
</tr>
<tr>
<td>earA&lt;sub&gt;Mva&lt;/sub&gt;-R</td>
<td>ACTACGGGTTCACCAAATTATAGTGCGGAAATC</td>
<td>MluI</td>
</tr>
</tbody>
</table>

---

**References:**

Fig. 1. Grouping of Euryarchaeota into five classes according to the genomic neighbourhood of earA. In all five classes, earA (shown in red) is located immediately upstream of a methyltransferase gene (shown in yellow). Class I: earA is located in a six-gene operon at a
distance from the fla operon (three downstream genes are missing in Mc. voltae). Also included in this class are Methanocaldococcus infernus and Methanothermobacter okinawensis. Class II and III: earA is located downstream (class II) or upstream (class III) of the fla operon. Other members of class II include Methanocaldococcus vulcanus, Methanocaldococcus fervens and Methanocaldococcus villosus. Also included in class III are Thermococcus barophilus, Thermococcus thioreducens, Thermococcus sibiricus, Thermococcus peptonophilus, Palaeococcus ferrophilus, Palaeococcus pacificus and Pyrococcus yannonisii. Class IV and V: earA is located between the fla operon and chemotaxis genes which are found typically in the opposite orientation. In class IV, earA is transcribed in the same orientation and found upstream of the fla operon. Additional members of class IV include Thermococcus piezophilus, Thermococcus celerilescens, Thermococcus eurythermalis and Thermococcus nautii. In class V, earA is transcribed in the same orientation and found upstream of chemotaxis genes. Other members of class V include Thermococcus zillii. Homologous genes are shown in the same colour. The fla operon is shown in blue and chemotaxis genes are shown in green.

located immediately upstream of che operons and transcribed in the same orientation.

**EarA**<sub>Mma</sub> binding site is found upstream of the fla operons in selected methanogens

Since EarA proteins are all ~110 aa in length with an annotated helix-turn-helix DNA binding domain, it is very likely that they all work as transcriptional regulators in their native host. To further explore whether these EarA homologues might serve the same function as EarA<sub>Mma</sub> in *Mc. maripaludis*, i.e. regulate transcription of the fla operon via binding to DNA motifs upstream of the fla promoter, the fla promoter region from selected archaeellated methanogens was first analysed for potential EarA binding sites. DNA sequences in the fla promoter region from *Mc. maripaludis* (GenBank version NC_000909.1), *Mcc. jannaschii* (GenBank version NC_006624.1), *Mvo* (<i>Mrc. voltae</i>) (GenBank version NC_000909.1), *Mcc. maripaludis* (GenBank version NC_006624.1) were also chosen for analysis as archaeellated non-methanogen representatives in which transcriptional work on the fla operon has been published [38, 39]. As shown in Fig. 2, the putative TATA boxes and BREs are highly conserved in the seven selected Archaea, although the distance between the TATA boxes and the start codon of the first gene in the fla operon varies.

In *Mc. maripaludis*, four 6 bp consensus sequences (one in the inverse orientation) located −136 bp to −74 bp upstream of the transcription start site (TSS) of the fla operon were identified as binding sites of EarA<sub>Mma</sub> [27]. At least one copy of this 6 bp consensus sequence is also found upstream of the TATA box in each selected methanogen: one in *Mcc. jannaschii* and *Mtc. thermolithotrophicus*, four in *Mvo*, and three in *Mcc. vannielii* (Fig. 2). The consensus sequences are all upstream of the putative BREs and TATA boxes, with the first consensus sequence found 22–36 nucleotides upstream of the BRE in all the methanogens (Fig. 2). We also examined the sequence upstream of the fla operon in T. kodakarense. Transcriptional analysis in this organism has shown that the operon begins with the first archaeellin and the methyltransferase gene located immediately upstream is transcribed separately [38]. In the intergenic region between these two genes, we identified a potential BRE and TATA box and one copy of the 6 bp consensus sequence located 32 nt upstream of BRE, the same distance upstream as in the methanogens. In *P. furiosus*, transcriptional analysis suggested that the earA homologue is co-transcribed with the methyltransferase and at least some of the fla operon genes, although a single transcript for the first fla gene, the major archaellin *flaB0*, was also reported [39]. Immediately upstream of a potential BRE and TATA box located in the intergenic region between the methyltransferase and *flaB0*, a single copy of the 6 bp consensus sequence was again found, 32 nt upstream of BRE.

**EarA homologues examined for complementation of an ΔearA<sub>Mma</sub> mutant**

The presence of the EarA<sub>Mma</sub> binding consensus sequence upstream of the fla operons in four other methanogens indicated the possibility that the EarA homologues of these methanogens might recognize and bind to the same DNA motif and thus activate transcription of the fla operon via the same mechanism. This hypothesis is further supported by protein alignments which also showed high sequence identity/similarity between EarA<sub>Mma</sub> and EarA<sub>Mva</sub>, EarA<sub>Jja</sub>, EarA<sub>Mbi</sub> and EarA<sub>Mvo</sub> (71%–93% identity, 83%–97% similarity). BLAST results are summarized in Table 3 while alignments of EarA<sub>Mma</sub> with each of the other EarA homologues are shown in Figs S1–S4, available in the online Supplementary Material). Among these methanogen EarA homologues, the highest similarity to EarA<sub>Mma</sub> was found with EarA<sub>Mva</sub> and EarA<sub>Jja</sub>, and the lowest to EarA<sub>Mbi</sub> (see also phylogenetic tree Fig. S7).

Several EarA homologues, i.e. EarA<sub>Mvo</sub>, EarA<sub>Jja</sub> and EarA<sub>Mbi</sub>, were examined for their ability to functionally complement in a ΔearA<sub>Mma</sub> deletion mutant. The various EarAs represent members of the Methanococcales that are mesophilic (*Mcc. vannielii*), thermophilic (*M. voltae*) and hyperthermophilic (*Mcc. jannaschii*). In addition, the EarA homologue from *P. furiosus* was also examined as a representative outside of the Methanococcales. earA<sub>Pfu</sub> is located immediately upstream of the fla operon with the methyltransferase gene, placing it in EarA class I (Fig. 1d). EarA<sub>Mva</sub> and EarA<sub>Pfu</sub> share 57% amino acid identity (76% similarity, Fig. S5). It should be noted that in *P. furiosus* the initial sequencing of the fla operon region was erroneous and had 771 bp missing. This missing sequence was later identified, leading to the
discovery of a third and the major archaellin (FlaB0) [39]. The revised order of the genes in *P. furiosus* is earA homologue, methyltransferase, flaB0, flaB1, flaB2, flaC, flaD, flaF, flaG, flaH, flaI and flaJ (Fig. 1c).

Plasmids carrying the various heterologous *earA* homologues were transformed into an ΔearA<sub>Mma</sub> deletion mutant. The various transformants were cultured in nitrogen-free medium supplemented with either L-alanine or NH₄Cl as

Fig. 2. Presence of EarA<sub>Mma</sub> consensus binding sites upstream of putative BREs and TATA boxes of *fla* operons in selected euryarchaea. The putative BRE and TATA boxes (shown in green) of the *fla* operons from *Mc. maripaludis* (*Mma*), *Mth. thermolithotrophicus* (*Mth*), *Mcc. jannaschii* (*Mja*), *Mc. vannielii* (*Mva*), *Mc. voltae* (*Mvo*), *T. kodakarensis* (*Tko*) and *P. furiosus* (*Pfu*) are aligned manually. The first gene in each *fla* operon is shown in red. Gene coding area upstream of the *fla* promoters is shown in grey. The 6 bp consensus sequences to which EarA<sub>Mma</sub> binds in *Mc. maripaludis* are highlighted in yellow.
the sole nitrogen source. After a minimum of three successive transfers, these cells were examined for evidence of heterologous complementation of EarA activity. Initially, FlaB2 expression from complemented cells harbouring vectors expressing EarA<sub>Mva</sub>, EarA<sub>Mtb</sub>, EarA<sub>Mja</sub>, EarA<sub>Pfu</sub>, as well as EarA<sub>Mma</sub> as a positive control, was analysed by Western blotting. Under alanine growth conditions (when transcription of earA from the nif promoter is induced), all three methanogen EarA homologues, as well as EarA<sub>Mma</sub> itself, were able to complement the function of EarA<sub>Mma</sub> as FlaB2 expression was restored (Fig. 3). In contrast, FlaB2 was not detected in the ΔearA<sub>Mma</sub> deletion mutant complemented with earA<sub>Pfu</sub> under the same growth conditions. All complemented cells were also cultured in nitrogen-free medium supplemented with NH₄Cl (where transcription from the nif promoter is repressed) and examined for the presence of FlaB2 by Western blot analysis, and in all cases FlaB2 production was not detected (Fig. 3).

<table>
<thead>
<tr>
<th>Protein</th>
<th>Amino acid identity (compared to EarA&lt;sub&gt;Mma&lt;/sub&gt;)</th>
<th>Amino acid similarity (compared to EarA&lt;sub&gt;Mma&lt;/sub&gt;)</th>
<th>EarA class</th>
</tr>
</thead>
<tbody>
<tr>
<td>EarA&lt;sub&gt;Mma&lt;/sub&gt;</td>
<td>100 %</td>
<td>100 %</td>
<td>I</td>
</tr>
<tr>
<td>EarA&lt;sub&gt;Mva&lt;/sub&gt;</td>
<td>93 %</td>
<td>95 %</td>
<td>I</td>
</tr>
<tr>
<td>EarA&lt;sub&gt;Mtb&lt;/sub&gt;</td>
<td>92 %</td>
<td>97 %</td>
<td>I</td>
</tr>
<tr>
<td>EarA&lt;sub&gt;Mja&lt;/sub&gt;</td>
<td>81 %</td>
<td>89 %</td>
<td>I</td>
</tr>
<tr>
<td>EarA&lt;sub&gt;Pfu&lt;/sub&gt;</td>
<td>71 %</td>
<td>83 %</td>
<td>II</td>
</tr>
<tr>
<td>EarA&lt;sub&gt;Mga&lt;/sub&gt;</td>
<td>57 %</td>
<td>76 %</td>
<td>III</td>
</tr>
</tbody>
</table>

**DISCUSSION**

Swimming is an important attribute for living organisms and it can contribute to a cell’s ability to cope with changing environments. Archaea use archaella as their swimming organelles; thus the activation of archaella synthesis might be critical for the survival of Archaea species under certain environmental conditions. The identification of EarA from *Mc. maripaludis* as the first transcriptional activator of the fla operon in the Euryarchaeota led us to investigate the phylogenetic distribution of this regulator, as an initial clue to the regulation of archaella synthesis in Euryarchaeota at the molecular level. Analysis of 237 archaeeal genomes using the Prokaryotic Synteny and Taxonomy Explorer SyntTax [33] showed that this protein is widespread in Euryarchaeota although unexpectedly absent in extreme halophiles, suggesting that EarA is likely to regulate archaella synthesis commonly, but not universally, in this phylum.

The absence of *earA* homologues in members of the Haloarchaea was surprising to us. Many extreme halophiles are archaellated, including ones like *Hbt. salinarum*, *Har. marismortui* and *Hfx. volcanii* that have been well studied in regard to archaellation and chemotaxis [40–50]. Examination of the genome region surrounding the fla genes in *Hfx. volcanii* did reveal the presence of a gene annotated as a transcriptional regulator. Although the encoded protein (WP_004043733.1) contained a helix-turn-helix DNA binding domain and was 112 aa in length, similar to the 110 aa length of EarA, it bears no significant sequence similarity to EarA (Fig. 5S), and was not retrieved in *blast* searches using EarA as query. The gene is located between *fla* genes and
che genes and is likely to be the transcriptional regulator for the fla operon in this organism. This regulator has homologues in many other extreme halophiles and, at least in the several cases we examined (including Hbt. salinarum R1, Halobacterium NRC-1 and Halorhabdus utahensis), the homologue also lies between fla genes and che genes. Interestingly, in the intergenic region between the end of the regulator gene and the start of the archaellin gene in Hfx. volcanii, there are no TACATA sequences to which EarA binds but we did identify one inverted repeat of this sequence (TATGTA). Recently, Legerme et al. [51] screened a Hfx. volcanii transposon library for motility mutants and, among the numerous motility-affected isolates, was one with a transposon insertion in gene hvo_0246 which encodes a predicted ArsR family transcription regulator, much larger than EarA at 320 aa in length (Fig. S6). This gene is located at a distance from the fla operon (beginning with archaellin hvo_1209), and the gene(s) it may regulate are still unknown. Mutations in many genes outside of the fla operon, such as certain ones involved in N-linked glycosylation, could result in motility defects [52, 53].

In Mc. maripaludis, earA_Mna is located upstream of a methyltransferase gene within a six-gene operon that is located approximately 35 kb from the fla operon [27]. Synteny analysis indicates that this arrangement of genes and the location of earA at a distance from the fla operon as found in class I archaea is not the norm. Instead, in most of the species in which an EarA homologue is found, earA is found in the immediate vicinity of the fla operon and often chemotaxis genes as well. The close association of earA and the methyltransferase gene suggests that the methyltransferase gene is also involved in archaellation synthesis. In P. furiosus, co-transcription of the earA homologue, the methyltransferase gene and at least some of the fla operon genes was demonstrated, although a single transcript for the first fla gene (flaB0) was also found [39]. Thus, in P. furiosus, the fla operon includes both the earA and methyltransferase homologues. These findings led to speculation that the methyltransferase might be modifying the archaellins [39]. Methylation of lysine residues in both flagellins and pilins of bacteria has been reported [54, 55]. We believe it is also possible that the methyltransferase interacts directly with EarA to modify its DNA-binding ability. We tried to investigate the potential role of MMP1717 in archaea regulation or archaellin modification by deleting mmp1717, but all attempts to do so were unsuccessful, suggesting that mmp1717, unlike earA or the fla operon genes, is an essential gene. If true, this means that if the methyltransferase does play a role in archaellation, it must have another, essential role or roles in the cell.

In EarA classes II–V, the earA genes, together with the associated mmp1717 homologues, are located adjacent to the fla operons. This close physical association of a regulator gene with the fla operon is also seen in S. acidocaldarius, with arrR and arrR1, encoding transcriptional activators of the fla operon, being located immediately adjacent the fla operon [23]. Furthermore, in EarA classes IV and V, chemotaxis operons, in addition to the fla operon, are located immediately adjacent to earA and the methyltransferase gene. Although archaella are genetically and structurally distinct from bacterial flagella, the chemotaxis system is highly conserved between the two domains and it is thought that Archaea che genes were introduced into the Euryarchaeota.
ancestor by lateral gene transfer from bacteria [56]. In bacteria, it is not uncommon for some late flagella genes, like those encoding flagellins, to be co-regulated with chemotaxis genes – often through the use of alternative fla-specific σ factors to direct transcription of these genes [57]. Archaea do not have σ factors [58, 59], but it is possible that euryarchaeota may co-regulate the expression of unlinked archaellum and chemotaxis genes through the use of a common transcriptional activator, like earA.

The DNA sequence upstream of the fla operon in several methanogens was examined for the presence of the 6bp consensus sequence to which EarA<sub>Mma</sub> binds. When the putative BREs and TATA boxes for Mc. vannielii, Mtc. thermodithrophilicus, Mc. voltae and Mcc. jannaschii fla operons were aligned with those from Mc. maripaludis, at least one copy of the 6bp consensus sequence was found immediately upstream, suggesting that EarA homologues from these methanogens likely use the same consensus sequence as binding sites for the activation of the transcription of the fla operon in their own host. Although the number of 6bp consensus sequences varies in the different methanogens from one to as many as four, the consensus sequence closest to the fla promoter in the four methanogens is located at a very similar distance upstream of the BREs and TATA boxes. Although earlier EMSA analysis confirmed the binding of purified EarA<sub>Mma</sub> to any of the four consensus sequences in Mc. maripaludis, in vivo mutation of the two consensus sites located furthest upstream of the promoter in the genome of Mc. maripaludis did not interfere with the transcription of the fla operon [27]. In contrast, mutation of all four sites abolished the transcription activation of the fla operon, indicating that at least one but not all four of the consensus sequences is required for the transcription activation of the fla operon in Mc. maripaludis by EarA [27]. The alignment of the most promoter-proximal consensus sequence in the selected four methanogens, and even in P. furiosus and T. kodakarense, suggests that this one plays a critical role in the EarA-dependent in vivo activation of the fla operon.

While the EarA homologues from the three Methanococcales were able to complement the ΔearA<sub>Mma</sub> deletion strain, the EarA homologue from P. furiosus, however, could not complement the function of EarA<sub>Mma</sub>. EarA<sub>Pfu</sub> shares 57% amino acid identity with EarA<sub>Mma</sub>, which is much lower than the 71% identity between EarA<sub>Mma</sub> and EarA<sub>Mmp</sub>, which is the lowest identity shared between EarA<sub>Mma</sub> and the three methanogen EarAs used in this study. This may explain the failure of the complementation of EarA<sub>Mma</sub> with EarA<sub>Pfu</sub>. Furthermore, since we used untagged versions of the EarA homologues in our study, we were unable to test whether the P. furiosus EarA was actually synthesized by Mc. maripaludis. There may be codon bias properties of the P. furiosus earA that preclude its optimal production in Mc. maripaludis. Analysis of the codon usage in the earA homologue of P. furiosus does reveal several codons not used in earA of Mc. maripaludis; not surprising perhaps considering the small size of the proteins. These codons are one GCG (alanine), one CAT (histidine), three ATC (isoleucine), one CTC (leucine), one CCC (proline), one TCC (serine) and two GTG (valine). Of these, GCG, CTC, CCC, TCC and GTG are used between 6–9% as the codon for that particular amino acid in Mc. maripaludis ([www.kazusa.or.jp/codon/], while CAT (42%) and ATC (16.5%) are commonly used codons. In the highly expressed Mc. maripaludis archaellins FlaB1 and FlaB2, the codons GCG, CCC and GTG are never used so their presence in the earA of P. furiosus may be a problem for expression. The sequence of a key area of the fla operon was missing in the initial P. furiosus genome sequence [60]. A 771 nt sequence which contained the first archaellin of the fla operon as well as the correct start of the next archaellin, flaB1, was later reported [39]. This sequence also contained some of the intergenic region between the end of the methyltransferase gene and the start of flaB0 missing in the original genome sequence. Examination of the complete intergenic region between the methyltransferase gene and flaB0 revealed potential BRE and TATA boxes, as well as a single TACATA sequence located at the same upstream position as found in the Methanococcales and Thermococcus sequences. Thus, while the failure of EarA<sub>Pfu</sub> to complement the Mc. maripaludis ΔearA deletion mutant might be explained by EarA<sub>Pfu</sub> using another consensus sequence upstream of the fla operon as its binding site compared to EarA<sub>Mma</sub>, this seems unlikely given the presence and location of a known binding sequence for EarA in the P. furiosus fla promoter region. P. furiosus grows in a temperature range between 70 and 103 °C, with the optimum growth temperature at 100 °C [61]. Considering that all of the complementation studies, including that employing EarA<sub>Pfu</sub>, were conducted at 35 °C, it is possible that EarA<sub>Pfu</sub> is not active simply due to the relatively low temperature at which the complementation assays were conducted compared to the normal habitat temperature of P. furiosus. EarA from the hyperthermophile Mcc. jannaschii, on the other hand, was able to complement at the much reduced temperature, but the temperature range for Mcc. jannaschii is 50–86 °C with an optimum of 80–85 °C [62].

Phylogenetic analysis of the first euryarchaeal archaellum regulator EarA<sub>Mma</sub> indicates the widespread distribution of this regulator in the phylum of Euryarchaeota, with the notable exclusion of the Halobacteriales. Species containing the earA gene are grouped here into five classes according to the genomic location of earA genes. In four out of the five classes (but not in Mc. maripaludis where EarA was first identified), the earA gene is located adjacent to the fla operon, and in two classes the chemotaxis operons are also found nearby. The function of three methanogen-derived EarA homologues as a transcriptional activator of the fla operon was confirmed via heterologous complementation studies. This work implies that in archaellated methanogens, and by extension in other euryarchaeotes harbouring an earA homologue, transcriptional activation of the fla operon...
is likely to occur via a mechanism similar to the EarA_Mma system in _Mc. maripaludis._

**Funding information**
Research in this paper was supported by grants from the Natural Sciences and Engineering Research Council of Canada (NSERC) awarded to K. F. J. and C. M. K.

**Acknowledgement**
The authors are grateful to Dr Frank Robb for the gift of _Pyrococcus furiosus_ DNA.

**Conflicts of interest**
The authors declare that there are no conflicts of interest.

**References**
motility and adhesion mutants. Screening of a


Syutkin AS, Pyatibratov MG, Galitzskaya OV, Rodriguez-Valera F, Fedorov OV. Haloarcula marismortui archaein genes as ecopara-

Beznosov S, Pyatibratov M, Veluri P, Mitra S, Fedorov O. A way to identify archaeins in Haloarcula marismortui archaea by FLA-

Pyatibratov MG, Beznosov SN, Rachel R, Tiktopulo EI, Surin AK et al. Alternative flagellar filament types in the haloarchaeon Halo-


Tripepi M, Imam S, Pohlscröder M. Halofex volcanii flagella are required for motility but are not involved in PibD-dependent sur-


Tripepi M, You J, Temel S, Önder Ö, Brisson D et al. N-glycosyla-

Kim YH, Park KH, Kim SY, Ji ES, Kim JY et al. Identification of tri-
methylation at C-terminal lysine of pilin in the cyanobacterium synechocystis PCC 6803. Biochim Biophys Res Commun 2011;404:
587–592.

1539–1547.


Helmann JD. Alternative sigma factors and the regulation of flag-


Burton SP, Burton ZF. The σ enigma: bacterial σ factors, archaeal TFB and eukaryotic TFII-B are homologs. Transcription 2014;5:
61.


Fiala G, Stetter KO. Pyrococcus furiosus sp. nov. represents a novel genus of marine heterotrophic archaeabacteria growing opti-
mally at 100°C. Arch Microbiol 1986;145:56–61.

Jones WJ, Leigh JA, Mayer F, Woese CR, Wolfe RS. Methanococ-
cus jannaschii sp. nov., an extremely thermophilic methanogen from a submarine hydrothermal vent. Arch Microbiol 1983;136:
254–261.

Edited by: S-V. Albers and F. Sargent

Five reasons to publish your next article with a Microbiology Society Journal

1. The Microbiology Society is a not-for-profit organization.
2. We offer fast and rigorous peer review – average time to first decision is 4–6 weeks.
3. Our journals have a global readership with subscriptions held in research institutions around the world.
4. 80% of our authors rate our submission process as ‘excellent’ or ‘very good’.
5. Your article will be published on an interactive journal platform with advanced metrics.

Find out more and submit your article at microbiologyresearch.org.