i. Title:

Sodium-powered stators of the bacterial flagellar motor can generate torque in the presence of phenamil with mutations near the peptidoglycan-binding region.

ii. Running Title:
Mutants in PG-binding motif of MotB impart motility in presence of phenamil.

iii. Authors: The full names of the authors;

Tsubasa Ishida[#1], Rie Ito[#1], Jessica Clark[2], Nicholas J Matzke[3], Yoshiyuki Sowa[*1,4], Matthew AB Baker[*2]
[#] equal contribution
[*] joint corresponding authors

iv. Contact Information:

[1] Department of Frontier Bioscience, Hosei University, Tokyo, Japan.
[2] School of Biotechnology and Biomolecular Science, University of New South Wales, Kensington, NSW, Australia.
[3] School of Biological Sciences, University of Auckland, Auckland, New Zealand.
[4] Research Center for Micro-Nano Technology, Hosei University, Tokyo, Japan.

v. Keywords:

bacterial motility
flagellar motor
ion channels
phenamil
sodium channels

Summary:

The bacterial flagellar motor (BFM) powers the rotation that propels swimming bacteria. Rotational torque is generated by harnessing the flow of ions through ion channels known as stators which couple the energy from the ion gradient across the inner membrane to rotation of the rotor. Here we used error-prone PCR to introduce single point mutations into the sodium-powered *Vibrio alginolyticus / Escherichia coli* chimeric stator PotB and selected for motors that exhibited motility in the presence of the sodium-channel inhibitor phenamil. We found single mutations that enable motility under phenamil occurred at two sites: 1) the transmembrane domain of PotB, corresponding to the TM region of the PomB stator from *V. alginolyticus*, and 2) near the peptidoglycan (PG) binding region that corresponds to the C-terminal region of the MotB stator from *E. coli.* Single cell rotation assays confirmed that individual flagellar motors could rotate in up to 100 µM phenamil. Using phylogenetic logistic regression, we found correlation between natural residue variation and ion source at positions corresponding to PotB F22Y, but not at other sites. Our results demonstrate that it is not only the pore region of the stator that moderates motility in the presence of ion-channel blockers.

Introduction

Motility imparts a large benefit to organisms competing for sparse nutrients. The bacterial flagellum is the oldest known form of motility (Rossmann and Beeby 2018), consisting of a propeller-like filament that rotates under the power of the bacterial flagellar motor (BFM). The BFM is 40 nm in diameter, powered by ion transit across the cell membrane through the stator protein complex, a heterodimer which forms a selective ion channel that transduces chemical energy into mechanical torque (Minamino et al. 2018). While most stators are proton powered (Sowa and Berry 2008), some from marine habitats are sodium powered (Yorimitsu and Homma 2001), others are powered by both sodium ions and protons (Paulick et al. 2015) and recently some have been discovered that are even powered by large divalent cations (Imazawa et al. 2016). In general, flagellar motors have adapted to function in various environments where bacteria live and survive (Terashima et al. 2017; Rossmann and Beeby 2018; Chaban, Coleman, and Beeby 2018). BFM diversity creates an ideal case study to investigate how macromolecular complexes adapt to different environments.

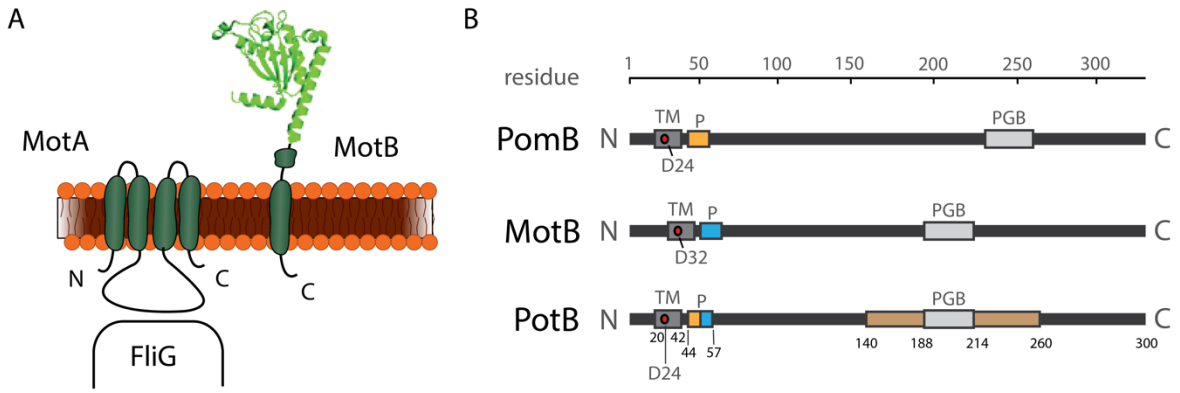
In *E. coli* the energy from the proton gradient is harnessed by the transit of protons through a heterodimeric stator complex of a homotetramer and a homodimer ($MotA_4MotB_2$) (Minamino et al. 2018). In *Vibrio* species this heterodimer is the sodium ion-powered $PomA_4PomB_2$. The total complex, in both cases, consists of four transmembrane (TM) domains of the A-subunit, a single TM domain of the B-subunit, and a large periplasmic region of the B-subunit which consists of a plug segment and a peptidoglycan-binding (PG-binding) region (Kojima and Blair 2001; Kojima et al. 2009). Recently, the structural rearrangements in PG-binding and stator activation were resolved, indicating that conformational rearrangements of the linker between the PG-binding region and the TM-domain are critical in stator activation (Kojima et al. 2018).

Whilst crystal structures have been solved for the C-terminal domains of MotB in multiple species (Roujeinikova 2008; Kojima et al. 2009; Zhu et al. 2014), the full structure of the TM-domain in any species has yet to be resolved beyond low-resolution cryoEM of the entire stator complex (Yonekura, Maki-Yonekura, and Homma 2011). Much of our understanding of stator function in the TM-domain relies on mutagenesis (Sharp, Zhou, and Blair 1995; Toshiharu Yakushi et al. 2006). Chimeric B-subunits have been engineered that combine the C-terminal peptidoglycan binding motif of the MotB subunit with the transmembrane domain of PomB that forms the ion-channel (in complex with two subunits of PomA) (Asai et al. 2003) (Fig. 1). These chimeric $PomA_4PotB_2$ stators utilise sodium motive force to drive flagellar rotation and can bind to the peptidoglycan (PG) layer thus enabling *E. coli* to swim under sodium-motive force. They have enabled the function of the motor to be investigated at low sodium concentration, and thus low energisation (Sowa et al. 2005; Lo et al. 2013).

Since this chimeric stator complex is driven by sodium ion flow across the inner membrane, it is natural to examine how sodium channel blockers affect this process. Phenamil, a known sodium-channel blocker (Garvin et al. 1985), has been used to probe sodium-interaction sites in sodium driven motors *in situ* in *Vibrio* species (Sugiyama, Cragoe, and Imae 1988; T. Atsumi et al. 1990). Work in *V. alginolyticus* demonstrated that motility under phenamil could be restored by mutations at the cytoplasmic end of the TM-domain of the stator complex (Kojima et al. 1997, 1999). *V. parahaemolyticus* has been widely studied as a model for dual-powered motility; it has lateral flagella that are proton powered but polar flagella that are sodium

94 powered (Tatsuo Atsumi, McCarter, and Imae 1992). Directed evolution approaches have
95 been applied to *V. parahaemolyticus* to examine which spontaneous mutations might induce
96 resistance to phenamil (Jaques, Kim, and McCarter 1999). Jaques et al. selected for resistance
97 to phenamil by selecting flagellar motors that were motile in the presence of 40 µM phenamil
98 methanesulfonate, yet interestingly only observed a single mutant in the TM-domain that
99 resulted in phenamil resistance.

100 Here we engineered a plasmid construct with enzyme cut sites adjacent to PotB to examine
101 the effects of randomly generated mutations on motility. We screened using phenamil to
102 measure the frequency and location of mutations that enabled motiity under phenamil in the
103 sodium powered $PomA_4PotB_2$ chimera. We induced mutations at a controllable rate using
104 error-prone PCR (error-prone PCR) and screened large populations of cells using streaking of
105 transformed error-prone PCR product onto swim agar in streaks for screening and subsequent
106 sequencing. This allowed a high-throughput screen to determine which mutations have
107 resulted in strains that were functional in the presence of phenamil. We examined how
108 frequently these mutations arose naturally using phylogenetics and parsimony reconstruction
109 to determine whether these adaptations were correlated with existing PomB/MotB
110 classification. By examining the structural and sequence location of these mutations, we have
111 assembled a model that outlines the relationship between PG-binding and stator activity, and
112 presented a framework for determining which residues hold evolutionary importance and
113 may constrain the possible adaptive pathways available to the stators.



114
115 **Figure 1: Schematic and structure of the stators of the bacterial flagellar motor.** (A) The stator complex in *E.*
116 *coli* consists of a heterodimeric complex composed of $MotA_4MotB_2$. Each stator has two ion channels that are
117 composed of 4 TM domains of MotA and 1 TM domain of MotB. (B) The PotB chimera consists of residues 1-50
118 from V. alginolyticus PomA and residues 59-308 of *E. coli* MotB with TM, plug, PGB and ompA-like regions (brown
119 bar) indicated on the schematic respectively. Structure in A from (Seiji Kojima et al. 2008).
120

121 **Results**

122

123 ***Mutagenesis of PotB.***

124 We generated mutations across the entirety of the PotB protein using error-prone PCR as per
125 Methods. We defined the following regions across the PotB protein (numbering via PotB
126 residue number): the TM domain (20-42), the plug (44-57), the middle domain (57-140) the
127 ompA-like domain (140-260) and the peptidoglycan binding domain (188-214). We screened
128 the combinatoric pool of mutagenesis outputs using swim streaking (Fig. S1) and
129 characterised the motility of 27 error-prone PCR plasmids that exhibited some motility when

selected for motility in the presence of phenamil (Fig. S2; Table S1). The distribution of all nucleotide mutations in all 27 plasmids, including silent mutations, is shown in Fig. 2.

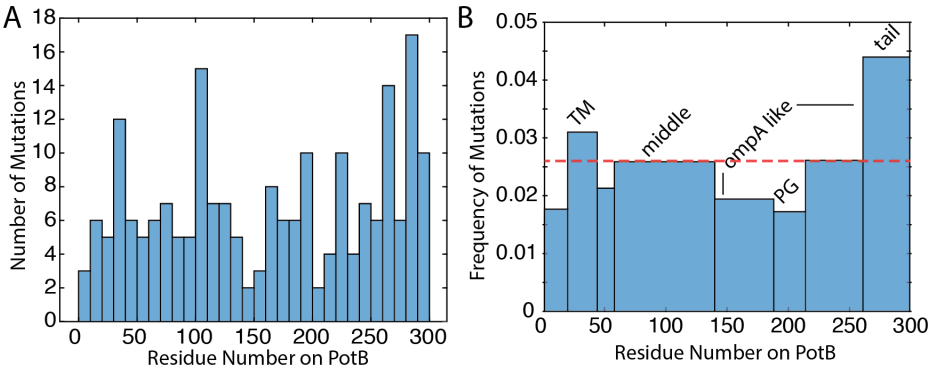


Figure 2: Histogram of mutant frequency. Histograms of mutations from error-prone PCR across every strain with at least partial motility in phenamil (from spread plate). (A) Binned evenly across PotB with each bin 10 residues in width showing raw counts of mutations across aggregated results from all 27 error-prone PCR plasmids. (B) Frequency of mutations binned according to specific regions of PotB: TM (20-42), plug (44-57), middle (58-139), ompA-like first section (140-187), PG region (188-214), ompA-like second section (215-260) and tail (261-300). Counts are normalised across each region and strain (average mutations per residue per strain), and the mean across the entire protein is indicated by the dashed red line (0.026 mutations per residue per strain)

Together, these histograms indicated that the mutations occurred roughly evenly throughout the protein. We quantified the mutation rate in each region of the protein by calculating the mean number of mutations per strain in each region and normalising for the length of each region to calculate a mutation rate per residue for each region (Fig. 2B). We observed that proportionately more mutations occurred in the TM and the tail region of PotB, that is, the mutation rate was 0.031 and 0.044 mutations/residue in the TM and tail region respectively, which was greater than the overall mutation rate across the entire protein, which was 0.026 mutations/residue.

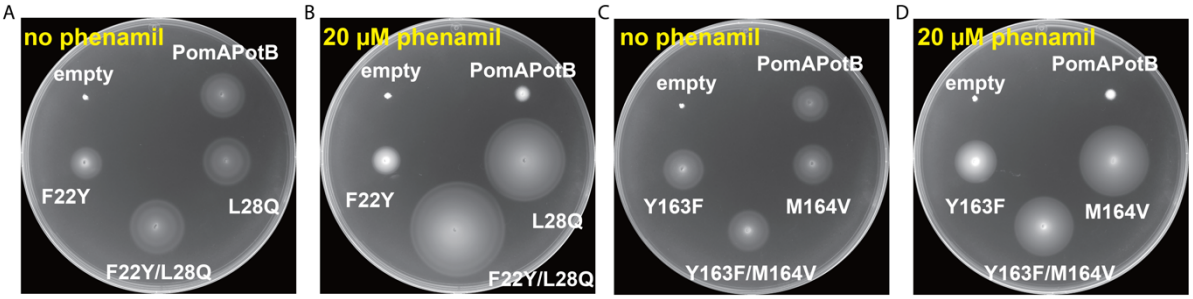
Point mutations in TM domain can enable motility in the presence of phenamil.

Out of the 27 strains (Table S1), 209 mutations were observed in total, with a median number of mutations per strain of 8. Of the 209 nucleotide mutations, 62 were silent mutations that did not result in an amino acid change. Six exhibited a strong motility phenotype when motility was characterised in the presence of 20 µM phenamil. Of particular interest, we focused on plasmid PotB-ep18 (Fig. S1-S2, Table S1) which contained only 3 mutations: F22Y, L28Q, K100Q and demonstrated increased motility in the presence of phenamil. To address the relative contributions of each mutation to the observed phenotype we synthesised single point mutants of each of the three mutations (Fig. S3). Both F22Y and L28Q alone enabled motility under phenamil to PotB (Fig. 3), or together (pSHU149, F22Y/L28Q), where they were slightly additive. Mutation K100Q appeared to offer no benefit in motility under phenamil and as such was classified a redundant mutation. These results indicate that single point mutations, particularly in the pore region of the TM domain of PotB, can enable motility under phenamil.

Point mutations near PG-binding domain enable motility in the presence of phenamil.

Plasmid PotB-ep9 (Fig. S1-S2; Table S1) displayed a phenamil-resistant phenotype and also had only three mutations: Y163F, M164V and L255M. This was targeted for further

171 investigation as these mutations were all far from the TM domain. We again characterised
172 the contributions of each mutation to motility and observed that Y163F and M164V each



173 separately enabled motility in the presence of phenamil and that L255M was a redundant
174 mutation (Fig. S3). When both mutations were combined in a single plasmid (pSHU146),
175 motility in the presence of phenamil was increased (Fig. 3).

176 **Figure 3: Mutants in TM-domain and PG-binding region enable motility under phenamil.** Motility of either TM-
177 domain mutants (A/B) or PG-binding mutants (C/D) was tested via 0.25% agar swim plating using the empty
178 vector (pBAD33) and PomAPotB wild-type as a negative and positive control respectively. A/C) In the absence
179 of phenamil, swimming of mutants is comparable to the positive control. B/D) In the presence of 20 µM
180 phenamil, the mutants are motile whilst PomAPotB wild-type is severely limited. A/C incubated at 30˚C for 10
181 hours, B/D incubated at 30˚C for 18 hours.

182

183 ***PG-binding double mutants rotate at lower sodium concentration than wild-type PotB***

184

185 To establish whether these mutants were generating torque from sodium motive force, or
186 had undergone a change in ion-specificity, we tested swimming in minimal sodium-free media
187 (Fig. S4). None of the mutants were motile in sodium-free media, yet all plasmids restored
188 swimming in the presence of sodium and in the presence and absence of phenamil (Fig S4).
189 We then tested individual cells for rotation whilst lowering the concentration of sodium to
190 determine whether the dependence on sodium-motive force (SMF) had changed in these
191 mutants (Fig. 4). This allowed us to confirm that our observed swim plate phenotype was
192 caused changes in flagellar rotation, and account for any compensatory changes in growth or
193 chemotaxis that might confound swim plate measurements. We examined the capacity of
194 'wild-type' PomAPotB (expressed via our plasmid pSHU1234, Fig. S5) and our mutant plasmids
195 to restore swimming in 'sticky-filament' strains. These sticky-filament mutants lack the
196 surface of the protein FliC and thus undergo hydrophobic interactions with glass that enable
197 filaments to stick to glass and allow the rotation of the body to be assessed in parallel
198 (Kuwajima 1988). In comparison with PomAPotB wild-type, the mutants in the PG-region
199 (pSHU146: Y163F/M164V) were highly motile at low sodium (further detail Fig. S6), with cells
200 rotating > 6 Hz in the presence of only 0.3 mM Na^+. This is in contrast to PomAPotB wild-type
201 restoring motility to < 2 Hz at the equivalent sodium motive force. In an opposite trend,
202 mutations in the TM region (pSHU149: F22Y/L28Q) did not demonstrate swimming below 1
203 mM Na^+ and had a lower maximum speed and slower response to increased [Na^+].

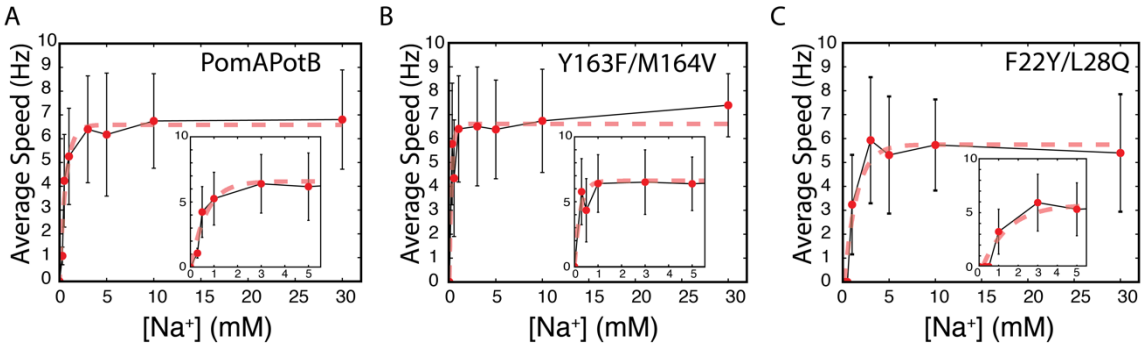


Figure 4: Y163F/M164V is motile at less than 0.3 mM [NaCl]. (A) The PomAPotB wild-type exhibits single cell rotation from 0.3 mM Na^+, increasing exponentially to a plateau above 6 Hz at full energisation. (B) PotB Y163F/M164V displays rotation at 0.3 mM Na^+, with plateau above 6 Hz at full energisation. (C) potB F22Y/L28Q displays no rotation below 1 mM Na^+, with a slower increase to a lower maximum speed plateau above 5 Hz and under 6 Hz. Inset in ABC shows zoom around 0-5 mM [Na^+]. Dashed lines indicate monoexponential fit ($\omega = ae^{-bx} + c$).

TM-domain double mutants are not affected by increasing concentrations of phenamil.

The response of single-cell speed with varying phenamil concentration was measured to examine the energy profile and structural basis of motility under phenamil in our two phenamil-resistant plasmids. In wild-type PomAPotB tethered cell rotation speed was reduced significantly when as little as 1 µM of phenamil was introduced, and cell speed plateaued at 5 µM of phenamil. The PG-binding double mutant (Y163F/M164V) demonstrated the same plateau effect in response to increasing phenamil but the concentration dependence of the decrease in rotational speed was much lower, that is, the decay constants for a monoexponential fit (fitting for *b* in: $\omega = ae^{-bx}$) were 0.8 mM^{-1} and 0.2 mM^{-1} for PomAPotB wild-type and Y163F/M164V respectively. Similarly, the I_{50}, the concentrations of phenamil at which the speed was half the maximum, was 0.9 µM and 3.1 µM for wild-type and Y163F/M164V respectively. In contrast, the speed of rotation for the double TM-domain mutant (F22Y/L28Q) exhibited no dependence on phenamil concentration up to 20 µM, and only a 46% reduction on maximum speed at 100 µM phenamil. This implies that the mechanism of resistance in the TM-domain mutants destroyed much of the efficacy of phenamil to inhibit motility altogether, whereas the PG-binding region mutants acted to stabilise rotation in the presence of phenamil, but this stabilisation could be overcome as phenamil concentration was increased.

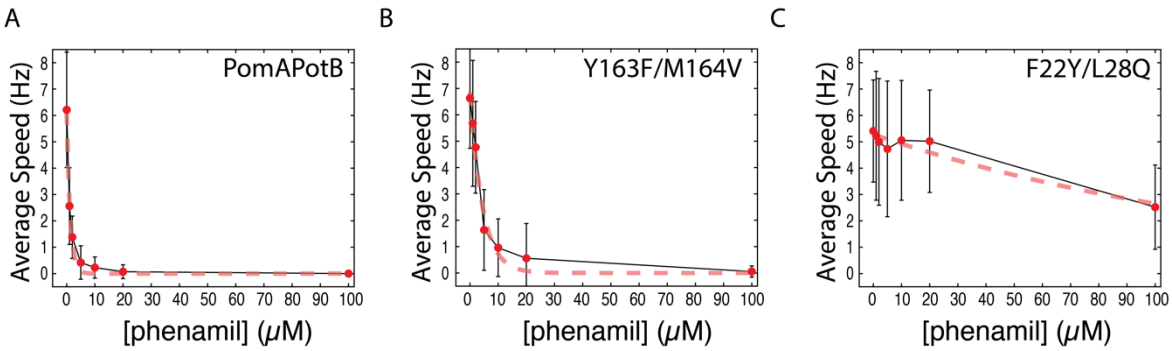


Figure 5: TM mutants enable motility in phenamil concentrations up to 100 µM. Dose response of averaged measurements of single-cell rotation speed with varying concentration of phenamil for: (A) PomAPotB wild-type, (B) double TM-domain mutant (PomAPotB F22Y/L28Q) and (C) double PG-binding mutant (PomAPotB

Y163F/M164V). Tethered cell data is fit with monoexponential decay curve to generate time constants for decay (PomAPotB wild-type: 0.8 mM^{-1}; PomAPotB Y163F/M164V: 0.2 mM^{-1}; PomAPotB F22Y/L28Q: 0.007 mM^{-1}).

F22Y is common and clusters in phylogenies along MotB/PomB lines; L28Q is very rare.

To explore the natural frequency of these residues and their correlation with ion specificity and motility outside the lab, a large low-level phylogeny of 948 MotB homologues was assembled which clearly divided into clades (Fig. 6A). We curated a subset of 34 commonly studied species where information on ion specificity exists in the literature (Fig. 6BCD). For each column of the subset alignment, we reconstructed ancestral residues with parsimony. At the F22Y site, residues clustered according to MotB/PomB identity (Fig. 6B), which is putatively linked to proton/sodium motility. To test if the correlation is greater than expected by chance, we used generalized linear models to perform phylogenetic and non-phylogenetic logistic regression (see Methods). This demonstrated that F/Y at position 22 had significant correlation with PomB/MotB respectively (Supplementary Table S3). In contrast, at L28Q (Fig 6D), the residue variation had no significant correlation with PomB/MotB. In fact, over the entire 948 species in the larger phylogeny, only one species had a glutamine at that site: *Myxococcus fulvus.*

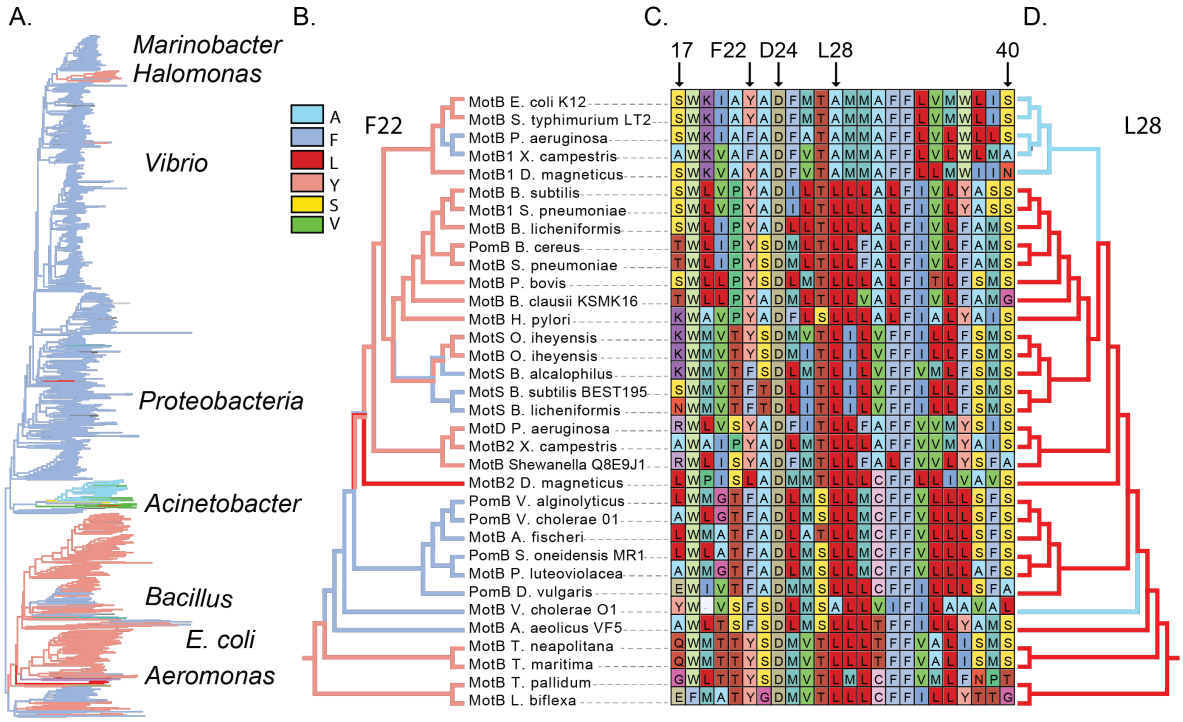


Figure 6: Parsimony Reconstruction for Y22 and L28. (A) Overall phylogeny of 948 MotB homologs coloured via sequence identity at site 22 showing clear signal (well-known clades are reproduced). Some major phyla indicated. (B) Parsimony reconstruction for site Y22 (PotB numbering) shown in colour over subset phylogeny of 34 representative species as labelled. (C) Sequence alignment across transmembrane domain of MotB/PomB showing conserved D24, and Y22/L28 mutations tested in this work. (D) Parsimony reconstruction for site L28A (PotB numbering) shown in colour over subset phylogeny.

Y163F/M164V are both rare and do not cluster along MotB/PomB lines.

Similarly, we aligned the sequences in the section 156-214, including the PG-binding site, and measured correlation with PomB/MotB (Fig. 7, Table S3). In both Y163F and M164V, logistic regression could not be run as these columns are not predominantly composed of 2 residues (our algorithm requires a binary predictor variable). However, visual inspection suggests the mutations do not cluster along PomB/MotB lines. A phenylalanine at site 163 was very rare across the full phylogeny, occurring only in *Yersinia pestis biovar Orientalis.* The valine at site 163 was observed in 13 species, mostly among members of the *Spirochaetes* and *Oceanospirilla* families (Supplementary Note).

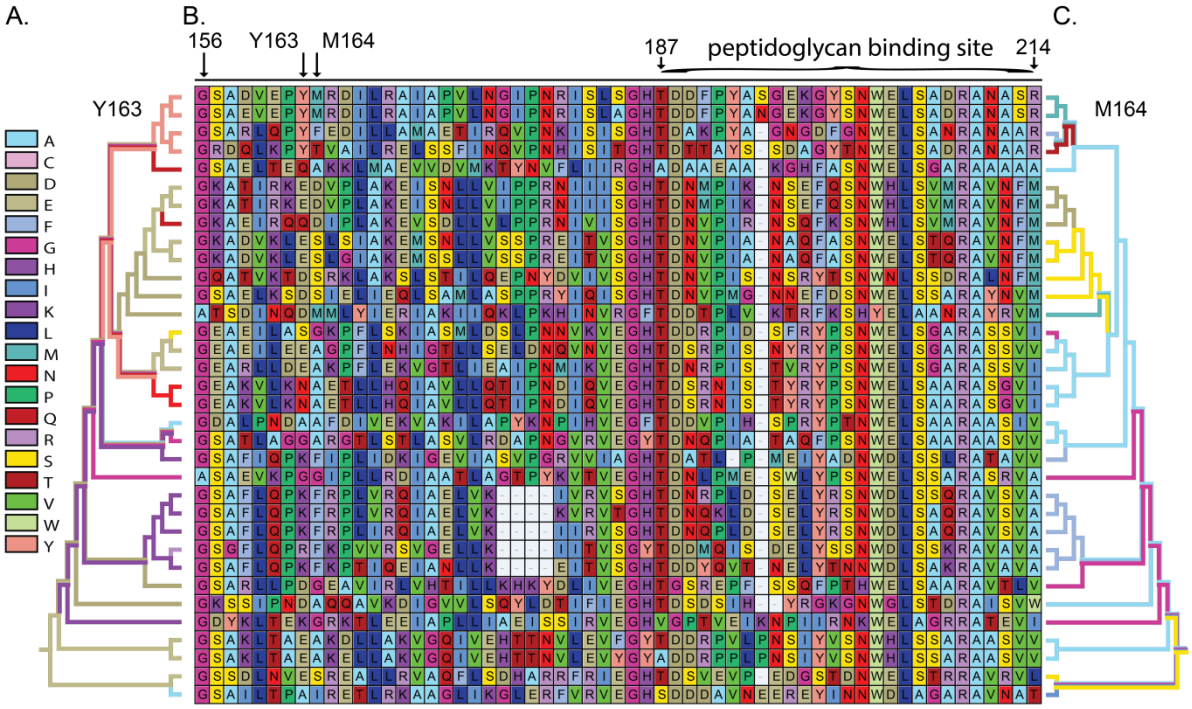


Figure 7: Parsimony Reconstruction for Y163 and M164. (A) Parsimony reconstruction for site Y163 (PotB numbering) shown in colour over subset phylogeny of 34 representative species from Fig. 5. (B) Sequence alignment over peptidoglycan binding site of MotB/PomB showing Y163/M164 mutations tested in this work. (C) Parsimony reconstruction for site M164 (PotB numbering) shown in colour over subset phylogeny.

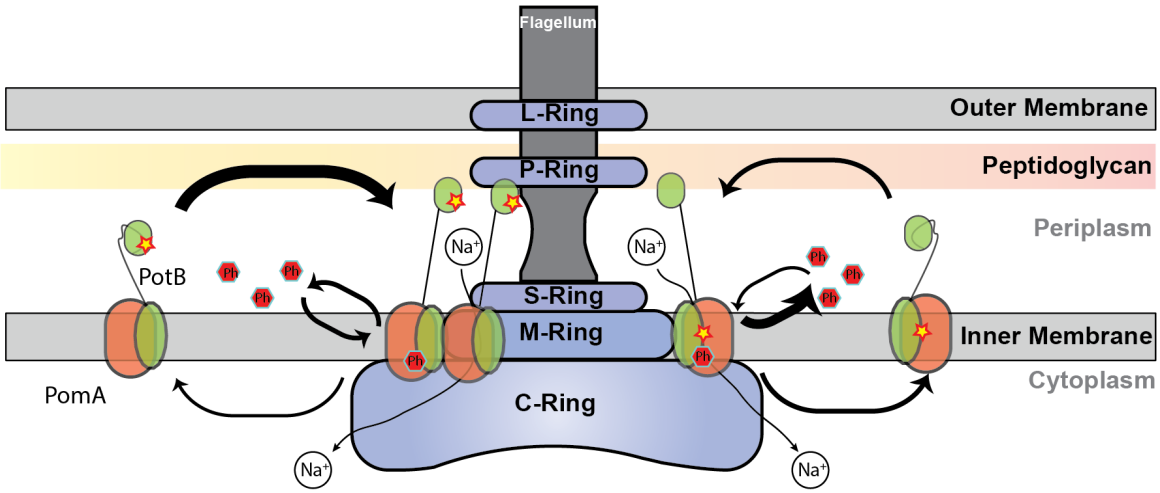


Figure 8: Model of activity. Schematic showing effects of mutations with respect to phenamil and sodium flow through stator complex. The pool of stator complexes ($PomA_4PotB_2$) diffuse in the inner membrane, and when they bind to the motor the PG-binding region pulls with tension on the transmembrane domains of PomA and PotB (Kojima et al. 2018). (left) Mutations in the 163/164 site affect PG-binding and the duration for which a stator is bound to the motor (dwell time). This could result in a slower off rate with respect to wild-type stators and a larger steady-state population of stators bound to the motor. In the presence of phenamil each stator has reduced functionality but in sum, with many stators, the motor is functional. (right) Mutations at the TM-domain near the putative phenamil binding site affect the binding and release of phenamil, thus phenamil is bound more weakly, and turns over more rapidly, allowing sodium ions to pass through the channel to generate motor rotation. An alternate possibility is that that mutations at 163/164 site cause a consequential alteration of the binding site and thus limit phenamil binding in a similar manner to mutations at 22/28 site.

Discussion

We have examined the effect of mutations in both the transmembrane domain and the peptidoglycan binding region of PotB to the binding of phenamil. Phenamil binding has been suggested to bind on the cytoplasmic face of the stator(S. Kojima et al. 1997, 1999; T. Yakushi, Kojima, and Homma 2004), and function by limiting Na^+ exit into the cell interior (Yoshida et al. 1990). However, the question remains how the mutations observed here, far from the pore, might influence phenamil binding and sodium transit. We observed that single point mutations in the TM region at L28 and F22 were sufficient to enable motility under phenamil. The mutation at F22Y is highly clustered around Pom/Mot stator subunits in varying species and presumably linked to sodium/proton specificity, which may show that residue F22 not only plays a role in sodium binding but also phenamil-binding and sodium blocking. Together with a mutation we observed as an error-prone PCR product at L28Q, these two mutations here have enabled motility under phenamil across a large range of phenamil concentrations, with stators remaining functional at concentrations of phenamil as high as 100 µM phenamil. This high tolerance to phenamil implies that they have affected or removed the binding site of phenamil inside the pore region of the stator complex.

Previous work in *V. Alginolyticus* extensively explored the effect of mutations at PomB site 22 on motility in the presence and absence of phenamil (Terauchi et al. 2011). Of these point mutations in PomB (as opposed to chimeric PotB), the substitutions F22S, F22W and F22N provided high motility under phenamil, however it was also observed that F22Y displayed a phenotype of motility under phenamil. Our results confirm the importance of F22 site in the

311 chimeric PotB and likely proximity to the Na^+ binding site and involvement in the ion-
312 conducting pathway, and show that even slight changes in the size of this residue can affect
313 the binding site and enable motility under phenamil.
314 Similarly, in the PG-binding region we observed that two clustered single mutations were also
315 sufficient to enable motility under phenamil. The PG-binding region is used to bind the stators
316 in place so that they act on the motor. With these two mutations at Y163/M164 we observed
317 a decaying concentration dependence of rotation to phenamil. These sites are far from the
318 pore, which is expected to be the phenamil binding site. We consider that mutations near the
319 PG region could allow motility under phenamil via two models: 1) changes to the PG region
320 can impact stator-rotor interface and alter the pore region and phenamil binding site; and 2)
321 that changes to the PG region can alter the binding strength of PotB to the PG layer in *E. coli*
322 and alter the dwell time of the stator on the rotor. The first option is supported by previous
323 evidence that a mutation in the PG region of MotB (P159I) could be suppressed by a mutation
324 on the rotor protein FliG (K192E) (Garza et al. 1995). This suggested that changes in the PG
325 region could influence the rotor-stator interface, and that this interface could be realigned
326 with a compensatory mutation in the rotor. Thus it is conceivable that our observed
327 mutations in the PG region (Y163/M164) are influencing the rotor-stator interface and thus
328 altering the ion channel and the phenamil binding site, in a similar mechanism to changed
329 directly at the binding site caused by L28Q and F22Y. The second model contends that a
330 change in the binding strength of the stator might increase the dwell time of the stator on the
331 rotor is result in more stators being bound more often (Fig. 8). In turn, if each stator is
332 fractionally functional, or occasionally able to pass a sodium ion in the presence of phenamil,
333 then in total, with more stators on the rotor, these sodium-powered flagellar motors are able
334 to function in the presence of phenamil. Recent structural evidence indicates that upon
335 binding to the PG layer there is a dynamic rearrangement where a helix (helix α1) is extended
336 into a non-helical structure to activates the stator and open the ion channel (Seiji Kojima et
337 al. 2018). Thus it is also conceivable that mutations in the PG region could affect the release
338 from the PG layer, the extension of helix α1, or the conformational change that removes the
339 plug from the channel. Further experiments are needed to distinguish between these models,
340 for example, the use of single cell studies and single molecule fluorescence to measure
341 changes in the dwell time of mutant stators on the motor, and tight regulation of stator
342 expression to measure for dependence of motility on stator type, concentration and
343 availability.
344 Phenamil resistance has been studied in other species, most notably in *Vibrio*
345 *parahaemolyticus*. This system is of interest as it has two motors, one sodium-powered and
346 one proton-powered and phenamil has been studied previously as a drug target for this
347 pathogen (Tatsuo Atsumi, McCartert, and Imae 1992; Jaques, Kim, and McCarter 1999).
348 Jaques et al. used natural selection over 12-18 days to screen for phenamil resistance (Jaques,
349 Kim, and McCarter 1999). They observed a single mutation in the TM-domain, but did not
350 observe the mutations that we saw in this paper. Their experiment, occurring through natural
351 mutagenesis and evolution in a laboratory setting will respond to pleiotropy and epistasis and
352 also inherent directionality in the process of stator adaptation. This means that the order of
353 specific mutations can restrict the possible outcomes. Naturally, our mutation at L28Q was
354 indeed very rare across our set of 948 MotB homologues, occurring only once in *Myxococcus*
355 *fulvus* ~0.1% of the species we examined. Error-prone PCR however, is agnostic to the order
356 of mutations and will insert mutations evenly across the protein of interest. This highlights
357 the power of error-prone PCR to explore novel mutations and to compare historical

358 evolutionary occurrences with those that can be discovered via directed and laboratory-
359 assisted experimental evolution.
360 Our experimental system allows us to examine phenotypes to determine which amino acids
361 have a functional cause in changing a phenotype. This allows us to address a common puzzle
362 in phylogenetics: because both mutations and traits are inherited on the phylogeny, they are
363 non-independent, and can thus exhibit "accidental" correlation without a functional
364 relationship (Uyeda, Zenil-Ferguson, and Pennell 2018). For example, in the case of MotB,
365 every clade has some amino acids unique to it, but this will be true whether or not those
366 residues have a functional role in the clade's motility phenotype. However, by characterising
367 upmotile mutations in the presence of phenamil and adaptations in ion-source, we can
368 directly test whether amino acid substitutions observed on the phylogeny are causally
369 associated with a particular ion source, or a particular environment. This allows us to identify
370 which amino acids, potentially in which order, have been part of historic adaptation events
371 such as a transition from sodium to proton-based motility. We can then use these approaches
372 to tease out the causality of events that constrain natural adaptation.
373
374 **Experimental Procedures**
375
376 *Bacterial strains and plasmids*
377 Bacterial strains and plasmids used in this study are listed in SI. For swim plating, plasmids
378 were transformed into RP6665 (Δ*motAmotB*) (Block, Blair, and Berg 1989) to restore motility.
379 For tethered cell assays, plasmids were transformed into JHC36 (Δ*cheY fliC-sticky* Δ*pilA*
380 Δ*motAmotB*) (Inoue et al. 2008).
381
382 *Error-prone PCR*
383 Plasmid pSHU1234 was constructed in a pBAD33 backbone with unique cut-sites for NdeI and
384 PstI directly upstream and downstream of PotB. Error-prone PCR was executed on pSHU1234
385 using primer 1176 and primer 0104 (Table S2), taq polymerase, unbalanced dNTPs and Mn^{2+},
386 as per (Wilson and Keefe 2000). The post-PCR product and pSHU1234 was then digested and
387 ligated with NdeI and PstI and transformed into RP6665(Δ*motAB*).
388
389 *Mutant isolation*
390 Tranformant was streaked in lines on low agarose motility plate (Fig. S1) consisting of 0.25%
391 TB soft agar (0.25% bactoagar, 1% bactotryptone, 0.5% NaCl), 25 µg/mL chloramphenicol, 1
392 mM arabinose and 20 µM phenamil. These plates were incubated at 30˚C for 20 hours. Flares
393 were selected from these plates and single colonies were generated. These colonies were
394 harvested, the plasmid was purified, transformed again into RP6665 and motility was tested
395 as below. Specific mutants derived from error-prone PCR outcomes were engineered using a
396 standard site-directed mutagenesis protocol with forward and reverse primers
397 (Supplementary Materials), template, and high-fidelity DNA polymerase (Pfu Ultra).
398
399 *Motility assays*
400 Motility testing was carried out using motility plate assays. Phenamil to a final concentration
401 of either 0, 20 or 100 µM was added to the plates. Swim plates were inoculated from single
402 colonies by toothpick. Plates were incubated for 10 or 18 hours respectively, as noted. For
403 standard motility assays, plates consisted of minimal media for minimal motility assays (Fig.
404 S4) consisted of 10 mM KPO_4, 0.1% glycerol, 0.1 mM Thr, 0.1 mM Leu, 0.1 mM His, 0.1 mM

405 Met, 0.1 mM Ser, 1 mM $MgSO_4$, 1 mM $(NH_2)_2SO_4$, 1 µg/mL Thiamine, 0.5% NaCl or KCl
406 respectively (85 mM [NaCl]; [67 mM KCl]), 1 mM arabinose, 25 µg/mL chloramphenicol and
407 the indicated concentration of phenamil (0, 20, 100 µM).
408
409 *Sequence analysis*
410 Sequencing was executed commercially by FASMAC (Japan) using forward and reverse
411 primers and template (Table S2). Three primers were used to read the entire PomAPotB
412 region.
413
414 *Single cell speed assays*
415 Cells were adhered spontaneously to glass coverslips via a sticky-filament and then imaged
416 onto a CMOS camera at 60 frames per second thorough 40x objective. Rotation of tethered
417 sticky-filament cells was analysed using custom software based upon LabView (National
418 Instruments).
419
420 *Phylogenetics and sequence diversity of MotB homologs*
421 To survey the known flagellar diversity for amino acid substitutions similar to the upmotile
422 mutations discovered by experiment, we estimated a very large phylogeny of MotB and
423 related proteins. As our goal was exploratory, we chose faster heuristic methods feasible for
424 a large dataset, rather than Maximum Likelihood or Bayesian methods that would be
425 impracticably slow. 948 MotB homologs were assembled by searching the UniProt90
426 database against key study taxa. They were aligned using Clustal Omega (Larkin et al. 2007)
427 on its most thorough settings (5 iterations of re-alignment). The phylogeny was then
428 estimated with Quicktree, a neighbour-joining method suitable for large datasets (Howe,
429 Bateman, and Durbin 2002). The phylogeny was midpoint-rooted in FigTree (Rambaut 2018).
430 As these are fast heuristic methods based on a single protein, deeper branches and rooting
431 should be considered uncertain, and we make no claim about e.g. inter-phylum relationships,
432 a very difficult phylogenetic problem (Pallen and Matzke 2006; Abby and Rocha 2012; Shih
433 and Matzke 2013; Koonin 2016). However, the tree was adequate for the purpose of showing
434 relationships between closer relatives, and for surveying MotB sequence conservation and
435 diversity.
436
437 To further explore conservation and diversity at key positions of the alignment, ancestral
438 amino acids were reconstructed on the phylogeny using parsimony in Mesquite (Maddison
439 and Maddison 2015). The 948-protein tree was subset to key study taxa for Fig. 6 and Fig. 7.
440
441 *Logistic Regression*
442 We classified each species in the 34-species subset phylogeny as Na+, H+, or N/A to create a
443 binary response variable for a logistic regression analysis. We were interested in cases where
444 this binary response might be correlated with a binary amino acid predictor, so we filtered
445 the columns of the subset alignment for columns where only 2 amino acids were
446 predominantly observed. We excluded any column where the two most common amino acids
447 added up to <85% of the observed residues. We also excluded any site where the first or
448 second most common amino acid had a <10% frequency. These steps excluded the 163/164
449 site, but included positions 22/28. We ran two models to correlate each amino acid or the
450 mutational pair with the ion source available: (1) non-phylogenetic logistic GLM (glm, R), and
451 (2) a phylogenetic logistic glm (phyloglm, R; Ives & Garland, 2010). We present the slope

452 estimates, errors, and *z* scores and *p*-values in Table S4, and our classification list of ionic
453 energy sources in Table S5.
454
455
456 **References**
457

458 Abby, Sophie S., and Eduardo P. C. Rocha. 2012. "The Non-Flagellar Type III Secretion
459 System Evolved from the Bacterial Flagellum and Diversified into Host-Cell Adapted
460 Systems." *PLOS Genetics* 8 (9): e1002983.
461 https://doi.org/10.1371/journal.pgen.1002983.
462 Asai, Yukako, Toshiharu Yakushi, Ikuro Kawagishi, and Michio Homma. 2003. "Ion-
463 Coupling Determinants of Na+-Driven and H+-Driven Flagellar Motors." *Journal of*
464 *Molecular Biology* 327 (2): 453–63. https://doi.org/{10.1016/S0022-2836(03)00096-
465 2}.
466 Atsumi, T., S. Sugiyama, E. J. Cragoe, and Y. Imae. 1990. "Specific Inhibition of the Na(+)-
467 Driven Flagellar Motors of Alkalophilic Bacillus Strains by the Amiloride Analog
468 Phenamil." *Journal of Bacteriology* 172 (3): 1634–39.
469 Atsumi, Tatsuo, Linda McCartert, and Yasuo Imae. 1992. "Polar and Lateral Flagellar
470 Motors of Marine Vibrio Are Driven by Different Ion-Motive Forces." *Nature* 355
471 (6356): 182–84. https://doi.org/10.1038/355182a0.
472 Block, Steven M., David F. Blair, and Howard C. Berg. 1989. "Compliance of Bacterial
473 Flagella Measured with Optical Tweezers." *Nature* 338 (6215): 514–18.
474 https://doi.org/10.1038/338514a0.
475 Chaban, Bonnie, Izaak Coleman, and Morgan Beeby. 2018. "Evolution of Higher Torque in
476 *Campylobacter*-Type Bacterial Flagellar Motors." *Scientific Reports* 8 (January): 97.
477 https://doi.org/10.1038/s41598-017-18115-1.
478 Garvin, J. L., S. A. Simon, E. J. Jr Cragoe, and L. J. Mandel. 1985. "Phenamil: An
479 Irreversible Inhibitor of Sodium Channels in the Toad Urinary Bladder." *J Membr*
480 *Biol* 87 (1): 45–54.
481 Garza, A. G., L. W. Harris-Haller, R. A. Stoebner, and M. D. Manson. 1995. "Motility
482 Protein Interactions in the Bacterial Flagellar Motor." *Proceedings of the National*
483 *Academy of Sciences* 92 (6): 1970–74. https://doi.org/10.1073/pnas.92.6.1970.
484 Howe, Kevin, Alex Bateman, and Richard Durbin. 2002. "QuickTree: Building Huge
485 Neighbour-Joining Trees of Protein Sequences." *Bioinformatics* 18 (11): 1546–47.
486 https://doi.org/10.1093/bioinformatics/18.11.1546.
487 Imazawa, Riku, Yuka Takahashi, Wataru Aoki, Motohiko Sano, and Masahiro Ito. 2016. "A
488 Novel Type Bacterial Flagellar Motor That Can Use Divalent Cations as a Coupling
489 Ion." *Scientific Reports* 6 (January): 19773. https://doi.org/10.1038/srep19773.
490 Inoue, Yuichi, Chien-Jung Lo, Hajime Fukuoka, Hiroto Takahashi, Yoshiyuki Sowa, Teuta
491 Pilizota, George H. Wadhams, Michio Homma, Richard M. Berry, and Akihiko
492 Ishijima. 2008. "Torque-Speed Relationships of Na+-Driven Chimeric Flagellar
493 Motors in Escherichia Coli." *Journal of Molecular Biology* 376 (5): 1251–59.
494 https://doi.org/10.1016/j.jmb.2007.12.023.
495 Ives, A. R. and T. Garland, Jr. 2010. "Phylogenetic logistic regression for binary dependent
496 variables." *Systematic Biology* 59:9-26.
497 Jaques, S., Y. K. Kim, and L. L. McCarter. 1999. "Mutations Conferring Resistance to
498 Phenamil and Amiloride, Inhibitors of Sodium-Driven Motility of Vibrio
499 Parahaemolyticus." *Proceedings of the National Academy of Sciences of the United*
500 *States of America* 96 (10): 5740–45.
501 Kojima, S., Y. Asai, T. Atsumi, I. Kawagishi, and M. Homma. 1999. "Na+-Driven Flagellar
502 Motor Resistant to Phenamil, an Amiloride Analog, Caused by Mutations in Putative

Channel Components." *Journal of Molecular Biology* 285 (4): 1537–47. https://doi.org/10.1006/jmbi.1998.2377.

Kojima, S., T. Atsumi, K. Muramoto, S. Kudo, I. Kawagishi, and M. Homma. 1997. "Vibrio Alginolyticus Mutants Resistant to Phenamil, a Specific Inhibitor of the Sodium-Driven Flagellar Motor." *Journal of Molecular Biology* 265 (3): 310–18. https://doi.org/10.1006/jmbi.1996.0732.

Kojima, Seiji, and David F. Blair. 2001. "Conformational Change in the Stator of the Bacterial Flagellar Motor." *Biochemistry* 40 (43): 13041–50. https://doi.org/10.1021/bi011263o.

Kojima, Seiji, Yukio Furukawa, Hideyuki Matsunami, Tohru Minamino, and Keiichi Namba. 2008. "Characterization of the Periplasmic Domain of MotB and Implications for Its Role in the Stator Assembly of the Bacterial Flagellar Motor." *Journal of Bacteriology* 190 (9): 3314–22. https://doi.org/10.1128/JB.01710-07.

Kojima, Seiji, Katsumi Imada, Mayuko Sakuma, Yuki Sudo, Chojiro Kojima, Tohru Minamino, Michio Homma, and Keiichi Namba. 2009. "Stator Assembly and Activation Mechanism of the Flagellar Motor by the Periplasmic Region of MotB." *Molecular Microbiology* 73 (4): 710–18. https://doi.org/10.1111/j.1365-2958.2009.06802.x.

Kojima, Seiji, Masato Takao, Gaby Almira, Ikumi Kawahara, Mayuko Sakuma, Michio Homma, Chojiro Kojima, and Katsumi Imada. 2018. "The Helix Rearrangement in the Periplasmic Domain of the Flagellar Stator B Subunit Activates Peptidoglycan Binding and Ion Influx." *Structure (London, England: 1993)* 26 (4): 590-598.e5. https://doi.org/10.1016/j.str.2018.02.016.

Koonin, Eugene V. 2016. "Horizontal Gene Transfer: Essentiality and Evolvability in Prokaryotes, and Roles in Evolutionary Transitions." *F1000Research* 5 (July): 1805. https://doi.org/10.12688/f1000research.8737.1.

Kuwajima, G. 1988. "Construction of a Minimum-Size Functional Flagellin of Escherichia Coli." *Journal of Bacteriology* 170 (7): 3305–9.

Larkin, M. A., G. Blackshields, N. P. Brown, R. Chenna, P. A. McGettigan, H. McWilliam, F. Valentin, et al. 2007. "Clustal W and Clustal X Version 2.0." *Bioinformatics* 23 (21): 2947–48. https://doi.org/10.1093/bioinformatics/btm404.

Lo, Chien-Jung, Yoshiyuki Sowa, Teuta Pilizota, and Richard M. Berry. 2013. "Mechanism and Kinetics of a Sodium-Driven Bacterial Flagellar Motor." *Proceedings of the National Academy of Sciences* 110 (28): E2544–51. https://doi.org/10.1073/pnas.1301664110.

Maddison, WP, and DP Maddison. 2015. *Mesquite: A Modular System for Evolutionary Analysis. Version 3.04*. https://www.mesquiteproject.org/.

Minamino, Tohru, Naoya Terahara, Seiji Kojima, and Keiichi Namba. 2018. "Autonomous Control Mechanism of Stator Assembly in the Bacterial Flagellar Motor in Response to Changes in the Environment." *Molecular Microbiology* 109 (6): 723–34. https://doi.org/10.1111/mmi.14092.

Pallen, Mark J., and Nicholas J. Matzke. 2006. "From The Origin of Species to the Origin of Bacterial Flagella." *Nature Reviews. Microbiology* 4 (10): 784–90. https://doi.org/10.1038/nrmicro1493.

Paulick, Anja, Nicolas J. Delalez, Susanne Brenzinger, Bradley C. Steel, Richard M. Berry, Judith P. Armitage, and Kai M. Thormann. 2015. "Dual Stator Dynamics in the Shewanella Oneidensis MR-1 Flagellar Motor." *Molecular Microbiology* 96 (5): 993–1001. https://doi.org/10.1111/mmi.12984.

Rambaut, Andrew. (2015) 2018. *FigTree v1.4.4, a Graphical Viewer of Phylogenetic Trees. 2014*. Java. https://github.com/rambaut/figtree.

Rossmann, F. M., and M. Beeby. 2018. "Insights into the Evolution of Bacterial Flagellar Motors from High-Throughput in Situ Electron Cryotomography and Subtomogram

Averaging." *Acta Crystallographica Section D: Structural Biology* 74 (6): 585–94. https://doi.org/10.1107/S2059798318007945.

Roujeinikova, Anna. 2008. "Crystal Structure of the Cell Wall Anchor Domain of MotB, a Stator Component of the Bacterial Flagellar Motor: Implications for Peptidoglycan Recognition." *Proceedings of the National Academy of Sciences* 105 (30): 10348–53. https://doi.org/10.1073/pnas.0803039105.

Sharp, Leslie L., Jiadong Zhou, and David F. Blair. 1995. "Tryptophan-Scanning Mutagenesis of MotB, an Integral Membrane Protein Essential for Flagellar Rotation in Escherichia Coli." *Biochemistry* 34 (28): 9166–71. https://doi.org/10.1021/bi00028a028.

Shih, Patrick M., and Nicholas J. Matzke. 2013. "Primary Endosymbiosis Events Date to the Later Proterozoic with Cross-Calibrated Phylogenetic Dating of Duplicated ATPase Proteins." *Proceedings of the National Academy of Sciences* 110 (30): 12355–60. https://doi.org/10.1073/pnas.1305813110.

Sowa, Yoshiyuki, and Richard M. Berry. 2008. "Bacterial Flagellar Motor." *Quarterly Reviews of Biophysics* 41 (02): 103–32. https://doi.org/10.1017/S0033583508004691.

Sowa, Yoshiyuki, Alexander D. Rowe, Mark C. Leake, Toshiharu Yakushi, Michio Homma, Akihiko Ishijima, and Richard M. Berry. 2005. "Direct Observation of Steps in Rotation of the Bacterial Flagellar Motor." *Nature* 437 (7060): 916–19. https://doi.org/10.1038/nature04003.

Sugiyama, S., E. J. Cragoe, and Y. Imae. 1988. "Amiloride, a Specific Inhibitor for the Na+-Driven Flagellar Motors of Alkalophilic Bacillus." *The Journal of Biological Chemistry* 263 (17): 8215–19.

Terashima, Hiroyuki, Akihiro Kawamoto, Yusuke V. Morimoto, Katsumi Imada, and Tohru Minamino. 2017. "Structural Differences in the Bacterial Flagellar Motor among Bacterial Species." *Biophysics and Physicobiology* 14 (0): 191–98. https://doi.org/10.2142/biophysico.14.0_191.

Terauchi, Takashi, Hiroyuki Terashima, Seiji Kojima, and Michio Homma. 2011. "A Conserved Residue, PomB-F22, in the Transmembrane Segment of the Flagellar Stator Complex, Has a Critical Role in Conducting Ions and Generating Torque." *Microbiology* 157 (8): 2422–32. https://doi.org/10.1099/mic.0.048488-0.

Uyeda, Josef C., Rosana Zenil-Ferguson, and Matthew W. Pennell. 2018. "Rethinking Phylogenetic Comparative Methods." *Systematic Biology* 67 (6): 1091–1109. https://doi.org/10.1093/sysbio/syy031.

Wilson, David, and Anthony D. Keefe. 2000. "Random Mutagenesis by PCR." In , edited by Frederick M. Ausubel, Roger Brent, Robert E. Kingston, David D. Moore, J. G. Seidman, John A. Smith, and Kevin Struhl. 8.3. John Wiley & Sons.

Yakushi, T., M. Kojima, and M. Homma. 2004. "Isolation of Vibrio Alginolyticus Sodium-Driven Flagellar Motor Complex Composed of PomA and PomB Łdots." *Microbiology*, January. http://mic.sgmjournals.org/cgi/content/abstract/150/4/911.

Yakushi, Toshiharu, Junghoon Yang, Hajime Fukuoka, Michio Homma, and David F. Blair. 2006. "Roles of Charged Residues of Rotor and Stator in Flagellar Rotation: Comparative Study Using H+-Driven and Na+-Driven Motors in Escherichia Coli." *Journal of Bacteriology* 188 (4): 1466–72. https://doi.org/10.1128/JB.188.4.1466-1472.2006.

Yonekura, Koji, Saori Maki-Yonekura, and Michio Homma. 2011. "Structure of the Flagellar Motor Protein Complex PomAB: Implications for the Torque-Generating Conformation." *Journal of Bacteriology* 193 (15): 3863–70. https://doi.org/10.1128/JB.05021-11.

Yorimitsu, Tomohiro, and Michio Homma. 2001. "Na+-Driven Flagellar Motor of Vibrio." *Biochimica et Biophysica Acta (BBA) - Bioenergetics* 1505 (1): 82–93. https://doi.org/10.1016/S0005-2728(00)00279-6.

Yoshida, S., S. Sugiyama, Y. Hojo, H. Tokuda, and Y. Imae. 1990. "Intracellular Na+ Kinetically Interferes with the Rotation of the Na(+)-Driven Flagellar Motors of Vibrio Alginolyticus." *The Journal of Biological Chemistry* 265 (33): 20346–50.

Zhu, Shiwei, Masato Takao, Na Li, Mayuko Sakuma, Yuuki Nishino, Michio Homma, Seiji Kojima, and Katsumi Imada. 2014. "Conformational Change in the Periplamic Region of the Flagellar Stator Coupled with the Assembly around the Rotor." *Proceedings of the National Academy of Sciences* 111 (37): 13523–28. https://doi.org/10.1073/pnas.1324201111.

Author Contributions

TI collected and analysed data. RI collected and analysed data. JC collected data. NJM analysed data and wrote the manuscript. YS conceived and supervised the project, collected and analysed data, and wrote the manuscript. MABB conceived and supervised the project, collected and analysed data, and wrote the manuscript.

Funding Sources

NJM was supported by ARC DECRA fellowship DE150101773 and Marsden grants 16-UOA-277 and 18-UOA-034. YS was supported by JSPS KAKENHI (18H02475 and 15K07034), MEXT KAKENHI (15H01332), Itoh Science Foundation, and Casio Science Promotion Foundation. MABB was supported by a UNSW Scientia Research Fellowship, a CSIRO Synthetic Biology Future Science Platform 2018 Project Grant and ARC Discovery Project DP190100497.