Ancestral reconstruction of the MotA stator subunit reveals that conserved residues far from the pore are required to drive flagellar motility

Md Imtiazul Islam[1], Pietro Ridone[1], Angela Lin[1], Katharine A Michie[1,2], Nicholas J Matzke[3], Georg Hochberg[4,5], Matthew AB Baker[1,6].

Affiliations:

[1] *School of Biotechnology and Biomolecular Sciences (BABS), University of New South Wales, Sydney, Australia*
[2] *Structural Biology Facility, Mark Wainwright Analytical Centre, University of New South Wales, Sydney, Australia*
[3] *School of Biological Sciences, University of Auckland, New Zealand*
[4] *Center for Synthetic Microbiology (SYNMIKRO), Philipps-University Marburg; Karl-von-Frisch-Str. 14, 11 35043 Marburg, Germany*
[5] *Max-Planck-Institute for terrestrial microbiology, Karl-von-Frisch Str. 10, 35043 Marburg, Germany.*
[6] *Centre of Excellence in Synthetic Biology, Australia.*

*Correspondence: Matthew A.B. Baker

Email: matthew.baker@unsw.edu.au

Abstract

The bacterial flagellar motor (BFM) is a rotary nanomachine powered by the translocation of ions across the inner membrane through the stator complex. The stator complex consists of two membrane proteins: MotA and MotB (in H^+ powered motors), or PomA and PomB (in Na^+ powered motors). In this study we used ancestral sequence reconstruction (ASR) to probe which residues of MotA correlate with function and may have been conserved to preserve motor function. We reconstructed ten ancestral sequences of MotA and found four of them were motile in combination with contemporary *E. coli* MotB and in combination with our previously published functional ancestral MotBs. Sequence comparison between wild-type (WT) *E. coli* MotA and MotA-ASRs revealed 30 critical residues across multiple domains of MotA that were conserved among all motile stator units. These conserved residues included pore-facing, cytoplasm-facing and MotA-MotA intermolecular facing sites. Overall, this work demonstrates the role of ASR in assessing conserved variable residues in a subunit of a molecular complex.

Introduction

The bacterial flagellar motor (BFM) is a rotary molecular motor responsible for motility in many bacteria aiding their survival and pathogenicity. The BFM is one of the largest (11 MDa), dynamically self-assembled, and membrane-spanning molecular machines in bacteria, requiring more than a dozen different proteins to assemble and function (Sowa and Berry, 2008; Beeby et al., 2020). It contains a cytoplasmic rotor and a ring of stator units that surround the rotor (Berg, 2003; Nakamura and Minamino, 2019). The stator powers the motor by transporting ions across the cell membrane, generating torque in interaction with the rotor, which is coupled to the flagellar filament (Minamino and Imada, 2015).

The stator of the BFM is a complex of two membrane proteins: MotA and MotB in H^+ powered motors in *Escherichia coli*, and PomA and PomB in Na^+ powered motors in *Vibrio alginolyticus* (Yorimitsu and Homma, 2001; Berg, 2003). MotA and MotB form a heterodimer with a stoichiometry of $MotA_5MotB_2$ that creates an ion channel at the rotor interface (Deme et al., 2020; Santiveri et al., 2020). The MotA subunit contains four transmembrane (TM) helices (TMH1–TMH4), two short periplasmic loops, and a cytoplasmic loop between TMH2 and TMH3 and the C-terminal cytoplasmic tail (Zhou and Blair, 1997; Hu et al., 2022). The cytoplasmic loop contains highly conserved charged residues that interact with the rotor protein FliG (Zhou et al., 1998a). Conversely, the MotB subunit contains a single TM helix that lines the ion pathway alongside the TMH3 and TMH4 helices of MotA (Kojima and Blair, 2001). Adjacent to the pore-lining MotB TM domain is a periplasmic region containing a plug that can block the pore with a spanner-like mechanism (Homma et al., 2021). A peptidoglycan-binding domain (PGD) anchors the stator through binding to the PG-layer (Hosking et al., 2006). Interactions between MotA and the rotor trigger a structural change of the PGD and the plug region, that allows binding of PGD to the peptidoglycan (PG) layer, unplugging of the pore and activating the stator complex to allow ion flow and torque generation (Kojima et al., 2018).

The TM domains of both MotA and MotB contain a series of conserved residues that are known to be important for flagellar function, ion selection and transfer (Asai et al., 2000a; Ito et al., 2005). The aspartic acid residue 32 at the TM region of MotB (MotB-D32) acts as a universally conserved ion-binding site (Zhou et al., 1998b; Che et al., 2008), and the D32s from both MotB subunits are exposed to the pentameric MotA ring at the subunit interface in proximity of MotA-T209 (Deme et al., 2020). Among other reported conserved and critical residues of the stator, MotA-R90 and MotA-E98 contribute to motor rotation with rotor protein FliG (Zhou and Blair, 1997; Takekawa et al., 2014), MotA-P173 and MotA-P222

interact with adjacent MotA monomers stabilizing the structure (Braun et al., 1999; Deme et al., 2020), and MotA-A180 is involved in ion specificity with MotB-Y30 (Biquet-Bisquert et al., 2021). Over the years, many mutagenesis techniques have been used to study the structural residues of BFM, such as random mutagenesis (Blair and Berg, 1990), tryptophan-scanning mutagenesis (Sharp et al., 1995), cysteine-scanning mutagenesis (Asai et al., 2000b), site-directed mutagenesis (Nakamura et al., 2009), photo cross-linking and disulfide cross-linking (Terashima et al., 2021). In this study, we used ancestral sequence reconstruction (Hochberg and Thornton, 2017; Garcia and Kaçar, 2019; Dube et al., 2022) to generate hypothetical ancestral states from phylogenies of extant proteins. We reconstructed and synthesized ten ancestral MotAs to determine which residues of MotA were conserved throughout the evolution of BFM and were essential for BFM function.

Results

Phylogeny of MotA and the selected Nodes for resurrection.

Two phylogenetic trees were calculated for two separate multiple sequence alignments comprising of 178 and 264 MotA homologous sequences (Fig. 1) Ten nodes across both phylogenies were selected for ancestral sequence reconstruction (ASR), focusing on a mixture of older and younger ancestors across known sodium- and proton- motile clades. Among the ten selected nodes, ASR180, ASR220, ASRN41, and ASR333 had descendants that only belonged to proteobacteria, whereas ASR244, ASRN65, ASR259, ASR266, ASR332, and ASR440 had descendants that also included terrabacteria (actinobacteria, firmicutes), aquificae, and spirochaetes. Pairwise internode distance and pairwise percent coidentity between each node is shown in Supplementary Fig. 1 & 2; posterior probability across each site for ASRs is shown in Supplementary Fig. 3. We performed structural prediction for all reconstructed MotA-ASRs using ColabFold (Mirdita et al., 2021, 2022) prior to synthesis to compare predicted structures of the ASRs with existing known MotA structures (PDBIDSupplementary Fig. 4-6).

Functional characterisation of MotA-ASRs.

We evaluated the functional properties of all ten reconstructed MotA-ASR sequences by testing their motility through plasmid-based expression in a Δ*motA* strain. Overall, we characterised semi-solid agar swimming, cell free swimming, and tethered-cell rotation in different concentrations of sodium. Swim plate assay results indicated that three out of the ten MotA-ASRs (ASR180, ASR220, and ASR333) were able to restore motility in combination

with contemporary WT *E. coli* MotB in Δ*motA* RP437 (Fig. 2A). ASR180, ASR220 and ASR333 were also motile in the free-swimming assay with mean speeds of 1.6 ± 0.8, 3.0 ± 1.3, and 2.7 ± 1.1 µm/s, respectively (Fig. 2B). In comparison, the Na^+ (*pomApotB*) and H^+ (*motAmotB*) powered swimming controls had an average speed of 3.3 ± 1.3, and 5.2 ± 1.5 µm/s respectively. Both motile and non-motile MotA-ASRs exhibited similar growth rates in combination with WT *E. coli* MotB in Δ*motA* RP437 (Supplementary Fig. 7).

Ionic power source driving rotation in functional MotA- ASRs.

We evaluated the ionic power source (H^+ or Na^+) of our motile MotA-ASRs in low Na^+ conditions (~1 mM) using a minimal swim plate without tryptone or yeast extract and where NaCl had been substituted with KCl (Islam et al., 2020; Ridone et al., 2022). Here, we checked whether the tested strains could generate swim rings or not in the low sodium environment. The result showed that the H^+-powered control (*motAmotB*) and each of the three motile MotA-ASRs (ASR180, ASR220 and ASR333) produced swimming rings on the minimal swim plate (Fig. 3A). In contrast, the Na^+ powered control (*pomApotB*) and other non-motile MotA-ASRs did not produce any swimming rings (Fig. 3A). Tethered cell assay results agreed: the rotational speeds of ASR180, ASR220, ASR333 and the H^+ powered control remained constant as sodium concentration was varied and were 2.0 ± 0.9 Hz, 2.9 ± 0.7 Hz, 2.75 ± 0.7 Hz and 3.8 ± 0.7 Hz respectively (Figure 3B). However, the rotation speed of the Na^+ powered control increased from stationary (0 Hz) to 2.8 ± 0.7 Hz with the increase of external sodium concentration from 0 to 85 mM (Fig. 3B). H^+ conductivity was further evaluated by examining growth curves for all MotA-ASRs when combined with plug-deleted $MotB_{\Delta 51\text{-}70}$. Growth of WT *E. coli* MotA and the three motile MotA-ASRs (ASR180, ASR220 and ASR333) was inhibited when combined with $MotB_{\Delta 51\text{-}70}$ (Figure 3CD and Supplementary Fig. 8). In contrast, growth was not inhibited when non-motile MotA-ASRs were combined with plug-deleted $MotB_{\Delta 51\text{-}70}$ (Figure 3D and Supplementary Fig. 8).

Functional compatibility of MotA-ASRs with MotB-ASRs and *Aquifex* MotBs.

We tested multiple stator combinations by examining a 150 element swim-plate array. We examined all combinations of 10 MotA-ASRs with 15 ancestral MotBs consisting of our previously published 13 MotB-ASRs and 2 *Aquifex* MotBs (Islam et al., 2020) in a Δ*motAmotB* strain to evaluate their intercompatibility. Four out of ten of the MotA-ASRs (ASR180, ASR220, ASR333 and ASRN41) were compatible with several MotB-ASRs and formed functional stator units. The expression of MotA-ASR220 restored motility with each of the 13 MotB-ASRs. In contrast, MotA-ASR180, MotA-ASR333, and MotA-N41 were able to restore motility with 11, 11, and 1 of 13 MotB-ASRs, respectively (Fig. 4). The other

six MotA-ASRs, ASR244, ASR259, ASR266, ASR332, ASR440, and ASRN65 did not show compatibility with any of the 13 MotB-ASRs and were non-motile (Fig. 4). None of the ten MotA-ASRs restored motility when combined with either of the two Aquifex MotBs (Fig. 4).

Conservation of residues between WT *E. coli* - MotA and the motile MotA-ASRs.

We examined residue conservation between WT *E. coli* MotA and four motile MotA-ASRs (ASR180, ASR220, ASR333 and ASRN41) by performing a multiple sequence alignment. The alignment of WT *E. coli* MotA and four motile MotA-ASRs (Supplementary Fig. 9) showed that 81 of the 295 sites (27.5%) in all motile MotA-ASRs were identical, 96 sites (32.5%) were similar (i.e., were replaced, relative to WT, with a residue with similar biochemical properties), and 62 sites (21%) were a mixture of both similar and distinct sites across all the motile MotA-ASRs when compared with *E. coli* MotA. The remaining 56 (19%) of 295 residues were distinct when compared with *E. coli*-MotA at the same position (i.e., were substitutions of amino acids with notable changes in hydrophobicity, polarity, or size, in all four functional ASRs) (Fig. 5). Sequence comparison and conservation was mapped onto an Alphafold model (Jumper et al., 2021) of the WT *E. coli* $MotA_5MotB_2$ complex (Fig. 5). The pLDDT and PAE plots of the *E. coli* MotAB Alphafold model, with colour-coded confidence overlay onto the structure are provided in Supplementary Fig. 10. The RMSDs for each chain of the *E. coli* MotAB Alphafold model in comparison with each of 6YKP, 6YKM and 6YSL (reported MotAB structures of *C. jejuni* and *B. subtilis*) are provided in Supplementary Table 1.

Determination of the residues critical for function from comparison of wild-type MotA with both motile and non-motile MotA-ASRs.

Sequence comparison between our resurrected MotA-ASRs and WT *E. coli* MotA revealed several residues critical to MotA function. Sequence comparison highlighted that 30 (10%) of the 295 residues in the motile MotA-ASRs differed from the corresponding residues in the non-motile MotA-ASRs, but were identical to the corresponding residues of WT *E. coli* MotA (Supplementary Fig. 11). We located each of the 30 identified critical residues on our *E. coli* $MotA_5MotB_2$ structural model and observed that these critical residues were distributed throughout the MotA domains (Fig. 6). Among the 30 residues, four were found in transmembrane helix 2 (TMH2), eight in cytoplasmic helix 1-3 (CPH1-3), four in transmembrane helix 3 (TMH3), six in transmembrane helix 4 (TMH4) and eight in cytoplasmic helix 4 (CPH4) (Fig. 6 and Supplementary Fig. 11). We measured the

intermolecular distances between our conserved residues and several previously reported residues that are essential to BFM function (Supplementary Table 2) and observed that these residues were in close proximity (ranging from 2.3-7.1 Å). In particular, we observed the proximity of MotA-F45 with MotA-I54 and V168, MotA-D128 with MotA-E262, MotA-Y217 with MotA-P173 and MotB-W26, and MotA-L223 with MotA-L64 (Supplementary Fig. 12).

Experimental verification of the role of the selected conserved residues in flagellar function.

We then tested whether the introduction of point mutations at specific residues could rescue motility in non-motile MotA-ASRs, or disrupt motility in either motile MotA-ASRs or WT *E. coli* MotA. We selected six MotA residues (A40, V178, A179, Y217, G218 and R262) of interest based on previous reports (Garza et al., 1996; Deme et al., 2020; Biquet-Bisquert et al., 2021). Replacement of residues at selected positions in WT *E. coli* and motile MotA-ASRs with residues from non-motile MotA-ASRs at the same site showed that V178I and A179G point mutations did not disrupt motility, but the Y217N mutation completely disrupted motility (Figure 7A and Supplementary Table 3). In contrast, the reintroduction of residues from motile MotA-ASRs at the selected position of non-motile MotA-ASRs showed that none of the tested mutations (G40A, I178V, G179A, N217Y, L/M/V/A218G, and K262R), individually or in combination, were able to rescue motility (Figure 7B and Supplementary Table 3).

Discussion

This study 1) verifies the roles of MotA residues previously reported as crucial, 2) provides evidence as to which regions of MotA are variable, and 3) reveals new residues that are conserved and essential for flagellar function. First, several of our identified conserved residues, including MotA-V178, MotA-A178, and MotA-Y217 (Supplementary Fig. 7), are located near the pore in transmembrane helices 3 and 4 (TM3/TM4) and these have been widely reported to be important for stator function (Zhou and Blair, 1997; Braun et al., 1999; Kim et al., 2008; Nakamura et al., 2009; Deme et al., 2020). Other conserved residues are found in the cytoplasmic region and in TM helix-2 (TM2), and a number of those, such as MotA-E33, MotA-A40, MotA-L81, MotA-Y83, MotA-D128, MotA-E142, MotA-E144, and MotA-R262, have also been previously shown to be important for function (Garza et al.,

1996; Togashi et al., 1997; Zhou and Blair, 1997; Hosking and Manson, 2008; Komatsu et al., 2016).

Secondly, this work adds to this list by confirming through point mutation that MotA-Y217N that MotA-Y217 are crucial for function. This finding is similar to the previous report where Deme et al. confirmed the importance of the contact between MotA-Y217 and MotB-W26 via MotB-W26A point mutation (Deme et al., 2020). We further confirmed via point mutations V178I and A179G that MotA-V178 and MotA-A179 were not essential for flagellar function, in contrast to proposals that MotA-V178, MotA-A179 and MotA-A180 residues might control BFM function and ion selectivity (Biquet-Bisquert et al., 2021). None of the point mutations introduced into our non-motile MotA-ASRs were able to restore motility, which could imply these residues were not solely responsible for disruption of motility. However, it must be noted that some of these non-motile ASRs may not be well-folded and we did not test for this.

Thirdly, from the opposite perspective, residues which differed between functional MotA-ASRs and wild-type *E. coli* MotA provided information about variable residues which were not critical for function. This informed the level of general tolerance for significant amino acid substitution across MotA. The presence of multiple distinct (highly variable) residues in the cytoplasmic C-terminal region of MotA, particularly in residues 280–295, indicates that this region is interchangeable, which is similar to previous findings (Muramoto and Macnab, 1998; Santiveri et al., 2020). Although MotA-A187 and MotA-I202 residues are conserved in several proton-powered motors, our functional MotA-ASRs with T187 and V202 indicated that those residues were tolerant to amino acid substitution.

All our motile MotA-ASRs appeared to be proton-powered, that is, they showed no torque-dependence on sodium and were able to rotate in the absence of sodium. Furthermore, growth was inhibited when each of our motile MotA-ASRs was combined with a plug-deleted MotB subunit, indicating that proton leakage was occurring and providing further evidence for proton conductivity in our motile MotA-ASR stator units. It has been proposed that sodium is the ancestral power source for the BFM based on the observation of sodium-dependence in the *Aquifex aeolicus* stator unit (Takekawa et al., 2015). To attempt to answer this using ancestral reconstruction would require well dated phylogenies with appropriate accounting for gene transfer among the stator subunits. We do not address this here, however we note

that, to date, none of our ancestral reconstructions of either MotB (Islam et al., 2020) or MotA (this work) have generated a sodium-powered stator unit.

One strength of the ancestral reconstruction approach is that it leverages sequence diversity in contemporary stator proteins. Our results show that even quick, heuristic methods in phylogenetics and ancestral reconstruction can generate functional stator units. Inferred nodes are only estimates, in some cases coarse ones, and may not correctly represent evolutionary history. Improved phylogenetic inferences may improve evolutionary accuracy and the chance of resurrected ancestral stator units being functional. Advanced amino acid substitution models that are selected for higher statistical fit (Minh et al., 2020) should result in phylogenies that are more accurate more deeply in the tree, reducing problems such as long-branch attraction that result from model misspecification (Naser-Khdour et al., 2019; Crotty et al., 2020), and accordingly, more accurate ASRs. Ultimately ASRs may generate non-functional proteins for several reasons, such as failure to fold or even express (Eick et al., 2017). Nevertheless, our results based on functional MotA-ASRs prove successful protein folding through positive function and highlight the role of important regions of MotA that are both conserved and essential for function.

Combinatorial testing of MotA-ASRs from this study with ASRs from our previous MotB-ASR work (Islam et al., 2020) showed that certain MotA-ASRs function only in combination with certain MotB-ASRs. This indicates that there is specificity between MotA and MotB proteins. The promiscuity of stator units (Ito and Takahashi, 2017) and the modular nature of the stator complex (Ridone et al., 2022) may mean that there may be combinations of A and B subunits which are able to power rotation and combinations which cannot. It is therefore suitable to resurrect MotBs that co-existed with MotAs. For this, gene trees of *motA* and *motB* would need to be matched, which should be possible since they are usually syntenic.

Overall, in this work we showcased the utility of ASR for synthesising a range of functional stator units. This delivered insight into the function and evolutionary constraints surrounding an ancient, canonical molecular complex. Future work will examine co-evolutionary history *motA* and *motB* and phylogenies and reconstructions of adjacent genes. This will provide an estimate of the ancestral stator complex while reducing possible complications that may arise from gene transfer.

Materials and Methods

Phylogenetic analysis

MotA genes were obtained from select groups of bacterial species and aligned using MAFFT (Katoh et al., 2019) under default settings (MSAs provided in Supplementary Material). Two phylogenies were calculated: the first phylogeny of 178 *motA* homologues was estimated using QuickTree, a neighbour-joining method suitable for large datasets (Howe et al., 2002), with the Kimura translation for pairwise distances. This phylogeny was midpoint-rooted in FigTree (Rambaut, 2014). The second phylogeny was generated using the RAxML-HPC v.8 on XSEDE (Stamatakis, 2014) tool through the CIPRES Science Gateway (Miller et al., 2011) using the PROTGAMMA protein substitution model, LG protein substitution matrix, and a parsimony seed value of 12345. Newick files with full Linnean names and Accession IDs are provided in Supplementary Material.

Ancestral sequence reconstruction

Both multiple sequence alignments and both Newick files for the two phylogenies in Fig. 1 were used as inputs for CodeML, a maximum likelihood program from the PAML package (Yang, 1997). PAML was run with an empirical model for the amino acid substitution rate using LG as the amino acid ratefile, matching the model used for the RAxML tree generation. As an additional search for functional ancestors, we tested a contemporary ASR method, GRASP (Foley et al., 2022). The 178 QuickTree-generated motA tree (Fig. 1A) was submitted to the GRASP reconstruction webserver with the LG model, and the most probable characters at each site, for each node, was exported for further consideration. Of these reconstructions, two were selected for further synthesis: N41, and N65, corresponding to the nodes 220 and 244 using the labelling system in PAML. All nodes that selected and the ancestral reconstructions that were subsequently synthesised for further testing are indicated in Fig. 1.

Structural prediction of selected MotA-ASRs

We used ColabFold: AlphaFold2_mmseqs2 notebook version v1.1 (Mirdita et al., 2021) and Alphafold 2.1.1 (Jumper et al., 2021) to predict the 3D structures of our MotA-ASRs and contemporary *E. coli* (RP437) MotAB ($MotA_5MotB_2$) stator complex, respectively. Monomeric structures were generated for individual MotA-ASRs using one copy of ancestral MotA sequence in the default setting. In contrast, an oligomeric structure was generated for the contemporary *E. coli* MotAB complex using five copies of MotA and two copies of MotB (as an oligomeric 5:2 structure) using Alphafold v2.1.1 in multimer mode with the

parameters: max template date = 24-1-2022 and prokaryotic_list = true. The resultant structures were aligned with the solved unplugged MotAB stator complex structure of *C. jejuni*: PDB-6YKP (Santiveri et al., 2020) in Pymol (version 2.5.1) for comparison. Monomeric structures were generated for individual MotA-ASRs using one copy of ancestral MotA in the default setting. These monomers were then aligned to individual chains of 6YKP (Supplementary Figure 4).

Bacterial strains and plasmids and growth conditions

The bacterial strains and plasmids used in this study are shown in full in Supplementary Table 4. The primary strains used in this work are Δ*motA* RP437 (this study) and Δ*motA*-Δ*motB* RP437 (this study). All the *E. coli* strains were cultured in LB broth and LB agar [1% (w/v) Bacto tryptone, 0.5% (w/v) Bacto yeast extract, 0.5% (w/v) NaCl and 2% (w/v) Bacto agar for solid media] at 37°C. According to the selective antibiotic resistance pattern of the plasmids, chloramphenicol (CAM), ampicillin (AMP) or kanamycin (KAN) were added to a final concentration of 25 µg/mL, 100 µg/mL and 25 µg/mL, respectively.

Cloning of selected MotA-ASR sequences

Selected MotA-ASR sequences were ordered as gBlocks from Integrated DNA Technologies (IDT) for cloning into pBAD33 based chimeric plasmid pSHU1234 (Nishino et al., 2015) carrying *pomA* and *potB* (a hybrid of *pomB* and *motB*). The cloning was performed using Circular polymerase extension cloning (CPEC) protocol (Quan and Tian, 2011) with slight modification. Briefly, primers were designed to prepare linearized vector and MotA-ASR inserts that contain overlapping sequences (15-35 bases) between the vector and the inserts. IDT synthesized the designed primers, and the list of all primers is provided in Supplementary Table 5. MotA-ASR inserts and linear vector were prepared by PCR amplification using Q5 high-fidelity (HF) DNA polymerase (NEB). The PCR reaction recipe and condition is provided to supplementary information. The amplified MotA-ASR insert, and linear vector sequences were then separated in a 1% agarose gel and purified using the Qiagen gel purification kit. Finally, the CPEC cloning assembly reaction was performed with the purified MotA-ASR insert and linearized vector using In-Fusion Snap Assembly master mix (Takara) following the manufacturer's protocol. All the cloned MotA-ASR plasmids were confirmed by colony PCR and sanger sequencing at The Ramaciotti Centre for Genomics (UNSW, Australia) using ASR sequence-specific primers (Supplementary Table 5).

Δ*motA* strain preparation Cas9-assisted Recombineering

Deletion of *motA* from *E. coli* RP437 was performed using a no-SCAR method, adapted from the previous report (Reisch and Prather, 2015). Briefly, the target strain to be edited (*E. coli RP437*) was sequentially transformed first with the pKD-sgRNA-3'MotA (Sm^+) plasmid, encoding a sgRNA sequence directed at the 3' end of motA, and then with the pCas9cr4 (Cm^+) plasmid to yield a parent strain harboring both plasmids. A slightly modified overlap extension PCR technique (Higuchi et al., 1988) was employed to assemble linear double stranded DNA molecules (dsDNA) using 3 starting dsDNA fragments. The resulting donor DNA was electroporated in the plasmid-bearing host using Gene Pulser Xcell Electroporation Systems (Bio-Rad) and the successfully edited clones were selected via colony PCR and Sanger sequencing of the motAB locus. The list of used primers is provided in Supplementary Table 5.

Evaluation of functional property of MotA-ASRs

We transformed all the MotA-ASR plasmids into Δ*motA* RP437 and Δ*mot*-Δ*motB* RP437 *E. coli* strains following the chemical transformation protocol from NEB for functional characterization. We evaluated the functional property of resurrected MotA-ASRs using swim plate motility assay and free-swimming assay following the previous protocol (Morimoto et al., 2017; Islam et al., 2020). Briefly, for swim plate motility assay, we inoculated fresh single colonies of ten MotA-ASR transformed Δ*motA* and Δ*mot*-Δ*motB E. coli* on (0.02%) arabinose and chloramphenicol (25 μg/ml) containing semi solid LB swim agar (0.25%) with a sterile toothpick and incubated at 30□ for 14h to allow proper development of a swimming ring. Swimming zones were visually checked, imaged with the Chemi Doc MP Imaging System (Bio-Rad).

For free swimming assay, the overnight culture of the MotA-ASR transformed Δ*motA E. coli* strains were subcultured with a 50-fold dilution into fresh TB broth and incubated for 5 h with 180 rpm at 30°C. At OD_{600} ~0.80, the cells were washed once with motility buffer (10 mM potassium-phosphate, 10 mM lactic acid, 100 mM NaCl, and 0.1 mM EDTA, pH 7.0). Then, cell suspensions were added to the tunnel slide and the swimming speed of the cells was observed using phase-contrast microscopy and a 20 s video was recorded at 20 frames per second (fps) through a 20x objective (Nikon) with a camera (Chameleon3 CM3, Point

Grey Research). Finally, the swimming speed of the cells in was calculated using LabVIEW 2019 software (National Instruments).

Determination and confirmation of the power source of MotA-ASRs

We used a minimal swim plate in low Na^+ condition (~1 mM) without tryptone or yeast extract and where NaCl had been substituted with KCl (Islam et al., 2020; Ridone et al., 2022), to determine which one between sodium (Na^+) and proton (H^+) was the power source of our MotA-ASRs. The minimal swim plates were inoculated with a fresh single colony of MotA-ASR transformed Δ*motA E. coli* with a sterile toothpick and incubated at 30□ for 16h. After incubation, we checked whether the tested MotA-ASRs could produce any swimming ring or not in the absence of Na^+ and compared the phenotype with the control Na^+ (*pomApotB*) and H^+ (*motAmotB*) swimmers.

We further used the tethered cell assay to confirm the H^+ dependent phenotype of the previous minimal swim plate assay following a modified version of earlier protocol (Nishiyama and Kojima, 2012). Briefly, the overnight culture of the MotA-ASR transformed Δ*motA E. coli* strains were subcultured with a 50-fold dilution into fresh TB broth and incubated for 5 h with 180 rpm at 30°C. At OD_{600} ~0.80, the flagella of the cells were sheared by passing the culture multiple times (~35) through a 26G needle syringe. After shearing the flagella, the cells were washed three times with motility buffer. Cells were then attached on glass slides pre-treated with an anti-*E. coli* flagellin antibody with a 1:10 dilution (Nishiyama and Kojima, 2012) and washed sequentially with a Na^+ concentration gradient containing motility buffer. The rotational speed of the cells was observed using phase-contrast microscopy (Nikon) and was recorded at 20 frames per second (FPS) through the 40X objective with a camera (Chameleon3 CM3, Point Grey Research). Rotational speed of the cells was analysed and calculated using Lab view 2019 software (National Instruments) and rotational speeds were calculated from 20 individual cells.

Growth curve assays

Growth curves of MotA-ASRs were determined using a slightly modified protocol as described (Islam et al., 2020). Briefly, overnight cultures of the test strains with OD600 1.0 were diluted 1:100 in 96-well plates (Corning) with fresh LB broth, 0.02% (w/v) arabinose and 25 µg/ml chloramphenicol and incubated at 37□. The OD600 was measured every hour in a microplate reader (FLUOstar OPTIMA, BMB LABTECH) with a brief shaking interval before each measurement. The experiment was performed in a triplicate.

Point mutations in WT MotA, MotA-ASRs and plug deletion from MotB:

We performed the selected point mutation to test the roles of our identified conserved residues for the disruption or rescue of motility in *E. coli* and functional MotA-ASRs through *QuickChange Site-Directed Mutagenesis Kit* (Agilent). Initially, mutagenic oligonucleotide primers were designed in Agilent quick change primer design software at https://www.agilent.com/store/primerDesignProgram.jsp (Supplementary Table 5.). Then the mutant strands were synthesized following the Agilent recommended protocol by PCR amplification. Finally, the desired point mutants were confirmed by Sanger sequencing before functional analysis. We constructed plug deleted pMotB by deleting residues 51-70 from *motB* using Site-directed, Ligase-Independent Mutagenesis (SLIM), following previous protocol (Chiu et al., 2008).

Acknowledgements:

The authors acknowledge use of facilities in the Structural Biology Facility within the Mark Wainwright Analytical Centre– UNSW, funded in part by the Australian Research Council Linkage Infrastructure, Equipment and Facilities Grant: ARC LIEF 190100165. MABB acknowledges funding support from a Scientia Fellowship from UNSW, a CSIRO Future Science Platform in Synthetic Biology Project Grant, and Australian Research Council Discovery Project DP190100497 and Human Frontiers Science Program Project Grant RGY0072/21.

References:

Asai, Y., Kawagishi, I., Sockett, R. E., and Homma, M. (2000a). Coupling ion specificity of chimeras between H+- and Na+-driven motor proteins, MotB and PomB, in Vibrio polar flagella. *EMBO J.* doi: 10.1093/emboj/19.14.3639.

Asai, Y., Shoji, T., Kawagishi, I., and Homma, M. (2000b). Cysteine-Scanning Mutagenesis of the Periplasmic Loop Regions of PomA, a Putative Channel Component of the Sodium-Driven Flagellar Motor in Vibrio alginolyticus. *J. Bacteriol.* 182, 1001–1007. doi: 10.1128/JB.182.4.1001-1007.2000.

Beeby, M., Ferreira, J. L., Tripp, P., Albers, S.-V., and Mitchell, D. R. (2020). Propulsive nanomachines: the convergent evolution of archaella, flagella and cilia. *FEMS Microbiol. Rev.* 44, 253–304. doi: 10.1093/femsre/fuaa006.

Berg, H. C. (2003). *The rotary motor of bacterial flagella*. doi: 10.1146/annurev.biochem.72.121801.161737.

Biquet-Bisquert, A., Labesse, G., Pedaci, F., and Nord, A. L. (2021). The Dynamic Ion Motive Force Powering the Bacterial Flagellar Motor. *Front. Microbiol.* 12. Available at: https://www.frontiersin.org/article/10.3389/fmicb.2021.659464 [Accessed May 31, 2022].

Blair, D. F., and Berg, H. C. (1990). The MotA protein of E. coli is a proton-conducting component of the flagellar motor. *Cell* 60, 439–449. doi: 10.1016/0092-8674(90)90595-6.

Braun, T. F., Poulson, S., Gully, J. B., Empey, J. C., Van Way, S., Putnam, A., et al. (1999). Function of Proline Residues of MotA in Torque Generation by the Flagellar Motor of Escherichia coli. *J. Bacteriol.* 181, 3542–3551. doi: 10.1128/JB.181.11.3542-3551.1999.

Che, Y.-S., Nakamura, S., Kojima, S., Kami-ike, N., Namba, K., and Minamino, T. (2008). Suppressor analysis of the MotB(D33E) mutation to probe bacterial flagellar motor dynamics coupled with proton translocation. *J. Bacteriol.* 190, 6660–6667. doi: 10.1128/JB.00503-08.

Chiu, J., Tillett, D., Dawes, I. W., and March, P. E. (2008). Site-directed, Ligase-Independent Mutagenesis (SLIM) for highly efficient mutagenesis of plasmids greater than 8kb. *J. Microbiol. Methods.* doi: 10.1016/j.mimet.2008.02.013.

Crotty, S. M., Minh, B. Q., Bean, N. G., Holland, B. R., Tuke, J., Jermiin, L. S., et al. (2020). GHOST: Recovering Historical Signal from Heterotachously Evolved Sequence Alignments. *Syst. Biol.* 69, 249–264. doi: 10.1093/sysbio/syz051.

Deme, J. C., Johnson, S., Vickery, O., Muellbauer, A., Monkhouse, H., Griffiths, T., et al. (2020). Structures of the stator complex that drives rotation of the bacterial flagellum. *Nat. Microbiol.* doi: 10.1038/s41564-020-0788-8.

Dube, N., Khan, S. H., and Okafor, C. D. (2022). Ancestral sequence reconstruction for evolutionary characterization of proteins. *Trends Biochem. Sci.* 47, 98–99. doi: 10.1016/j.tibs.2021.10.009.

Eick, G. N., Bridgham, J. T., Anderson, D. P., Harms, M. J., and Thornton, J. W. (2017). Robustness of Reconstructed Ancestral Protein Functions to Statistical Uncertainty. *Mol. Biol. Evol.* 34, 247–261. doi: 10.1093/molbev/msw223.

Foley, G., Mora, A., Ross, C. M., Bottoms, S., Sützl, L., Lamprecht, M. L., et al. (2022). Engineering indel and substitution variants of diverse and ancient enzymes using Graphical Representation of Ancestral Sequence Predictions (GRASP). 2019.12.30.891457. doi: 10.1101/2019.12.30.891457.

Garcia, A. K., and Kaçar, B. (2019). How to resurrect ancestral proteins as proxies for ancient biogeochemistry. *Free Radic. Biol. Med.* 140, 260–269. doi: 10.1016/j.freeradbiomed.2019.03.033.

Garza, A. G., Bronstein, P. A., Valdez, P. A., Harris-Haller, L. W., and Manson, M. D. (1996). Extragenic suppression of motA missense mutations of Escherichia coli. *J. Bacteriol.* 178, 6116–6122. doi: 10.1128/jb.178.21.6116-6122.1996.

Higuchi, R., Krummel, B., and Saiki, R. K. (1988). A general method of in vitro preparation and specific mutagenesis of DNA fragments: study of protein and DNA interactions. *Nucleic Acids Res.* 16, 7351–7367. Available at: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC338413/ [Accessed August 5, 2022].

Hochberg, G. K. A., and Thornton, J. W. (2017). Reconstructing Ancient Proteins to Understand the Causes of Structure and Function. *Annu. Rev. Biophys.* 46, 247–269. doi: 10.1146/annurev-biophys-070816-033631.

Homma, M., Terashima, H., Koiwa, H., and Kojima, S. (2021). Putative Spanner Function of the Vibrio PomB Plug Region in the Stator Rotation Model for Flagellar Motor. *J. Bacteriol.* 203, e00159-21. doi: 10.1128/JB.00159-21.

Hosking, E. R., and Manson, M. D. (2008). Clusters of Charged Residues at the C Terminus of MotA and N Terminus of MotB Are Important for Function of the Escherichia coli Flagellar Motor. *J. Bacteriol.* 190, 5517–5521. doi: 10.1128/JB.00407-08.

Hosking, E. R., Vogt, C., Bakker, E. P., and Manson, M. D. (2006). The Escherichia coli MotAB Proton Channel Unplugged. *J. Mol. Biol.* 364, 921–937. doi: 10.1016/j.jmb.2006.09.035.

Howe, K., Bateman, A., and Durbin, R. (2002). QuickTree: Building huge neighbour-joining trees of protein sequences. *Bioinformatics*. doi: 10.1093/bioinformatics/18.11.1546.

Hu, H., Santiveri, M., Wadhwa, N., Berg, H. C., Erhardt, M., and Taylor, N. M. I. (2022). Structural basis of torque generation in the bi-directional bacterial flagellar motor. *Trends Biochem. Sci.* 47, 160–172. doi: 10.1016/j.tibs.2021.06.005.

Islam, M. I., Lin, A., Lai, Y.-W., Matzke, N. J., and Baker, M. A. B. (2020). Ancestral Sequence Reconstructions of MotB Are Proton-Motile and Require MotA for Motility. *Front. Microbiol.* 11. doi: 10.3389/fmicb.2020.625837.

Ito, M., and Takahashi, Y. (2017). *Nonconventional cation-coupled flagellar motors derived from the alkaliphilic Bacillus and Paenibacillus species*. doi: 10.1007/s00792-016-0886-y.

Ito, M., Terahara, N., Fujinami, S., and Krulwich, T. A. (2005). Properties of motility in Bacillus subtilis powered by the H +-coupled MotAB flagellar stator, Na+-coupled MotPS or hybrid stators MotAS or MotPB. *J. Mol. Biol.* doi: 10.1016/j.jmb.2005.07.030.

Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., et al. (2021). Highly accurate protein structure prediction with AlphaFold. *Nature* 596, 583–589. doi: 10.1038/s41586-021-03819-2.

Katoh, K., Rozewicki, J., and Yamada, K. D. (2019). MAFFT online service: multiple sequence alignment, interactive sequence choice and visualization. *Brief. Bioinform.* 20, 1160–1166. doi: 10.1093/bib/bbx108.

Kim, E. A., Price-Carter, M., Carlquist, W. C., and Blair, D. F. (2008). Membrane Segment Organization in the Stator Complex of the Flagellar Motor: Implications for Proton Flow and Proton-Induced Conformational Change. *Biochemistry* 47, 11332–11339. doi: 10.1021/bi801347a.

Kojima, S., and Blair, D. F. (2001). Conformational change in the stator of the bacterial flagellar motor. *Biochemistry* 40, 13041–13050. doi: 10.1021/bi011263o.

Kojima, S., Takao, M., Almira, G., Kawahara, I., Sakuma, M., Homma, M., et al. (2018). The Helix Rearrangement in the Periplasmic Domain of the Flagellar Stator B Subunit Activates Peptidoglycan Binding and Ion Influx. *Structure*. doi: 10.1016/j.str.2018.02.016.

Komatsu, H., Hayashi, F., Sasa, M., Shikata, K., Yamaguchi, S., Namba, K., et al. (2016). Genetic analysis of revertants isolated from the rod-fragile fliF mutant of Salmonella. *Biophys. Physicobiology* 13, 13–25. doi: 10.2142/biophysico.13.0_13.

Miller, M. A., Pfeiffer, W., and Schwartz, T. (2011). The CIPRES science gateway: a community resource for phylogenetic analyses. in *Proceedings of the 2011 TeraGrid Conference on Extreme Digital Discovery - TG '11* (Salt Lake City, Utah: ACM Press), 1. doi: 10.1145/2016741.2016785.

Minamino, T., and Imada, K. (2015). *The bacterial flagellar motor and its structural diversity*. doi: 10.1016/j.tim.2014.12.011.

Minh, B. Q., Schmidt, H. A., Chernomor, O., Schrempf, D., Woodhams, M. D., von Haeseler, A., et al. (2020). IQ-TREE 2: New Models and Efficient Methods for Phylogenetic Inference in the Genomic Era. *Mol. Biol. Evol.* 37, 1530–1534. doi: 10.1093/molbev/msaa015.

Mirdita, M., Ovchinnikov, S., and Steinegger, M. (2021). ColabFold - Making protein folding accessible to all. 2021.08.15.456425. doi: 10.1101/2021.08.15.456425.

Mirdita, M., Schütze, K., Moriwaki, Y., Heo, L., Ovchinnikov, S., and Steinegger, M. (2022). ColabFold: making protein folding accessible to all. *Nat. Methods* 19, 679–682. doi: 10.1038/s41592-022-01488-1.

Morimoto, Y. V., Namba, K., and Minamino, T. (2017). Measurements of Free-swimming Speed of Motile Salmonella Cells in Liquid Media. *Bio-Protoc.* 7, e2093–e2093. Available at: https://bio-protocol.org/e2093 [Accessed July 12, 2021].

Muramoto, K., and Macnab, R. M. (1998). Deletion analysis of MotA and MotB, components of the force-generating unit in the flagellar motor of Salmonella. *Mol. Microbiol.* 29, 1191–1202. doi: 10.1046/j.1365-2958.1998.00998.x.

Nakamura, S., and Minamino, T. (2019). Flagella-Driven Motility of Bacteria. *Biomolecules* 9, 279. doi: 10.3390/biom9070279.

Nakamura, S., Morimoto, Y. V., Kami-ike, N., Minamino, T., and Namba, K. (2009). Role of a Conserved Prolyl Residue (Pro173) of MotA in the Mechanochemical Reaction Cycle of the Proton-Driven Flagellar Motor of Salmonella. *J. Mol. Biol.* 393, 300–307. doi: 10.1016/j.jmb.2009.08.022.

Naser-Khdour, S., Minh, B. Q., Zhang, W., Stone, E. A., and Lanfear, R. (2019). The Prevalence and Impact of Model Violations in Phylogenetic Analysis. *Genome Biol. Evol.* 11, 3341–3352. doi: 10.1093/gbe/evz193.

Nishino, Y., Onoue, Y., Kojima, S., and Homma, M. (2015). Functional chimeras of flagellar stator proteins between E. coli MotB and Vibrio PomB at the periplasmic region in Vibrio or E. coli. *MicrobiologyOpen*. doi: 10.1002/mbo3.240.

Nishiyama, M., and Kojima, S. (2012). Bacterial motility measured by a miniature chamber for high-pressure microscopy. *Int. J. Mol. Sci.* doi: 10.3390/ijms13079225.

Quan, J., and Tian, J. (2011). Circular polymerase extension cloning for high-throughput cloning of complex and combinatorial DNA libraries. *Nat. Protoc.* 6, 242–251. doi: 10.1038/nprot.2010.181.

Rambaut, A. (2014). FigTree v1.4.2, A Graphical Viewer of Phylogenetic Trees. *Available 〈httptreebioedacuksoftwarefigtree〉*.

Reisch, C. R., and Prather, K. L. J. (2015). The no-SCAR (Scarless Cas9 Assisted Recombineering) system for genome editing in Escherichia coli. *Sci. Rep.* 5, 15096. doi: 10.1038/srep15096.

Ridone, P., Ishida, T., Lin, A., Humphreys, D. T., Giannoulatou, E., Sowa, Y., et al. (2022). The rapid evolution of flagellar ion selectivity in experimental populations of E. coli. *Sci. Adv.* 8, eabq2492. doi: 10.1126/sciadv.abq2492.

Santiveri, M., Roa-Eguiara, A., Kühne, C., Wadhwa, N., Hu, H., Berg, H. C., et al. (2020). Structure and Function of Stator Units of the Bacterial Flagellar Motor. *Cell.* doi: 10.1016/j.cell.2020.08.016.

Sharp, L. L., Zhou, J., and Blair, D. F. (1995). Features of MotA proton channel structure revealed by tryptophan-scanning mutagenesis. *Proc. Natl. Acad. Sci. U. S. A.* 92, 7946–7950. doi: 10.1073/pnas.92.17.7946.

Sowa, Y., and Berry, R. M. (2008). Bacterial flagellar motor. *Q. Rev. Biophys.* 41, 103–132. doi: 10.1017/S0033583508004691.

Stamatakis, A. (2014). RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 30, 1312–1313. doi: 10.1093/bioinformatics/btu033.

Takekawa, N., Kojima, S., and Homma, M. (2014). Contribution of Many Charged Residues at the Stator-Rotor Interface of the Na+-Driven Flagellar Motor to Torque Generation in Vibrio alginolyticus. *J. Bacteriol.* 196, 1377–1385. doi: 10.1128/JB.01392-13.

Takekawa, N., Nishiyama, M., Kaneseki, T., Kanai, T., Atomi, H., Kojima, S., et al. (2015). Sodium-driven energy conversion for flagellar rotation of the earliest divergent hyperthermophilic bacterium. *Sci. Rep.* doi: 10.1038/srep12711.

Terashima, H., Kojima, S., and Homma, M. (2021). Site-Directed Cross-Linking Identifies the Stator-Rotor Interaction Surfaces in a Hybrid Bacterial Flagellar Motor. *J. Bacteriol.* 203, e00016-21. doi: 10.1128/JB.00016-21.

Togashi, F., Yamaguchi, S., Kihara, M., Aizawa, S. I., and Macnab, R. M. (1997). An extreme clockwise switch bias mutation in fliG of Salmonella typhimurium and its suppression by slow-motile mutations in motA and motB. *J. Bacteriol.* 179, 2994–3003. doi: 10.1128/jb.179.9.2994-3003.1997.

Yang, Z. (1997). PAML: a program package for phylogenetic analysis by maximum likelihood. *Bioinformatics* 13, 555–556. doi: 10.1093/bioinformatics/13.5.555.

Yorimitsu, T., and Homma, M. (2001). Na+-driven flagellar motor of Vibrio. *Biochim. Biophys. Acta BBA - Bioenerg.* 1505, 82–93. doi: 10.1016/S0005-2728(00)00279-6.

Zhou, J., and Blair, D. F. (1997). Residues of the cytoplasmic domain of MotA essential for torque generation in the bacterial flagellar motor11Edited by M. Gottesman. *J. Mol. Biol.* 273, 428–439. doi: 10.1006/jmbi.1997.1316.

Zhou, J., Lloyd, S. A., and Blair, D. F. (1998a). Electrostatic interactions between rotor and stator in the bacterial flagellar motor. *Proc. Natl. Acad. Sci.* 95, 6436–6441. doi: 10.1073/pnas.95.11.6436.

Zhou, J., Sharp, L. L., Tang, H. L., Lloyd, S. A., Billings, S., Braun, T. F., et al. (1998b). Function of protonatable residues in the flagellar motor of Escherichia coli: A critical role for Asp 32 of MotB. *J. Bacteriol.* doi: 10.1128/jb.180.10.2729-2735.1998.

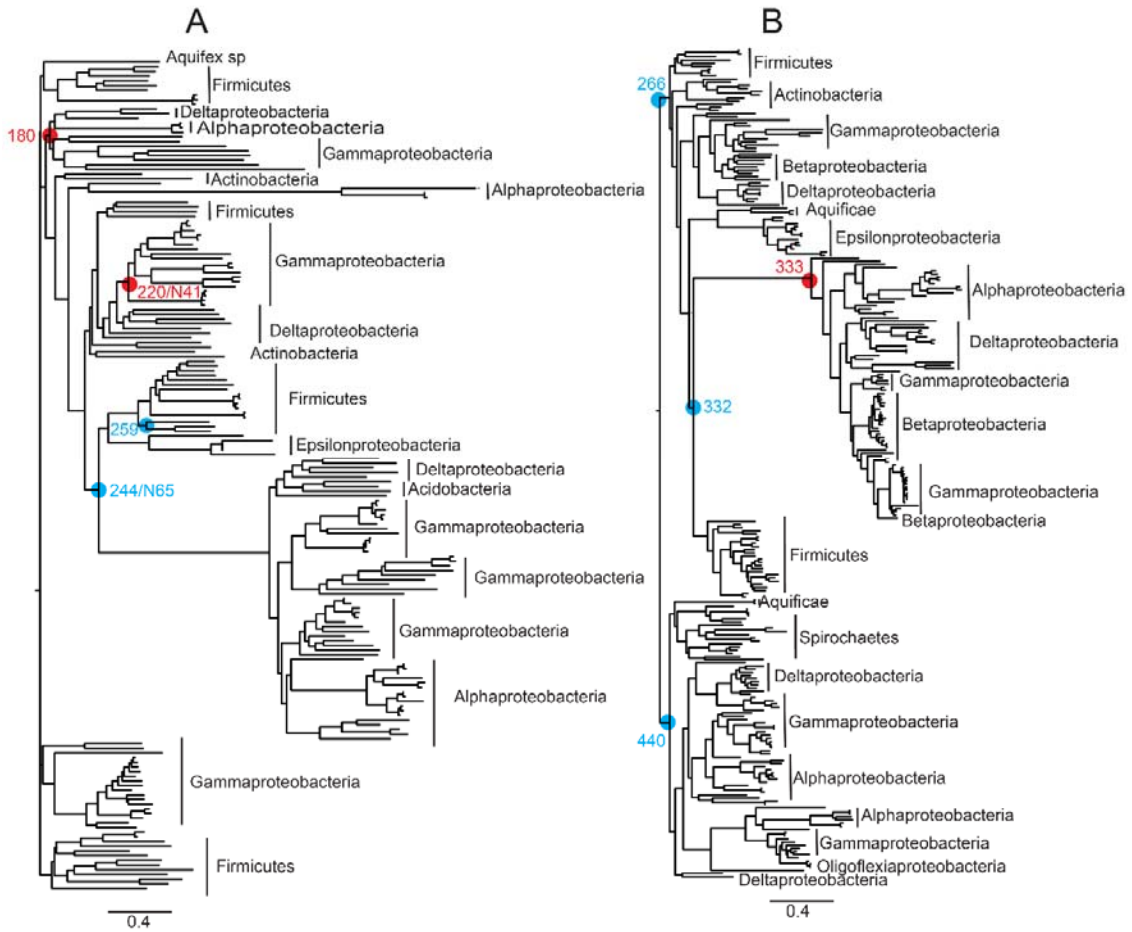

Figure 1: Phylogenetic trees of 178 and 264 *motA* homologs. Ten nodes across both phylogenies were selected for synthesis (red circles: motile MotA-ASRs; blue circles: non-motile MotA-ASRs). (A) Neighbour-joining phylogeny generated using QuickTree (Howe et al., 2002) for 178 *motA* sequences. (B) Maximum-likelihood tree generated using RAxML (Stamatakis, 2014) for 264 *motA* sequences. Scale bar indicates substitutions per site. Newick files for the phylogenies, sequence alignments, and ASR quality measurements are available in the Supplementary Material.

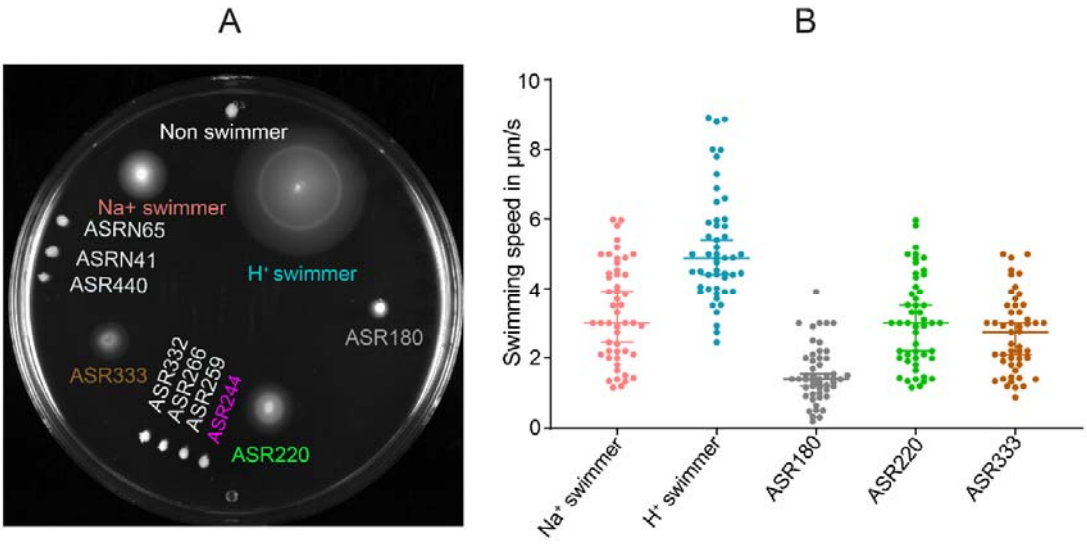

Figure 2: Functional properties of MotA-ASRs. (A) Swim plate assay (LB + 0.25% agar) of Mot-ASRs, H+ powered (*motAmotB*) and Na+ powered (*pomApotB*) control swimmers and a control non-swimmer (empty vector, pBAD33) on LB swim agar. After 14 h incubation at 30°C, ASR180, ASR220 and ASR333 produced swimming halos (coloured labels) while the other 7 MotA-ASRs did not produce swimming halos (white labels) (B) Swimming speeds of three functional MotA-ASRs (ASR180, ASR220 and ASR333 and the control H^+/Na^+ swimmers in a free-swimming assay. Colours in B match labels in A. The speed of individual cells was measured in µm/s and is shown with the scattered plot; median speeds are indicated by the horizontal line with a 95% confidence interval (n = 50 cells per ASR).

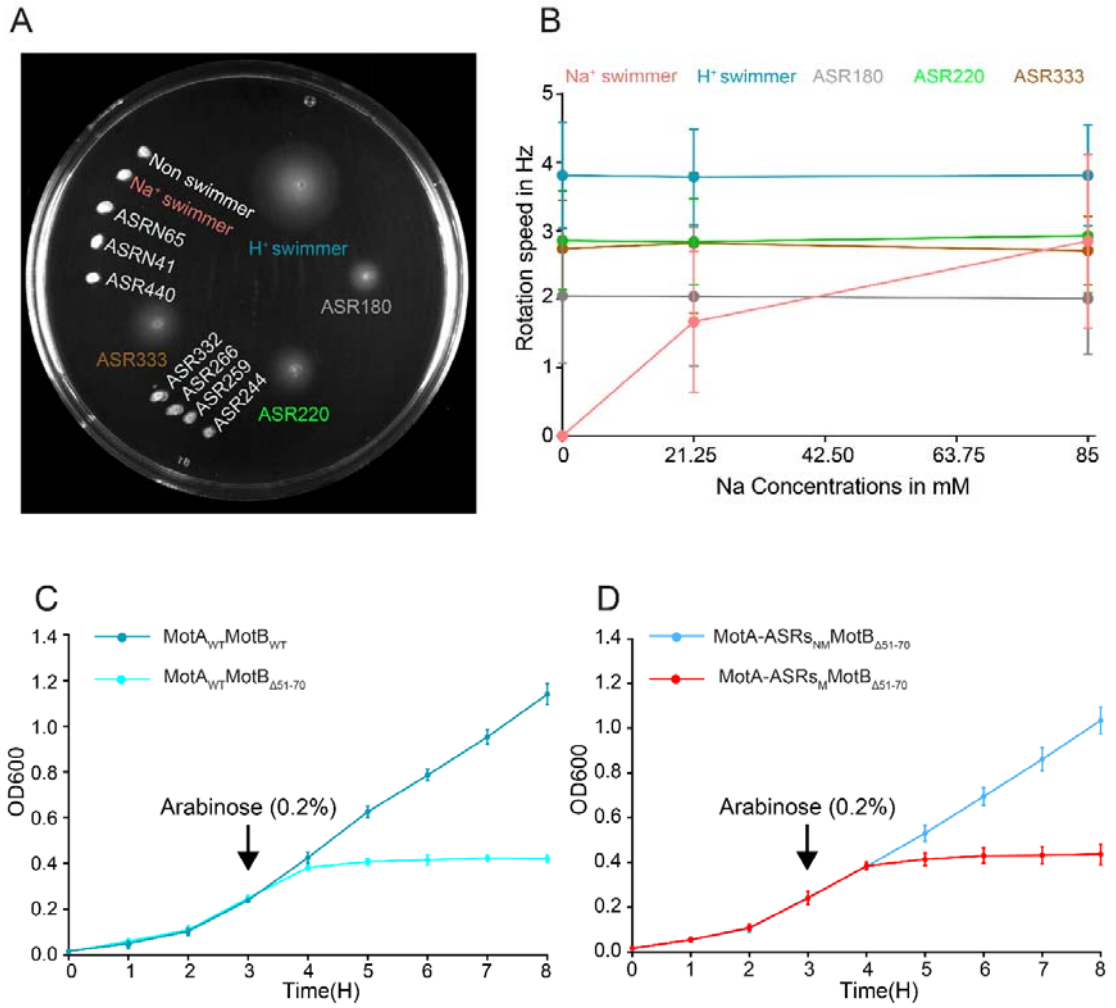

Figure 3: Characterization of the ion selectivity of MotA-ASRs (A) Minimal swim plate assay with KCl instead of NaCl (~1 mM [Na^+]) for all MotA-ASRs and the Na^+ and H^+ powered control. After 18 hours incubation at 30°C, control Na^+ swimmer (*pomApotB*) and all non-functional MotA-ASRs were non-motile. In contrast, control H^+ swimmer (*momAmotB*) and the tested functional MotA-ASRs (ASR180, ASR229, ASR333) were motile and produced swimming rings. (B) The rotational speed of MotA-ASRs and the Na^+ and H^+ powered controls were calculated in the presence of 0, 21.25, and 85 mM NaCl from the tethered cell assay (measured in Hz). The average rotational speed of all tested strains (mean ± SD, n = 50 cells per ASR) is shown in the line graph. Colours in B match labels in A. (C) Growth curves of wild type MotA ($MotA_{WT}$) in combination with wild type MotB ($MotB_{WT}$) and plug deleted MotB ($MotB_{\Delta 51\text{-}70}$). (D) Growth curves of motile MotA-ASRs ($MotA\text{-}ASRs_M$) and non-motile MotA-ASRs ($MotA\text{-}ASRs_{NM}$) in combination with plug deleted MotB ($MotB_{\Delta 51\text{-}70}$). The averaged growths (mean ± SD) of three motile (red line) and

six non-motile MotA-ASRs (blue line) were showed. Full data for each ASR is shown in Supplementary Figure 8.

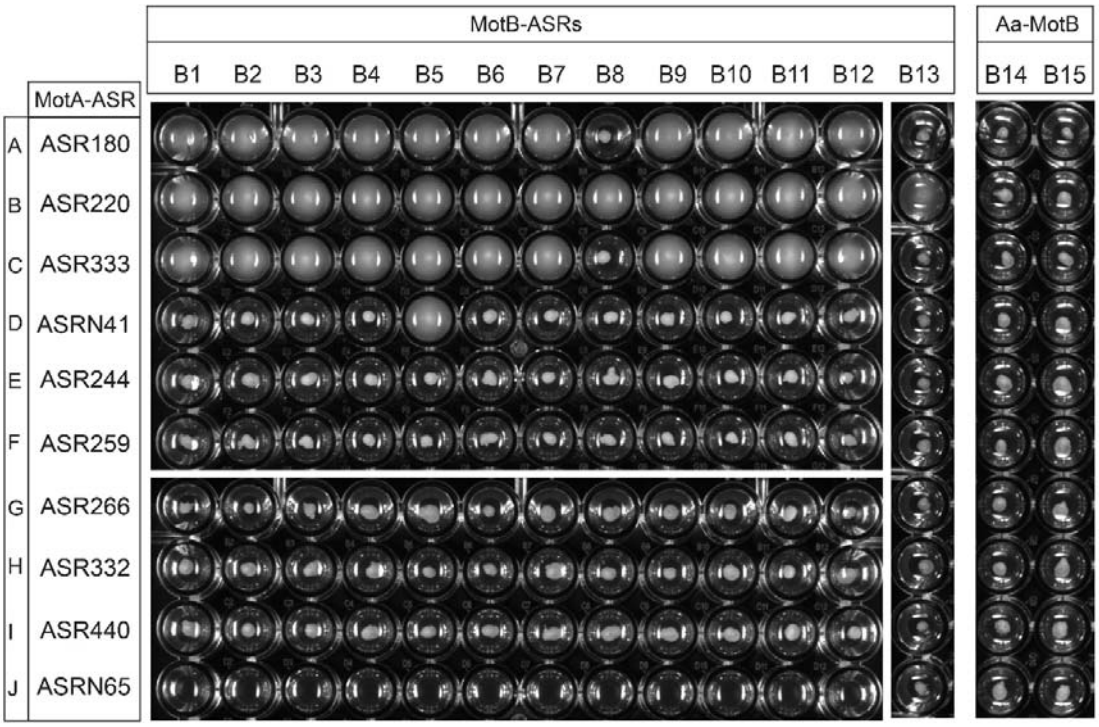

Figure 4: Compatibility between MotA-ASRs and MotB-ASRs. Swim plate assay in 96 well plate showing functional compatibility of MotA-ASRs with MotB-ASRs from (Islam et al., 2020). Each well contains 0.25% LB swim agar and the plate was incubated 14 hours at 30°C after inoculation of the test combination. A-J rows contain MotA-ASRs, and B1-B13 and B14-B15 columns contain tested MotB-ASRs and *Aquifex aeolicus* (Aa) MotBs, respectively. Row-A: MotA180; Row-B: MotA220; Row-C: MotA333; Row-D: MotAN41; Row-E: MotA244; Row-F: MotA259; Row-G: MotA266; Row-H: MotA332; Row-I: MotA440; and Row-J: MotAN65; Col-B1: MotB758; Col-B2: MotB759; Col-B3: MotB760; Col-B4: MotB765; Col-B5: MotB908; Col-B6: MotB981; Col-B7: MotB1024; Col-B8: MotB1170; Col-B9: MotB1239; Col-B10: MotB1246; Col-B11: MotB1457; Col-B12: MotB1459; Col-B13: MotB1501; Col-B14: Aa MotB1 and Col-B15: Aa MotB2.

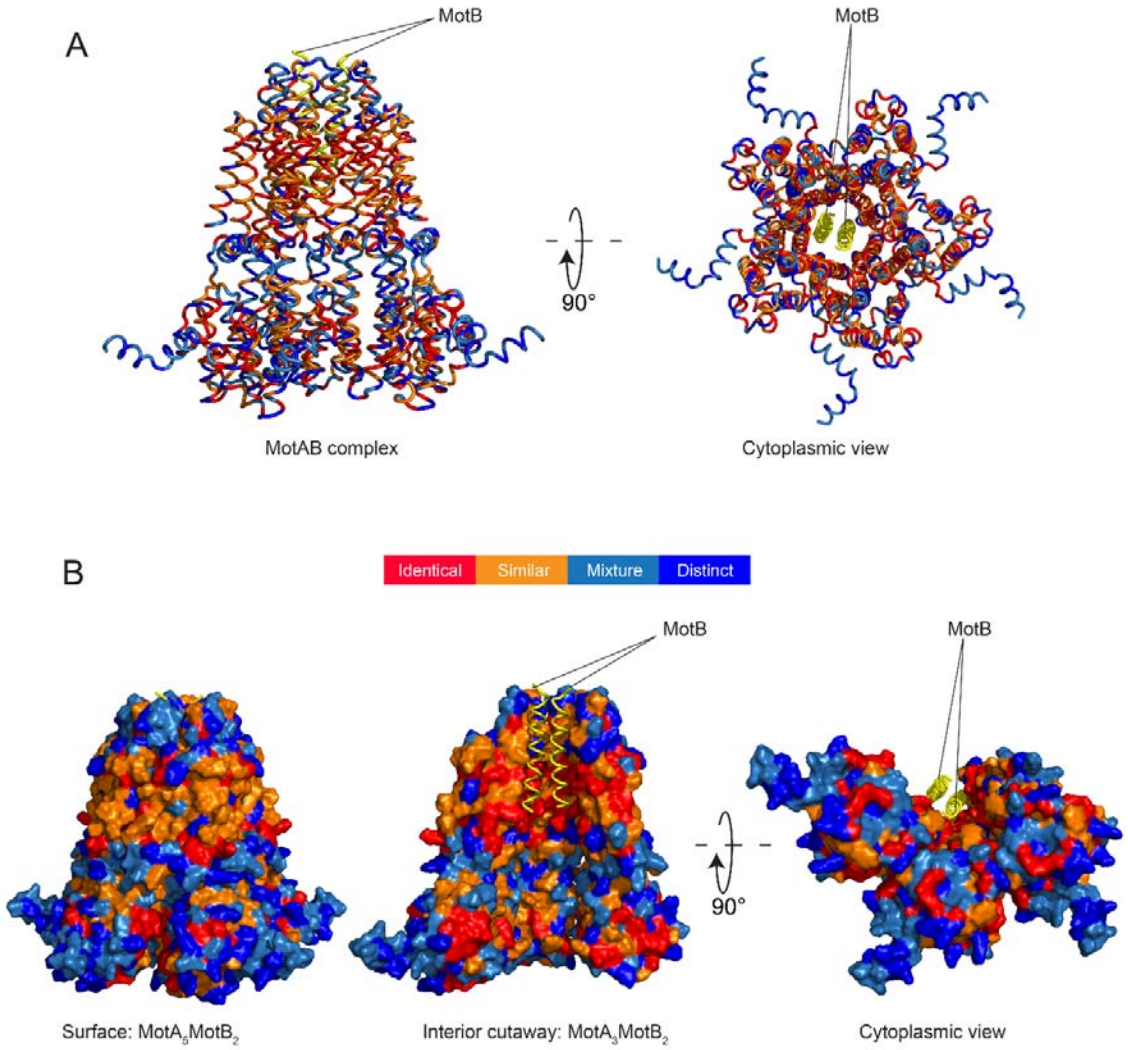

Figure 5: Residue conservation and variation between WT and reconstructed MotAs mapped onto an Alphafold model of *E. coli* $MotA_5MotB_2$. (A) Cartoon model shows the identical residues across the functional MotA-ASRs (red), similar (orange), mixed (light blue) and distinct (dark blue), visualised from the side (left) and from the cytoplasm (right). MotB is coloured yellow. (B) Surface model of MotAB complex (A_5B_2) with MotB is shown as cartoon. Side view showing the external surface conservation of MotA (Left), interior cutaway of the $MotA_3B_2$ complex showing the internal conservation (middle), and bottom cytoplasmic view of the interior cutaway of the $MotA_3B_2$ complex (right).

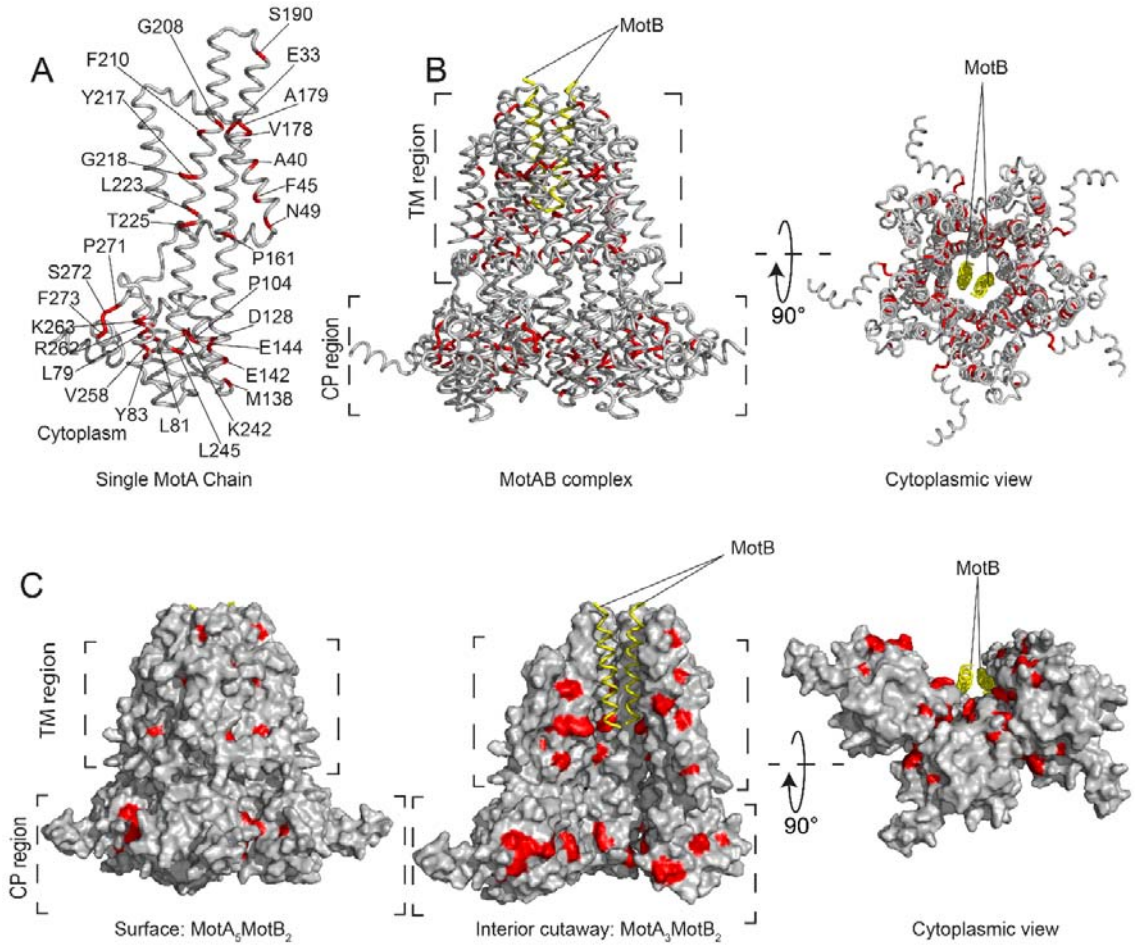

Figure 6: Revealed critical residues of MotA mapped onto an Alphafold model of *E. coli*.. (A) 30 critical residues from functional and non-functional MotA-ASRs are highlighted in red (conserved) with their identity on an isolated MotA chain. (B) The distribution and location of the 30 critical residues are showed on the complete $MotA_5MotB_2$ (side view and bottom cytoplasmic view). (C) Location of 30 residues labelled in (A) coloured in red on the surface and interior of the $MotA_5MotB_2$ complex (left) side view of the outside of the surface with critical residues, (centre), interior view with MotA trimer and MotB dimer and bottom cytoplasmic view (right).

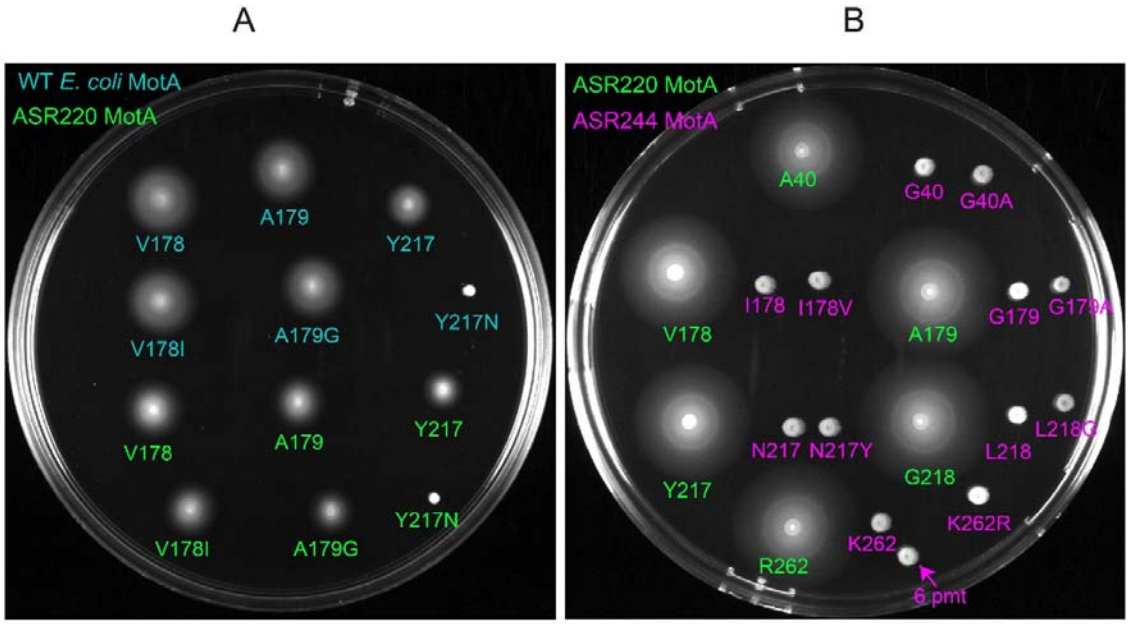

Figure 7: Role of the point mutations in disrupting or rescuing motility of MotA. (A) Swim plate assay (LB + 0.25% agar) comparing motility of WT *E. coli*-MotA (coloured in deep teal) and the functional MotA-ASR220 (coloured in green): V178, A179 and Y217 with their point mutants:V178I, A179G, and Y217N. (B) Swim plate assay (LB + 0.25% agar) comparing motility of functional MotA-ASR220 (coloured in green): A40, V178, A179, Y217, G218 and R262, non-functional MotA-ASR244 (coloured in magenta): G40, I178, G179, N217, L218 and K262 and point mutants of non-functional MotA-ASR244: G40A, I178V, G179A, N217Y, L218G and K262R. Magenta arrow indicates strain with all 6 point mutations (6 pmt).