Functional dynamics of the bacterial flagellar motor driven by fluorescent protein tagged stators and by evolutionary modified foreign stators

Minyoung Heo

Minyoung Heo. Functional dynamics of the bacterial flagellar motor driven by fluorescent protein tagged stators and by evolutionary modified foreign stators. Molecular biology. Université Montpellier, 2016. English. NNT : 2016MONTT080 . tel-01972611

THÈSE
Pour obtenir le grade de
Docteur

Délivré par **l'Université de Montpellier**

Préparée au sein de l'école doctorale Sciences Chimiques et Biologiques pour la Santé (ED168)

Et de l'unité de recherche Centre de Biochimie Structurale (CBS) - CNRS UMR 5048

Spécialité : Biophysique de la molécule unique, la microscopie à fluorescence et évolution expérimentale

Présentée par Minyoung HEO

Dynamique fonctionnelle du moteur flagellaire bactérien entraîné par des stators marqués par des protéines fluorescentes et par des stators étrangers modifiés par évolution

Soutenue le 25 Novembre 2016 devant le jury composé de

M. Francesco PEDACI, Directeur de recherche, Centre de biochimie Structurale (CNRS INSERM)	Directeur de thèse
M. Emmanuel MARGEAT, Directeur de recherche, Centre de biochimie Structurale (CNRS INSERM)	Directeur de thèse
M. Bertus BEAUMONT, Professeur des Universités, Delft University of Technology	Co-encadrant de thèse
M. Christophe DANELON, Directeur de recherche, Delft University of Technology	Rapporteur
M. Peter GALAJDA, Directeur de recherche, Hungarian Academy of Sciences	Rapporteur
M. Andrea PARMEGGIANI, Professeur des Universités, Université de Montpellier	Président du jury

FUNCTIONAL DYNAMICS OF THE BACTERIAL FLAGELLAR MOTOR DRIVEN BY FLUORESCENT PROTEIN TAGGED STATORS AND BY EVOLUTIONARY MODIFIED FOREIGN STATORS

Minyoung Heo

Thesis submitted for the degree of Doctor of Philosophy,
Sciences Chimiques et Biologiques pour la Santé (CBS2)
University of Montpellier, 2016

Abstract

The bacterial flagellar motor (BFM) is a macromolecular complex which allows bacteria to swim in liquid media. Located at the base of the flagellum, this remarkably small (~45nm) yet powerful rotary motor rotates each flagellum up to ~1000 revolutions per second in both counterclockwise (CCW) and clockwise (CW) directions. The motor rotation is generated at the interface between the two key components of the motor: the stator protein complexes (each composed of 4 MotA and 2 MotB proteins) and the C- ring protein complex at the base of the rotor. The stator complexes are structurally and functionally discernible modules of the motor, and their dynamical association and dissociation around the rotor controls the torque generation. Previously, dynamic nature of the stators has been demonstrated by the fluorescence detection of the stators fused to a fluorescent protein (FP). When a FP is fused to MotB, the motor is functional, but a reduced motility of the cells has been observed. The precise reasons for such reduction in motility is yet to be determined.

The first project aims to investigate how the FP tag on the stator protein modifies the torque generation and switching of the motor. This is important because the fluorescent protein tag lies at the interface between stator and rotor, where torque and switching are produced. Three different FPs (eGFP, YPet, Dendra2) were fused to MotB. Interestingly, despite the high similarity of these three FPs' structures, our analysis revealed that the three fusion stators generate different torques by single stator. Furthermore, the motors driven by the fusion stators showed significantly impaired switching abilities. When switching direction of the rotation, the absolute value of the speed of WT motors does not change, whereas this symmetry of speed upon switching is not observed in the fusion stator motors and switching can be accompanied with a significant (~30%) decrease in absolute speed. Both the impaired torque generation and the switching ability were improved by introducing a rigid linker between the stator and the FP tag. Taken together, this study provides a further insight into the dynamics of the stator and rotor interaction at its interface.

When the cells carrying the fluorescently labeled stators were observed in a custom made TIRF-fluorescence microscope with single molecule capability, the fluorescence signals were detected as concentrated clusters in the membrane as expected for the stator proteins around the motors, together with a population of stators diffusing in the membrane. The fluorescent clusters were also observed at the center of rotating cells tethered to the glass slide by a single flagellum, confirming that the fluorescent spots observed were attributable to the functioning motors.

In a second project developed in Bertus Beaumont lab at TU Delft, taking BFM as an experimental evolutionary model system, its modularity and evolvability have been explored to learn the molecular details of the evolution of molecular machines. When the stators of *E.coli* (K-12) have been replaced by a set of 21 homologue foreign stators, some of the foreign stators were immediately compatible with the *E.coli* motor, while some of the non-compatible stators were positively modified by a chemotaxis evolution experiment. More than half of those evolved motors accumulated beneficial mutations in the functional domains of their foreign stator genes. Motilities of the evolved motors were investigated and compared at the level of population (chemotaxis), single cell (swimming) and single motor. This three-levels of functional investigation enabled detailed functional characterizations of the evolved or foreign motors. Especially, the single motor level assays revealed that those beneficial mutations improved the torque generation and/or the switching ability. The detailed genotype and phenotype investigations of the evolutionary modified BFM may bring an insight into how molecular machines such as BFM have evolved as well as the functional effects of the beneficial mutations that facilitate functional integration.

Résumé de la thèse en français

Le moteur flagellaire bactérien (BFM) est un complexe moléculaire qui permet aux bactéries de nager dans un milieu liquide. Situé à la base de la flagelle et fixé sur la membrane cellulaire, ce moteur rotatif remarquablement petit (~45nm) mais puissant entraîne la rotation de chaque flagelle à une fréquence pouvant atteindre 1000Hz, et en alternant son sens de rotation. La rotation du moteur est générée à l'interface entre deux éléments clés: les protéines formant le stator (MotA and MoB) et l'anneau C "switching complex" à la base du rotor. Les stators sont des modules du moteur structurellement et fonctionnellement différentiables du reste du moteur, et leurs association et dissociation dynamique autour du rotor contrôle la génération du couple. Puisque les protéines du stator génèrent le couple de rotation du moteur, il a été démontré qu'une protéine fluorescente fusionnée avec MotB permet d'observer et d'étudier la nature dynamique de cette protéine par une détection de fluorescence. Quand une protéine fluorescente (PF) est fusionnée à MotB, le moteur est en état de marche mais une réduction générale de la mobilité de la cellule a été observée. La raison précise d'une telle réduction de mobilité n'a pas été étudiée.

Le but de cette étude est de comprendre comment la fusion PF de la protéine du stator modifie la génération du couple et le sens de rotation du moteur. C'est particulièrement important car le tag FP se trouve à l'interface entre le stator et le rotor, là où le couple et le changement du sens de rotation sont produits. Trois différentes PFs (eGFP, YPet, Dendra2) ont été fusionnées à la protéine MotB. Malgré la haute similarité de leurs structures, notre analyse a montré que les trois stators fusionnés génèrent des couples différents. Les stators marqués avec YPet produisent un couple moyen similaire au WT (stators sans tag PF), alors que les stators marqués avec eGFP et Dendra2 produisent respectivement 70% et 40% du couple moyen du WT. De plus, les moteurs utilisant les stators fusionnés ont montré des capacités de changement de sens de rotation réduites. Lors d'un changement de sens de rotation, la valeur absolue de la vitesse des moteurs WT ne change pas. Cette "symétrie" de vitesse lors du changement n'apparaît pas avec les moteurs à stators fusionnés et le changement peut être accompagné d'une importante diminution (~30%) de la vitesse absolue.

En observant par microcopie TIRF avec détection de molécules uniques, des stators marqués dans un moteur en état de marche, les signaux de fluorescence sont détectés à la membrane comme prévu pour ces protéines, montrant une population de stators diffusant dans celle-ci. Les clusters fluorescents étaient visibles au centre des cellules en rotation, attachés au couvre-glace par une seule flagelle, confirmant que le tag de fluorescence peut être visualisé dans des moteurs en état de marche.

Dans un second projet développé dans le laboratoire Bertus Beaumont à TU Delft, en prenant le BFM en tant que système modèle d'évolution expérimentale, sa modularité et son « évolubilité » ont été explorés pour apprendre les détails au niveau moléculaire de l'évolution de ce type de machine. Les stators de *E.coli* ont été échangés par un set de 21 stators étrangers homologues. L'expérience a révélé que les protéines du stator peuvent être échangées entre espèces de bactéries distantes et certains stators non compatibles peuvent être modifiés positivement par un procédé d'évolution pour devenir fonctionnels. Au cours de cette évolution, les bactéries ont accumulé des mutations avantageuses dans leurs gènes motA et motB étrangers, tout particulièrement dans leur domaine fonctionnel. Des mutations identiques dans des stators différents ont été observées, indiquant que l'évolution peut se reproduire. L'analyse fonctionnelle au niveau d'un seul moteur a révélé que ces mutations avantageuses amélioraient la génération du couple et/ou la capacité du moteur à changer de sens. Les investigations

détaillées du génotype et du phénotype du BFM modifié par évolution apportées par cette étude, pourraient donner une idée sur la façon dont des machines moléculaires comme le BFM ont évolué, et les effets fonctionnels des mutations bénéfiques qui facilitent l'intégration fonctionnelle.

Acknowledgement

I would like to thank my supervisor Dr. Francesco Pedaci for his supports and advices during my PhD study years in Montpellier. Thank you for introducing me what biophysics is. It was a great opportunity for me to expand my expertise in science, especially in quantitative data analysis. I am grateful to my co-supervisor Dr. Bertus Beaumont for his advices in many aspects of scientific research, such as paying attention to details and more rigorous sciences. I really enjoyed learning about experimental evolution as well. I was lucky to have to two supervisors in two domains of science: physics and biology. It would not have been possible to complete this highly interdisciplinary study without them.

Special thanks to Dr. Emmanuel Margeat for his advises, help on the French abstract and for being an essential member on each thesis committee meeting. I would like to thank Dr. Andrea Parmeggiani for his encouragement as well as Dr. Chistopher Daleon and Dr. Peter Galajda for being jury members of my thesis. I would also like to thank Dr. Labesse Gilles for helpful scientific discussions.

I would also like to thank many members of the lab, Regis Flohr introduced me the project he was working on, Erwin van Rijn, Carten Blom, Bert Hubert, and so many members of the Bionanoscience department of TUDelft, Dr. Ashley Nord for her various helps on fluorescence microscope set up and for sharing her knowledge on flagellar motors and experiments throughout the whole years, Dr. Zhanna Santybayeva and Dr. Luca Lanotte for making the office a fun and stimulating place, lastly all the members of CBS in Montpellier for a helpful and friendly working environment. I have learned so much from each one of them.

Table of Contents

Chapter 1. Molecular Motors and the Bacterial Flagellar Motor

1.1 Molecular Motors

1.1.1 Introduction to the life in motion

Growth, sensing and reproduction of cellular life require certain degree of dynamical physical movements in both inside and outside of the cell. Likewise, one of the main features that distinguish cellular life to inorganic matters is their capacity to move autonomously. Having a power to move spontaneously is particularly a vital characteristic of animal life, and arguably, self-movement could have already presented at the origin of life [1]. Biophysics is interdisciplinary science that studies the complexity and functionality of life using mathematical and physical tools. Biophysics look for the patterns in life, analyze them with math and physics [2]. Hence, the motility in cellular life, which embraces both simplicity of physical laws and complexity of cellular life, is one of the main subjects in biophysics.

There are several biological macromolecules known as molecular motors or biological nano-machines. As the name suggests, mechanical movements are essential for their performance. The mechanical action of these motors is driven by chemical energy, such as ATP hydrolysis or ion translocation, which result in conformational changes of the molecular motors [3-6]. Understanding how the conformational change results in the mechanical action of the motor is essential to understand their complete mechanism. The molecular motors play various roles in cellular life, such as genetic information processing, molecules transporting and rotating an external flagellum. Eukaryotic cellular life has three types of motors: 1) nucleic acid DNA machines (replisome) [7-12], 2) cytoskeletal motors- kinesin, myosin and dynein [13-17], and 3) rotary motors – ATPases [18-21]. Prokaryotes also contain those motors but in simpler versions with respect to the eukaryotic motors. In addition, prokaryotes have another rotary motor that facilitates their unicellular movement, which is called the bacterial flagellar motor [22-25]. General functions of these eukaryotic and prokaryotic motors will be introduced in the following sections of this chapter, and the bacterial flagellar motor will be discussed in details in chapter 1.2 as this is the molecular motor studied in this thesis.

Lastly, why study motor proteins? First reason is to expand our existing knowledge in biology, as these motors play essential roles in cellular processes. With this, secondly, we can apply for two possible practical outcomes: 1) engineering nano-motor [17, 26, 27] as understanding the design principles of molecular motors may guide us to construct efficient nanoscale machine, and 2) in relevance to medicine, these motor proteins can be unique therapeutic targets as inhibition or enhancement of their activity may cause various therapeutic benefits. [6, 12].

1.1.2 Nucleic acid motors and linear motors

Nucleic acid motor proteins refer to all the molecular motors involved in synthesis or replication of DNA (deoxyribonucleic acid) and RNA (ribonucleic acid) polymers. The most well-known nucleic acid motor is the DNA replicating enzyme- DNA polymerases (DNAP). There are more than 30 different types of DNA polymerases discovered, some of which have distinct structures of the catalytic subunits [7-8]. The main function of the DNA polymerases is in DNA repair and replication, but they also play diverse roles in cell-cycle check points, homologous recombination and development of the immune system [8]. The DNAP works together with a primase, which is a special type of polymerase that initiates polymerization of a short RNA primer by identifying a specific sequence [4,11]. Without this primase, DNAP cannot initiate DNA replication. Together with the DNAPs and primases, there are helicases and topoisomerases in eukaryotic replisome. The replisome is a multiprotein DNA replication machinery, loosely analogous to a sewing machine [9, 30-31] (Figure 1.1a). Helicase unwinds a double strained (ds) DNA into single strained (ss) DNA so that the DNAP can replicate the ssDNA [9-12]. Topoisomerases are responsible for untangling super-coiled and interlinked DNA structures during replication (decatenation of DNA) [4, 28].

Another DNA template directed nucleic acid motor is the RNA polymerase (RNAP), which transcribes various types of RNAs from DNA templates. Just like other molecular motors, the RNAP consists of large number of subunits, and there are multiple different types of RNAPs in eukaryotic cells [4]. Another nucleic acid motor is the ribosome, which is responsible for synthesizing a chain of amino acids (polypeptides) from a messenger RNA (mRNA). Functionally important sites of a ribosome can be found in between the ribosomal subunits [29], implying the importance of the interactions between subunits. The driving force of these nucleic acid motors is primarily ATP/GTP hydrolysis, though any enzymatic interactions with DNA/RNA and small molecules (ligands, co-factors etc) or metal ions may involve in the generation of their mechanical action.

Cytoskeletal motors (kinesin, myosin and dynein) are linear motors, tracking down the microtubules and actin filaments to transport organelles, vesicles and chromosomes for the redistribution and organization of the cellular components. Their activities are essential in cellular transport processes, and they utilize the energy of ATP hydrolysis [3-4,13-17]. Cytoskeletal motors can attach and detach actively to the actin filaments or microtubules, and they walk along the filament uni-directionally either towards plus-end or towards minus-end of the linear cytoskeletal filaments. The filaments are linear but placed in a complex manner with crossovers. Since kinesin and myosin motors are functional in *in vitro* system in simple cellular environments, they have been studied extensively at the single molecule levels using optical-trap force spectroscopy. In the single-molecule experiment, a bead is chemically attached to the motor protein and the displacement of the bead is monitored by a laser beam [3,14-15]. The bead particle trajectories obtained in optical-trap force spectrometry can reveal the motions of these motor proteins, such as single step size and run lengths. The global architectures of the kinesin and the myosin are shown in figure 1.1b. They are dimers consisting of two identical protein chains, walking in a ‘hand-over-hand’ fashion (alternating one of the two identical protein chains that bound to the microtubule) with a step size of ~8 nm [13-16].

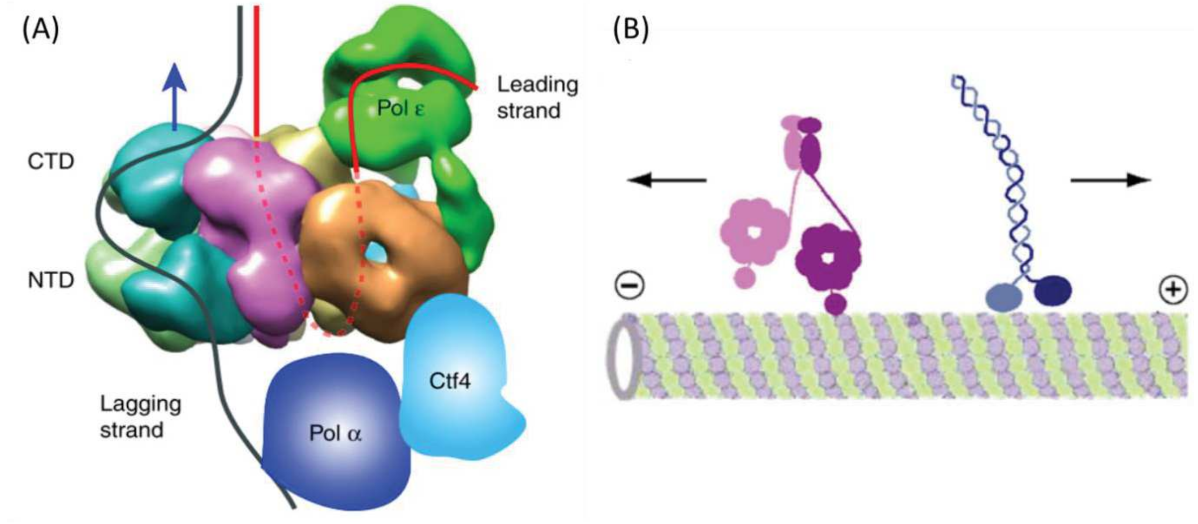


Figure 1.1 Architectures of the eukaryotic replisome (A) and the cytoskeletal motors, kinesin and myosin (B). (A) 3D electron microscopy (EM) reconstruction of *S. cerevisiae* replisome structure: Cdc45-Mcm2-7-GINS (CMG) helicase is shown in multiple color subunit (not-labelled parts), two DNAPs (Pol α and Pol ε) is shown in green and blue, Ctf4 (a fork-stabilization factor that connect helicase and polymerase) is in cyan. Replisome move in the direction of the blue arrow, and the red and black lines illustrate possible leading- and lagging-strand DNA. The size of helicase and pol ε complex is about 20 nm. In comparison to the other linear motors and rotary motors, the available information on how replisome work is very limited, most likely due to their very high complexity [5]. (B) Stepping kinesin (pink) and dynein (blue) motors that walk along the microtubule utilizing ATP hydrolysis. The microtubule has two different ends, a plus and a minus end, and the diameter of the microtubule is 25 nm. The figures (A,B) are adapted from [31] and [14], respectively.

1.1.3 Rotary motors

There are two known rotary motors: ATPases and the Bacterial Flagellar Motor (BFM). These motors have been studied for more than 40 years, yet there are still many open questions, especially regarding the detailed “power stroke” mechanism. ATPases are ATP-fueled ion pumps that transport solutes or ions across the plasma membrane and the intracellular compartment. Therefore, they are important to maintain normal physiological cellular life [18-21]. Vacuolar ATPase (V-ATPase) and F-type ATPase (F-ATPase) exhibit opposite cellular functions, though they share a highly similar overall architecture. The V-ATPases is facilitated by the ATP hydrolysis to transport protons, and produce a proton gradient. The roles of V-ATPases are involved in the intracellular membrane traffic, the entry of various viruses and toxins and tumor cell invasiveness [19]. The transmembrane Na^+/K^+ ATPase extrude three Na^+ cations and bring two K^+ cations into the cell by ATP hydrolysis, and its implication in neurological disorders has been identified [20]. F-type ATPases (ATP synthases), on the other hand, synthesize ATP- the biological energy carrier using the potential energy from the proton gradients [21]. The BFM is another powerful reversible rotary motor, driven by a proton flux, that rotates the bacterial flagella, and they will be discussed in depth in chapter 1.2.

In summary, there are five common features of all these (nucleic acids, linear and rotatory) molecular motors. First, they are nano-meter size efficient biological machines. As compared to man-made machines with 10~30 % efficiency, theses molecular machines may have a maximal possible efficiency [3]. Second, they are composed of multiple subunits, also called modules or components. Third, the mechanochemical mechanisms involve the hydrolysis of nucleotides (i.e., ATP or GTP) or the electrochemical gradient (i.e., proton ion force). This chemical energy induces conformational change of the proteins, which transfers further into the motion and other mechanics of their active subunits. Fourth, as they are small within nanometer size scales, the viscous forces dominate their activity in comparison to the inertial forces [115]. Fifth, the molecular motors are composed of the two functional units- the rotary part (rotor) that move and the stationary part (stators) that do not move and generate mechanical movement of rotor.

1.2 The Bacterial Flagellar Motor (BFM)

1.2.1 Bacteria motility

In 1683, Antony van Leeuwenhoek observed bacteria for the first time, which he referred to animalcules. He could discover them owing to their impressive motion which had captured his attention [22]. The bacterial flagella and the populations of swimming cells directed towards light source or oxygen were first observed in the nineteenth century [22]. Each cell possesses about five flagella, which are arranged polarly or laterally around the cell body, and the flagella play a key role in the bacterial movement in liquid environments [32-36]. The bacterial swimming motility takes place when individual cells swim in a random direction powered by their rotating flagellar [22-25, 32-38]. They swim in a smooth forward swimming movement (about 1 s, on average), and they also tumble, a rapid switch of direction (about 0.1 s, on average). To allow the forward swimming, the flagella form a bundle when they all rotate in the same CCW (counter-clockwise) direction. When one or more flagellar filaments switch their direction of rotation to CW (clockwise), the flagella bundle flies apart, and the cell stops, and rotates randomly, so that the following run will be in a new direction (figure 1.2).

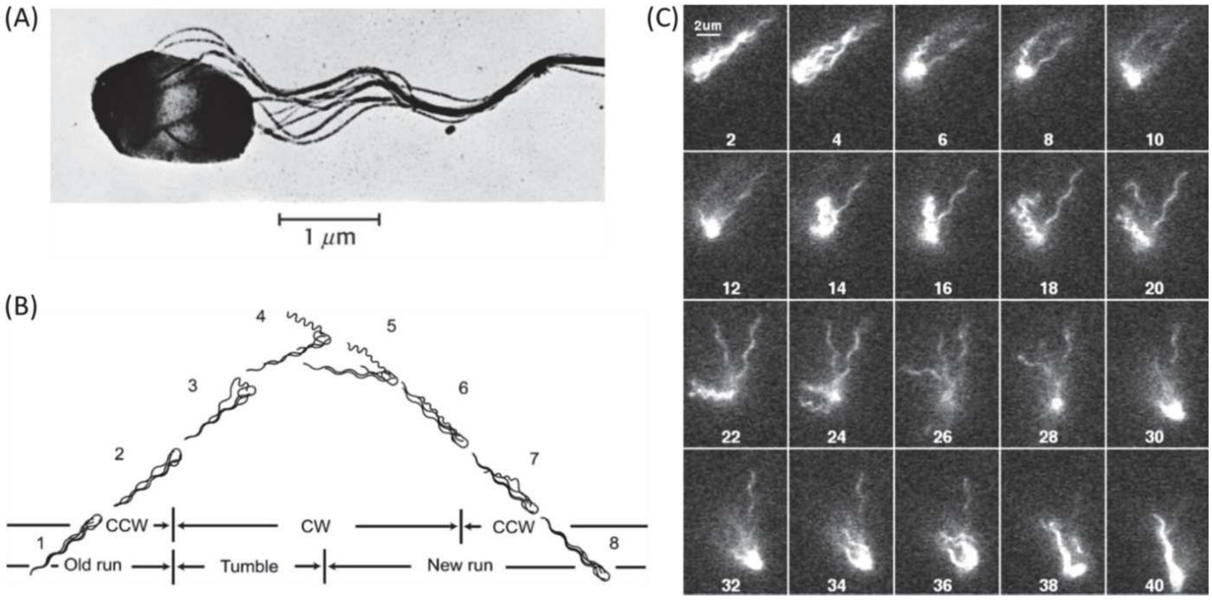


Figure 1.2 (A) a bundle of rotating helical flagella of *Salmonella*. (B) Bacterial cells swim by alternating run and tumble modes. The "run" phase is facilitated by a bundle of flagella that all rotate in CCW (counter-clock wise) direction. The "tumble" phase is facilitated by one or a few flagella reversing their rotational direction to CW (clock wise), resulting in the separation from the bundle. Then, the cell swim

in a new direction in the next run. (C) *E. coli* cell swimming with a bundle of several flagellar filaments, one filament falls out from the bundle and switch its direction of rotation. The figures are adapted from [37] (for A, B) and [38] (for C).

Besides their swimming motility in liquid environment, bacterial cells possess swarming, twitching, gliding and sliding motilities [37, 39-42] (Note: *Escherichia coli* species do not possess all those motilities). Swarming is a bacterial motility facilitated by the flagella to move over solid surfaces. The swarming motility is also operated by rotating flagella, but can be distinguished from the swimming motility for their multicellular group movement rather than individual cells movement. The swarming motility requires surfactant, increased cell to cell interactions and increased number of flagella per cell [39-42]. Twitching motility is also a surface motility, induced by the pili on the surface. The extension and retraction movement of the pilum pull the cell closer to where the pili were attached [39-42]. Gliding is another active surface movement without flagella or pili. Instead, it uses focal-adhesion complexes, which is a putative cell surface-associated complex [39, 42]. Sliding is a surface spreading movement without an active motor [37, 39]. The sliding motility relies on the secretion of surfactants which then reduce surface tension, and the cell growth further promote the colony to spread away. Thus, sliding is driven by the outward pressure of cell growth, without any aids of flagellar motors or pili.

1.2.2 Structure of the BFM and its molecular components

The bacterial flagellar motor (BFM) is the macromolecular complex which rotates each flagellum to promote bacteria to swim in liquid media [22-25]. This remarkably small (~45nm in diameter), yet powerful biological rotary electric motor can rotate up to 300 Hz for *E.coli* and up to 1700 Hz for *Vibrio* species [22-24, 44-48]. BFM is distinguishable from other molecular motors for its bi-directionality: the capacity to switch the direction of rotations from CCW to CW (or vice versa) [33-36, 51]. The motor is composed of about 30 multiple protein components [51].

The flagellum consists of at least three parts: the basal body, the hook and the filament (figure 1.3). The hook and the filament are in the cell exterior, and the basal body (rotary motor) is embedded within the cell membrane. The filament is composed of a single protein called flagellin (FliC), and how this protein is bonded each other determines the global degree of twist of the filament [22, 52-53]. Thus, although the flagella are relatively stiff, their long cylindrical structure can change between distinct polymorphic forms, either left-handed or right-handed filaments each with distinct values of curvature and twist [22, 38, 52]. These different helical waveforms of filaments may aid an efficient thrust. The hook is found between the basal body and the filament. Unlike the stiff flagellar filament, the hook is a flexible short filament (~50 nm). It functions as a universal joint that connects and transfer the torque produced by the motor to the helical filament [51, 54]. The structure of the hook is similar to the filament [53]. Each subunit of the hook (protofilament) can cycle between short and long forms, creating the curved and flexible hook. It is also shown that the disrupted bundle formation can result in atypical swimming behavior [54].

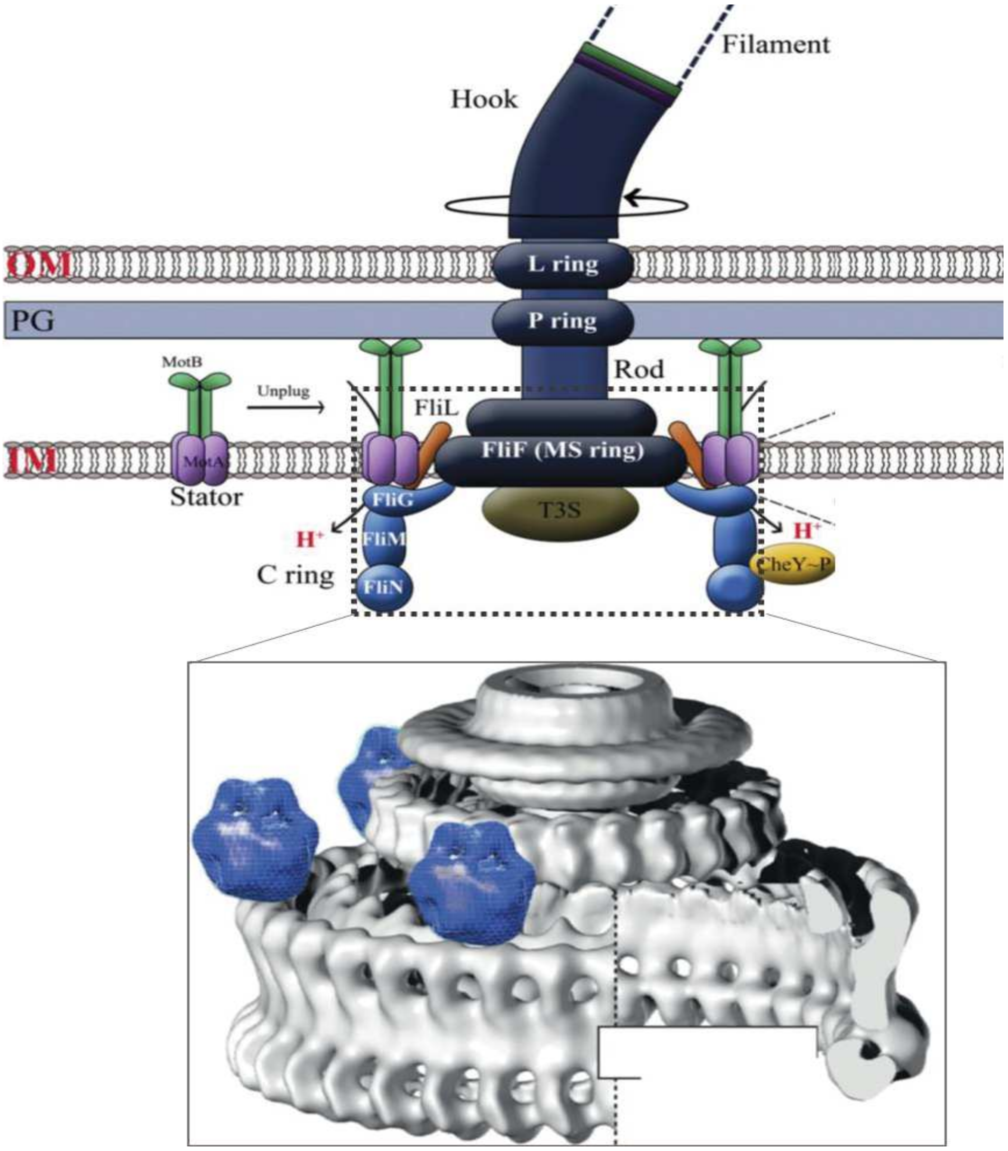


Figure 1.3 (Top) A typical macromoleular complex bacterial flagellum. It is composed of the two external components: a long helical filament and a hook (acting as a universal joint). Then, embedded in the membrane, the structure called a basal body is found, which includes four main rings (L ring, P ring, MS ring and C ring) around the periplasmic rod. MotAB are the motor stators. Acting as proton translocation channels, they can diffuse in the membrane and bind and unbind to the motor. FliL is presumably placed between the stator and the C-ring. The stators are stationary, while all the basal body is rotating. The C-ring is composed of the three proteins, FliG, FliM and FliN. This ring is called a switching complex because it is responsible for the switching from the default CCW state to CW upon the binding of the switching regulator CheY-P. Figure is adapted from [42]. (Bottom) Architecture of the bacterial flagellar motor based on 3D EM reconstructions of the CW-locked motor (grey) from *Salmonella typhimurium* (EMDB accession code: 1887) and the stator complex (blue). The stator is positioned above the outer lobe of the C-ring where it is seen in electron cryotomograms by [49]. L-ring and P-ring on the outer membrane and peptidoglycan layer are not shown. The figure is adapted from [50].

A structure of the BFM is shown in figure 1.3. The two rings L-ring and P-ring are located in the outer membrane and peptidoglycan layer. The basal body is composed of the rod, the MS-ring (26-fold symmetry) and the C-ring (34-fold symmetry), which are the rotating part of the motor [33-36]. The C-ring, composed of the three proteins, FliG, FliM and FliN, is responsible for both torque generation and switching activity. The stator complexes, bound to the peptidoglycan layer, are torque-generating transmembrane proteins MotA and MotB (figure 1.3). FliL is a recently discovered member of the motor. It is suggested that FliL may interact with both the stators and the rotor complex, thereby it can influence on the motor speed and directional bias [55-56]. Overall the structure of BFM share some features in common with man-made rotary motors, which are also composed of the rotating rotor part and the stationary stator part with ratchets that generates thrust. The C-ring complex (a part of the rotor) and the stator complexes are the main torque generating elements of the motor, hence their functions, structures and dynamics will be discussed in detail in the following sections.

1.2.3 Torque generation at the stator and rotor interface

The motor rotation is generated at the interface between the stators and the C- ring switching complex at the base of the rotor [57-60]. Torque can be defined as the twisting force that causes rotation, according to $\tau = F * R$, where F is the force applied at a distance R from the axis of rotation. The stator complex is composed of four MotA and two MotB proteins, and the topology of MotA is shown in figure 1.4a [61-62] and their structure is not known. MotA has four transmembrane (TM) regions with two long cytoplasmic loops in between TM2 and TM3 and in after TM4. These two cytoplasmic loops are known to interact with a FliG protein in the C-ring, and produce torque. Some of the charged residues of MotA that showed a critical role in torque generation are shown in figure 1.4. MotB has only one TM and a long stretch of amino acid sequences in the periplasmic space, ending in a domain known to bind to the PG. This domain in MotB enables the stator to anchor to the peptidoglycan layer, so that the stator can remain stationary while rotating the rotor [63]. The structure of stator complex is not known, though a partial structure of MotB [48] and a 3D electron microscopy (EM) reconstruction image of stator are available [63]. Stators are proton channels, transmitting protons across the membrane [23, 63]. The proton ion translocation is regulated by the proton motive force (PMF) in *E.coli*. The PMF consists of two parts: membrane potential and transmembrane ion concentration gradient.

The C-ring complex is composed of the three proteins, FliG, FliM and FliN, as shown in figure 1.3. The C-ring is about 45 nm for *E.coli* motor (other bacterial species can have larger C-ring [74]), and the stoichiometry of these proteins are estimated to be: 26 (or 34) for FliG, between 32 and 37 for FliM and 120 ± 26 for FliN [64-67]. Interestingly, the copy numbers of FliM and FliN can vary depends on the directional bias of flagellar rotation, revealing highly dynamic nature of the BFM [65, 67]. The stoichiometry between FliM and FliN is estimated to be 1:4, while mismatched stoichiometric ratio between FliG and FliM is expected [64-66]. The C-ring complex, specifically the C-terminus of FliG, interacts with the cytoplasmic region of MotA. The C-terminus of FliG protein contains multiple conserved charged residues (figure 1.4). In summary, ion translocation induces a conformational change in the cytoplasmic region of MotA, where an electrostatic interaction with FliG could occur [57-60], generating torque and. causing FliG to move.

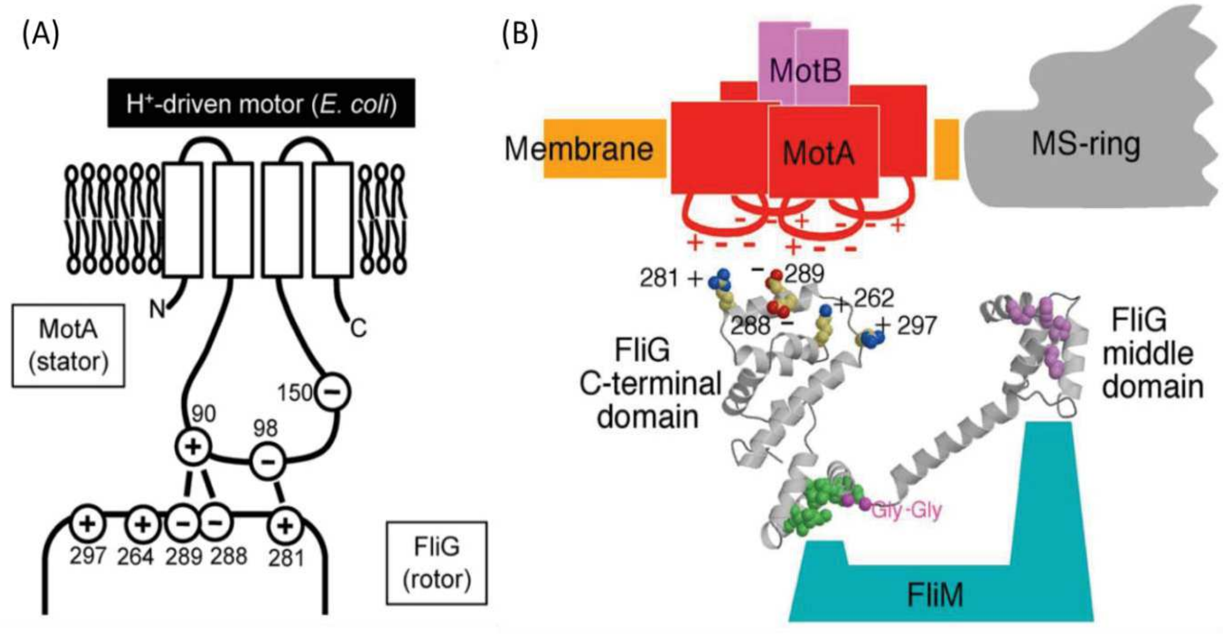


Figure 1.4 (A) MotA topology showing some of the MotA-FliG interaction sites. The charged residues at the MotA cytoplasmic loop play a critical role in torque generation. Figure is adapted from [57]. (B) The same illustration as in (A) but with the structures of the stator complex (4 MotA and 2 MotB) and the middle and C-terminal domains of the rotor protein FliG. The FliG C-terminal domain contains charged residues that interact with the charged residues of the cytoplasmic domain of MotA. Figure is adapted from [23].

It is well accepted that the proton translocation through the stator complex results in the conformation change of the stator complex, and this conformation change can push the rotor to generate torque. However, the detailed mechanics behind this energy conversion remain poorly understood. A recent study [68] proposed a model describing how torque can be generated in the BFM. Here, the two fundamental forces are electrostatic and steric. Contrary to the previous electrostatic model [60], it is proposed that the electrostatic forces ensure the proper positioning of the stators around the rotor, while the steric forces generate the actual "power stroke". The authors [68] proposed that a proline residue (Pro173) on the TM3 helix of MotA may act as a "hinge" [75], causing "bending" (fig 1.5 a). This conformational change leads to a steric push imposed on FliG, generating the power stroke. This model also proposes that two pairs of MotA are required to complete one step for CCW rotation, while the other two pairs are active for CW rotation (fig 1.5b, c).

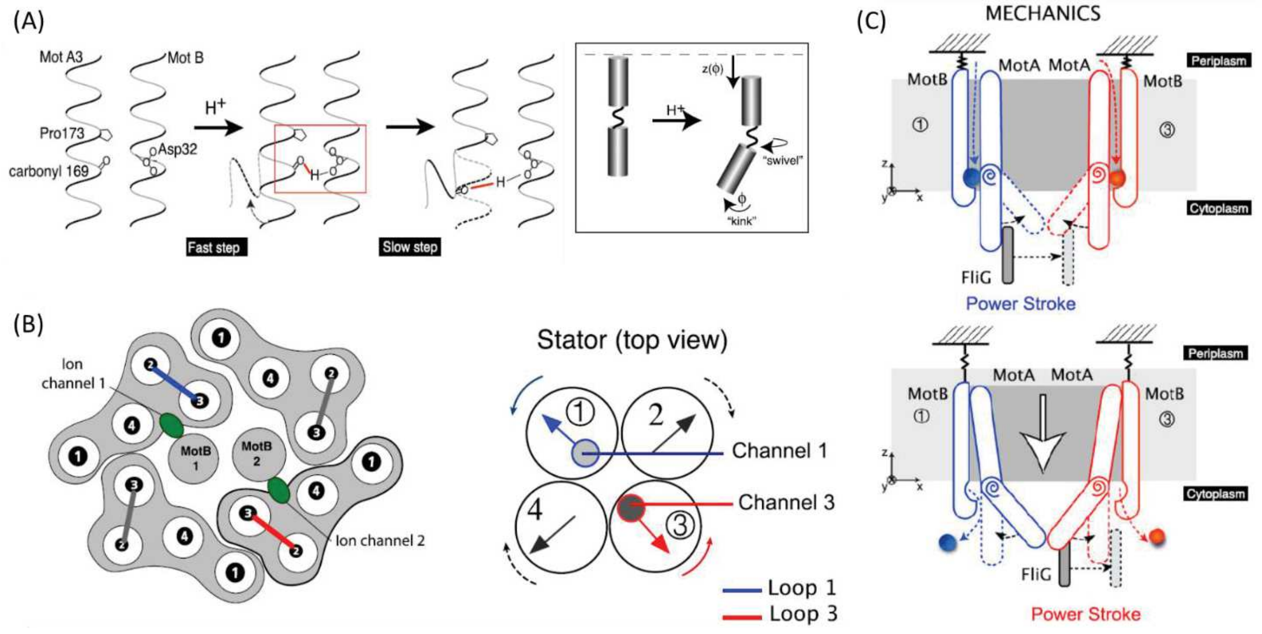


Figure 1.5 A model by [68] (A) a proton binds to Asp32 of MotB, causing re-organization of hydrogen bonds, and further into a local elastic strain in the MotA helix. This results in a drastic "swivel" and "kink" conformational change of the stator protein. (B) Proposed arrangement of stator membrane domains (four α-helices of four MotA subunits and a single helix of a MotB). The loops between helices 2 and 3 of MotA (shown as solid bars) interact with the rotor FliG. The green spots represent the ion channels. They are associated by loop 1 and 3, but not by loop 2 and 4. Loops 1 and 3 drive CCW rotation, and loops 2 and 4 drive CW rotation. (C) Proposed mechanics of the power stroke. Note that the figure A and C are a 2D depiction of a 3D process. Figures are adapted from [68]

1.2.4 Fluorescent protein tags on the BFM molecular components

Being able to observe cellular processes happening underneath the membrane brings a tremendous advantage in biological study, and it can be a way to elucidate the complex and dynamic nature of macromolecular complex. Since most of the biological molecules are not fluorescent, they need to be labelled with additional fluorescent probes such as fluorescence proteins (FP), organic dyes, quantum dots, etc. Numerous fluorescence labeling strategies and diverse advanced fluorescent probes have been developed in recent years [69-71]. To visualize BFM components, fluorescent protein labelling has been favored for its advantages in direct genetic encoding, 1:1 direct stoichiometry and less toxicity.

In general, choosing which labelling strategy to take and which fluorophore to use are critical to generate functional fusion proteins, because the addition of a fluorescent protein tag may alter the original functions of the protein of interest. Especially, when the target protein forms many functional contacts with other proteins, such as the dense arrangement of the BFM, it is important to measure functionality for each fusion protein [73]. For example, it is observed that when a fluorescent protein is attached to either N- or C- terminal sites of MotA, the motor became non-functional as measured by chemotaxis swarm plates [72]. Another example is Tar-tdEos and Tar-mEos fusion proteins in *E.coli*: tdEos tag is non-functional whereas mEos is partially functional, revealing the FP dependent functional specificity of Tar protein [73]. There are four general requirements for functional fusion proteins. First, the tag must allow functional interactions between all the interacting proteins and allow space around the functional domains or binding sites, and not sterically interfere with specific surfaces of the protein. Second, the tag must not aggregate or multimerize, which may affect the function or the location of the protein. Third, the tag must hold proper folding, stability and protein expression properties. Fourth, the tag must not trigger the target of the protein for degradation or modification [72]. The functionality of a particular fusion protein cannot be predicted; therefore, it is necessary to test combinations of different FPs fused to either the N- or C- termini of proteins, with or without a linker.

In BFM, several motor protein components (stator protein MotB and C-ring proteins) have been fused to a FP to perform fluorescence microscopy that uncover the dynamic nature of the motor functions in response to the different loads [76-78], ion motive force [44, 79], mutant stator proteins [80] and rotational bias [65,67]. To study the dynamics of the stators, FP tagged MotB has been used [82]. An example study that explored the dynamic properties of stator units is shown in figure 1.6. Although FP tagging at the C-terminal is generally better in preserving the localization of the native protein than tagging at the N-terminal [81], the FP was often attached to the N-terminus of MotB since the C-terminus of MotB is in periplasm. The motor driven by FP-MotB stators is functional but with a reduced motility [82], and the precise reasons for such reduced motility has not been investigated. Despite the fact that the FP tag on MotB are lying at the stator-rotor interface, where the torque and switching are generated, it was only speculated that the FP tag on MotB somehow interferes with the power-stroke mechanism and no studies were available to understand how the FP tag on the stator affects the torque generation and switching mechanism. Therefore, we aim to perform detailed functional investigation of the FP-MotB fusion stator motors in this thesis (Chapter 3).

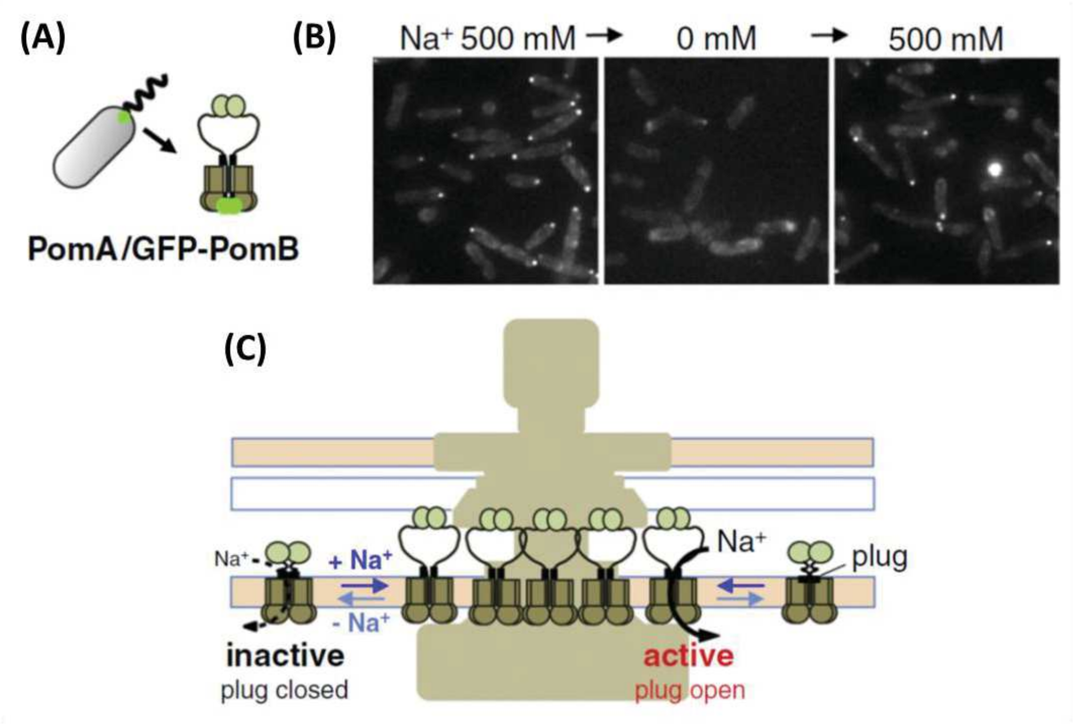


Figure 1.6 Dynamic properties of the stator proteins revealed by fluorescence microscopy. (A) Schematic drawing of the PomA/GFP-PomB stator complex. Unlike MotA/MotB, the PomA/PomB is a sodium translocating stator. (B) The stators localization around the rotor is found in the polar flagellum of *Vibrio alginolyticus*, only when the sodium ions are present. This demonstrates that the sodium is required for stator assembly. (C) Schematic drawing of the dynamical engagement/disengagements of the sodium-dependent stator to the rotor. The PomA/PomB stator units associate to the motor only in the presence of Na^+ ions, and change the conformation from the inactive state to the active state. Figure is adapted from [44].

1.2.5 Stator dynamics

Successive incorporation of stators in the motor have observed since 1984 [83]. This successive incorporation is referred to "resurrection", because it is observed that the paralyzed MotB knockout motors began to spin again within several minutes after inducing the MotB gene, with their rotational speed changing in a series of steps [83]. The number of stators bound to the motor is directly proportional to the rotation speed. The maximum number of torque-generating units in the motor was first suggested to be 8 [84], then later it is suggested to be at least 11 [88]. The observation that torque can be restored in a series of equally spaced steps, especially by the re-introduction of the protein MotA and MotB, confirmed the critical role of stators in torque generation, revealed their dynamical activity in the motor and provided a useful tool (resurrection experiment) to study the torque generation mechanism at the single torque generating unit [83-92, 79].

Dynamical turnover of stator proteins from the membrane to the motor was also observed [82]. The stators have been observed to remain bound into the motor with a mean dwell time of ~30 seconds, before disengaging and diffusing away in the membrane. A similar remodeling of the motor components was also observed for FliN and FliM, the two C-ring proteins [65-67]. The stator incorporation is dependent on the applied load on the motor as well as the ion motive force (IMF). Load dependent stator incorporation was observed by three different labs [76-78]: the higher load on the motor, the higher number of stator bound to the motor. It is suggested that the periplasmic domain of MotB may be involved in load-dependent of the motor [93], but further studies will be required to identify the detailed mechanisms of mechano-sensing behavior of the motor. When IMF is disrupted, the stators diffuse away from the motor. Then, when the disrupted IMF is restored, the stators can incorporate back to the motor, showing that IMF presence is necessary for continued motor integrity [44, 79]. All these stator dynamics were studied using fluorescence microscopy. As a result of the two recent discoveries of the load- and IMF- dependent stator dynamics, the number of the stators bound in the motor can be controlled 1) by the control of the expression level [83-84, 88] or 2) by imposing a low mechanical load or 3) by the control of the IMF. The stator incorporation can be observed by the step-wise speed changes of the rotating bead attached to the flagellum.

1.2.6 Switching dynamics

The BFM rotates in both counterclockwise (CCW) and clockwise (CW) directions. This bidirectional torque generation is symmetric as shown by [110]. The switch of the direction of rotation is initiated when CheY-P proteins (the switching regulating- phosphorylated CheY protein) bind to the switching complex, which presumably cause a conformational change in FliM. This conformational change triggers allosterically a rotation of the orientation (~180 degree) of the charged ridge in the C-terminus of FliG, which interacts with the charged cytoplasmic loop of MotA. This 180° rotation of the charged FliG domain presumably causes different electrostatic interactions with the stators, and this process is thought to be responsible for change in rotation direction (figure 1.7a) [97, 101-104]. The structure of FliG and its rotational conformational changes were proposed [101-103]. The structure of the full-length FliG protein from *Aquifex aeolicus* consists of distinct amino-terminal(N), middle (M) and carboxy-terminal(C) globular domains. These domains are connected by the two long helices (helixNM and helixMC). The N-terminal domain is responsible for the flagellar assembly and binding to FliF proten in MC ring. The C-terminal domain is responsible for torque generation and rotational switching (inset in figure 1.7a) [101]. Numerous mutagenesis studies have already shown the significant roles of the charged residues in helix C5 for torque generation [57-60, 109]. The structural analysis predicts that 34 copies of FliG form the C-ring (figure 1.7b), a number which is higher than the previously proposed 26-fold symmetry, and that more or less than 34 FliG proteins create improbable domain arrangements and inconsistency with EM data (~45 nm in diameter) [101].

The domain right before the C-terminus of FliG, specifically the 245MFXF248 motif and Asparagine 216 in the ARMc domain, is suggested to be acting as a hinge that generates a smooth conformational rotation of the helix C5 [102,111]. A mutagenic analysis of FliG also revealed that residue substitutions in the middle domain of FliG (so called connector residues 180~199) generate wide range of unusual functionalities such as strongly CW biased, infrequent switchers, rapid switchers, and transiently or permanently paused [105]. It is also found that when FliG has a three-amino-acid deletion (ΔPro-Ala-Ala) at positions 169 to 171, it became an extreme clockwise (CW) motor [100]. This mutant is called a CW-locked mutant of *Salmonella* (PAA).

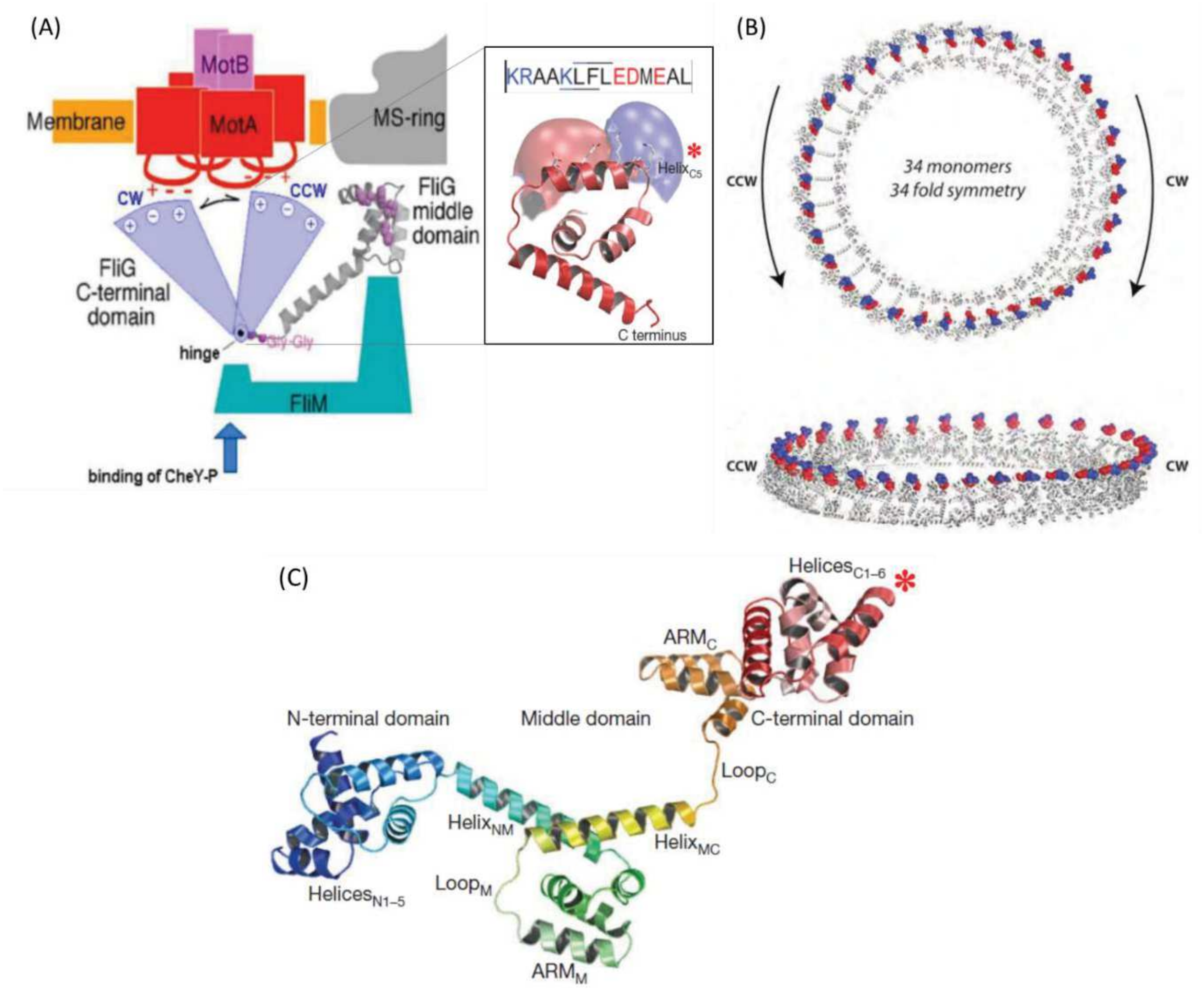


Figure 1.7 (A) An illustration showing a hypothetical FliG movement and its interaction with the cytoplasmic loops of MotA. A CheY- P molecule bound to FliM trigger the FliG to have two conformations: CW rotation states or CCW rotation state. Whether the FliG can adapt another conformational state than these two simple two-state (CW/CCW) is still an open question. A crystal structure of the torque generating helix C5 is shown in inset. Positive and negative charged residues are in red and blue, respectively. (B) A FliG ring displaying opposing switch states viewed from the top and side. The structural model by [101] predicts that having two different conformational state (the right side of the rings are in a CW state whereas the left side are in a CCW state) in a single C-ring is possible. (C) Structural overview of the full-length FliG monomer. Torque helixC5 is labeled with a red asterisk. The conformational change of the ARMc and loop$_c$ (just below the C-terminal domain) may rotate the C-terminal domain, and promote the switching of the rotational direction. Figures adapted from [99] and [101].

In strains lacking CheY (the chemotaxis regulator), motors rotate exclusively CCW, and cells swim smoothly [132]. The structural mechanism of how a CheY-P binding is associated with a switching event of the motor remains unknown. However, conformational spread model describes the coupling between CheY-P binding and the cooperative FliG subunits conformation change [98]. In this model, switching events follow the allosteric cooperativity coupled with the conformational spread model. The CheY-P binding initiates conformational change of the CCW state FliG to CW state FliG, and these changes can be spread to adjacent FliG subunits to share a identical conformational state (either CCW or CW state) [113]. Switching is not an instantaneous event- and typically, it has an exponential distribution with a characteristic time of 4-8ms where the fastest events take less than 1ms and the slowest take up to 100ms. The switching time distribution does not change with the direction of the switch (CCW-CW or CW-CCW) [98]. Between switching events, flagellar motors rotate in CCW and CW at about the same speed. When this symmetry of rotation speed is not reached during a single switch, the switch is termed incomplete. The conformational spread model propose that incomplete switches can occur when the conformational change of the entire C-ring is incomplete, with the result of collapsing back to the previous stable (CW or CCW) state (figure 1.8a).

Switching events occur randomly about once a second in average. Switching frequency responses to chemotactic stimuli transiently. For example, addition of attractant or removal of repellent causes reduced switching frequencies [95-96]. Switching dynamics under low load (60-nm- diameter gold spheres on a hook) revealed linearly increased switching rates with respect to motor torque, with a maximum frequency at the intermediate rotation speed [107]. Switching measurement on the motors lacking flagella is important to eliminate any possibility that the filaments contribute to the switching dynamics [114]. Consequently, it is suggested that switching regulation is coupled to torque generation, meaning that the number of stators bound to the motor have an effect on the switching dynamics [99]. In this model, stators are hypothesized to accelerate the FliG conformation flipping rates (conformational change between CCW and CW), thereby a motor is predicted to switch more often with more stators, with a maximum switching rate at intermediate speed (figure 1.8 b and c). The switching rate shows a maximum at intermediate speed, because a stator 1) accelerates the conformation flipping rates of individual rotor switching units, which favors slower motor speed and 2) increases torque, affecting more switching units within unit time, which favors faster speed.

Load dependent switching dynamics was also observed [106]. At high load (a bead bigger than 1.44 µm), CW intervals lengthened appreciably, whereas CCW intervals remained about the same for all loads. In other words, the probability of the motor rotating in CW (CW time) increases under high load, especially when the motor spins slowly [106]. Since CW rotation is associated with tumbling and changing the direction of swimming, the cells may have evolved to rotate longer in CW at high load, which they sense as a viscous environment. The "dwell time" indicates the time that the motor remains in each CW or CCW state. The distributions of CCW and CW intervals follow an exponential distribution, which agrees with the two-state Poisson process motor switching model [107].

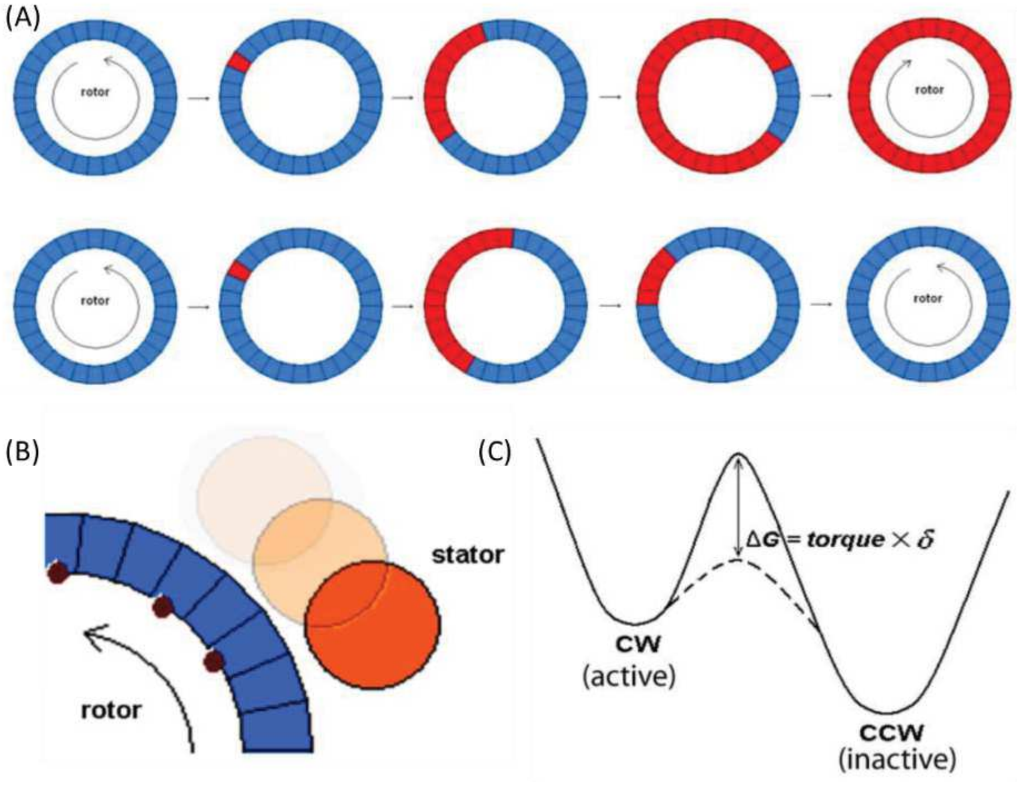


Figure 1.8. (A) In conformational spread model, successful growth of conformational flips will eventually encompass the whole ring (complete switching); but more frequently, the event may collapse back (incomplete switching). Figure adapted from [112]. (B-C) How switching events can be influenced by torque generation. (B) A cartoon showing stator acting on more rotor switching unit (RSU) with increasing motor speed per unit time. (C) Stator torque lowers the activation energy barrier for those contacting RSUs (easier to move from CW state to CCW state, vice versa). In this case, the flipping rate in the conformational spread model is increased. Figure is adapted from [99]

1.3 Evolutionary dynamics of the bacterial flagellar motor

1.3.1 Biological complexity, modularity and evolvability

Biological complexity can be found at the level of organisms, organs, cell and macromolecular protein complexes. Molecular motors, such as the ribosome, the spliceosome and the BFM, are macromolecular proteins that possess great biological complexities. They are assembled by multiple differentiated proteins that are structurally linked each other to perform biological functions. Asking a question like how these complex molecular machines have evolved has been one of the biggest challenges in biology. Neodarwinian biological evolution is driven by two mechanisms: genetic variation and natural selection [116]. Genetic variations can arise by multiple molecular mechanisms, and the natural strategies to modify genomic DNA sequences (mutations) can be categorized in: 1) small local sequence changes by replication of infidelities, 2) intra-genomic rearrangement of DNA segments by re-combinational reshuffling of genomic DNA segments, 3) the addition of a foreign organism DNA segment (i.e., horizontal gene transfer) [116]. In principle, genetic variation is a random process rather than a targeted response to a specified need. Natural selection arises when the genetic variations have a beneficial effect such that the organisms can fit into the new environment better. As a result, it is the natural selection that directs the process of evolution, leading evolutionary adaptation (increased fitness) in altered physiological or environmental conditions.

Epistasis refers to the phenomenon of the effects of genotype changes at one locus affects the fitness contribution of another locus [117]. When a mutation on a gene is largely constrained by another gene, they are linked by an epistatic interaction. The principle of protein interactions can be seen as epistatic interactions. When a system is composed of many interacting components like molecular motors, epistasis is likely strong. This implies that evolution towards a novel function on the basis of small random mutations along a step-wise adaptive trajectory is statistically unlikely, because typically mutations compromise existing functions rather than generate new ones [117-118]. Therefore, the two-step process of adaptive evolution (genetic variations followed by natural selection) alone cannot fully explain the complexity found in macromolecular complexes. Numerous population genetics and comparative genomic studies suggest another evolution force. It is suggested that complex molecular motors may have evolved by compositional evolution that is either neutral or adaptive [133].

Constructive neutral evolution takes into account non-adaptive evolutionary forces such as genetic drift, genetic variations as well as the emerging concept of modularity and evolvability of biological components [120-122]. This mode of evolution explains that biodiversity and complexity have emerged through horizontal gene transfer, gene duplication and divergence. The failure of a new introduced gene to correctly interact with preexisting protein subunits can provide a potential ground for neutral evolution to cause complexification of protein complexes (figure 1.8). The directional force of the neutral evolution can be the increased fitness of the foreign protein subunit of the new host rather than or not limited by the natural selective forces such as adaptation. Adaptive evolution, on the other hand, focuses on how genetic variations are translated into phenotypes and achieving a higher fitness based on mutations and natural selection [123]. This mode of evolution may play a role in piecing together different pre-existing proteins and promote their further functional integration. Therefore, in the eyes of compositional evolution, the collaborative efforts by the constructive neutral evolution and adaptive evolution may take place to generate biological complexity and diversity in the process of biological evolution.

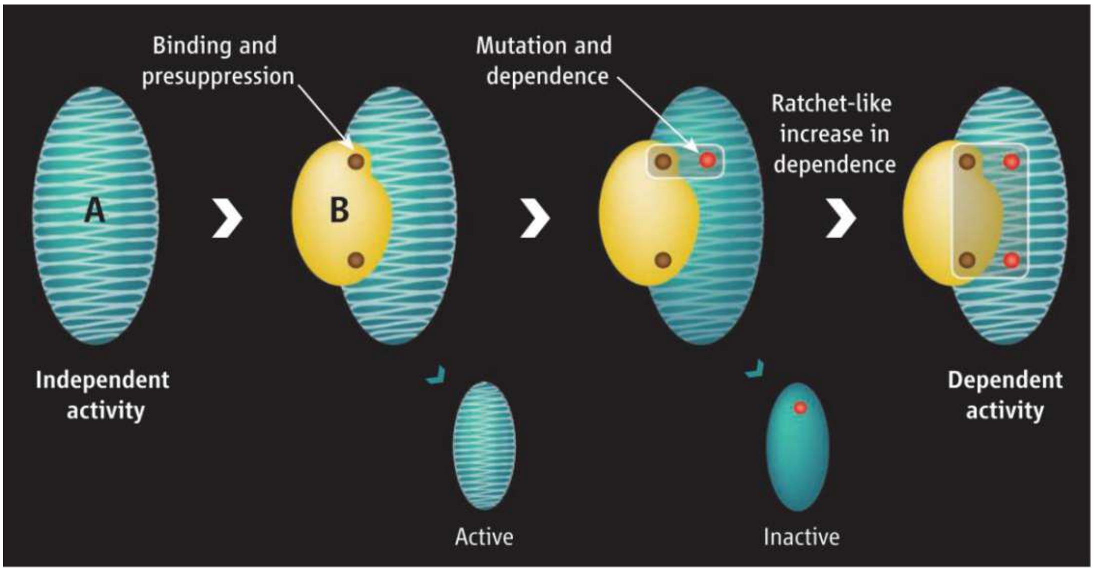


Figure 1.8 Constructive neutral evolution towards a macromolecular complex. Independent cellular components A and B fortuitously interact, then, gradually they become interdependent, because the interaction allows (presuppresses) mutations that would otherwise inactivate A and increase the chances of further mutations. This process eventually limits the chance of each component to reverse back to function independently. In this illustration, two modules A and B evolved collectively to generate a single dependent activity. Figure is adapted from [122].

If the interactions between protein components in macromolecular complexes are relatively few, weak and non-specific, instead of many, strong and specific, this mode of interactions can form modules. Modularity in protein complexes allows each component to interact with only a subset of other components in a system, and promote functional independence [117, 124]. Biological systems' modularity can be described as modular organizations of biological structure [125]. If the prospective components of a protein complex possess a degree of modularity—that is to say, a degree of functional and structural independence—then evolution might be capable of forging new functional interaction between them. Modularity affects how biological systems evolve. First, it reduces epistasis. Owing to the less epistatic interactions, less constraints on adaptation can be achieved [126]. Second, due to the limited epistasis interactions, it prevents potentially harmful protein-protein interactions that lead to non-beneficial functionalities, but promote an efficient protein network [124]. Third, it may enable modules to acquire multiple functions, because the functional independence of each module can allow each module to adapt novel functionalities depending on where and when they are expressed in cellular backgrounds. Forth, it allows component exchange, resulting in horizontal gene transfer (HGT), which is important in shifting the global phenotype and a central force in microbial evolution [127]. In conclusion, modularity is an important as it eliminates unnecessary needs of searching/trying out the entire space of possibilities (random mutations towards a gradual trajectory of beneficial mutations). Modularity can limit the search to a subspace of the modular solutions, instead of the entire space of possibilities. It can also facilitate evolution of modular protein complexes since separated independent components enable more evolution (from each discernable module) to occur than from a single complex system [125]. Therefore, with the help of modularity, the biological system enhances evolvability- the capacity of a system for adaptive evolution [117].

Compositional evolution suggests that complex systems like molecular motors evolve gradually by the addition and exchange of pre-existing components (which work as modules) [119]. However, extensive experimental evidences that provide molecular mechanisms underlying this mode of evolution process is lacking. In order to provide an experimental insight into how evolution can cause functional integration of pre-existing proteins with a degree of fortuitous pre-adaptation, Bertus Beaumont lab developed an evolution experiment with a BFM modified by a component exchange. Exploring the evolutionary dynamics of the macromolecular machines may lead us to a better understanding of how that particular system evolved as well as their unique functional features, and in the ends, perhaps it can help unravel the origins of biological complexity.

1.3.2 Experimental evolution of BFM

Biological component exchange between homologues has been studied extensively and shown that inter-species gene replacement, also called HGT, often results in either neutral or deleterious phenotype [128]. The lab of Bertus Beaumont, using the BFM as the experimental model system, systematically exchanged stator proteins (both MotA and MotB proteins) and generated a series of stator chimeric BFMs, termed cBFM. The list of cBFMs can be found in Table 5.1. We use the term *foreign* stator to indicate the exchanged stators. The stator complex was exchanged instead of other BFM protein components, because they are the main torque generating element of the motor and structurally independent part of the motor. Stators diffuse around the membrane, and dynamically associate and disassociate from the rotor complex during rotation. In addition, several studies showed signs of component exchange of stator proteins across bacterial species [129, 130]. For example, *Shewanella oneidensis MR-1* possess two stator systems: MotAB and PomAB stators, which confer a different mode of ion translocations (protons and sodium ions), and it is speculated that MotAB were acquired by HGT as a consequence of adaptation to a low-sodium environment [129]. Some bacteria species also encode two distinct flagellar systems, and utilize more than one mode of motility, which are adaptable based on the environment [131]. Therefore, the stator exchange in BFM across the bacteria kingdom provides an excellent potential ground in which modularity and evolvability can be tested.

Compositional evolution explains that when a foreign component is integrated into a system, they can be immediate compatible or potentially compatible (evolutionarily generated mutations modified the component to be compatible) or incompatible [119]. The compatibility deals with the functionality as well as the capacity to transcribe and express the exchanged foreign proteins in adequate level. Importantly, it is also related with the capacity to generate high probabilistic mutations that lead to an increased fitness (evolved compatibility). In order to gain an experimental insight into compositional evolution, Bertus Beaumont lab challenged these compatibility schemes in a series of cBFM strains. The cBFM experimental evolution was carried out by applying a selective advantage of the motility during chemotaxis assay. When cells are motile (immediately compatible), they swim towards more nutrient rich environment in the semi-soft agar, moving further away from where they were first introduced. When cells are non-motile (incompatible), they remain at the original spot without making further migration. Those that were not motile immediately were kept on growing in the semi-soft agar, to check if they could evolve motility (potentially compatible) by observing the appearance of migration. This “later born” motile group appears as a population of a flare in semi-solid agar. In summary, in this evolutionary experiment, the following questions were asked: will the motor still be functional if we swap the stator component? If not, can the non-functional motor be evolved to become functional? (figure 1.9)

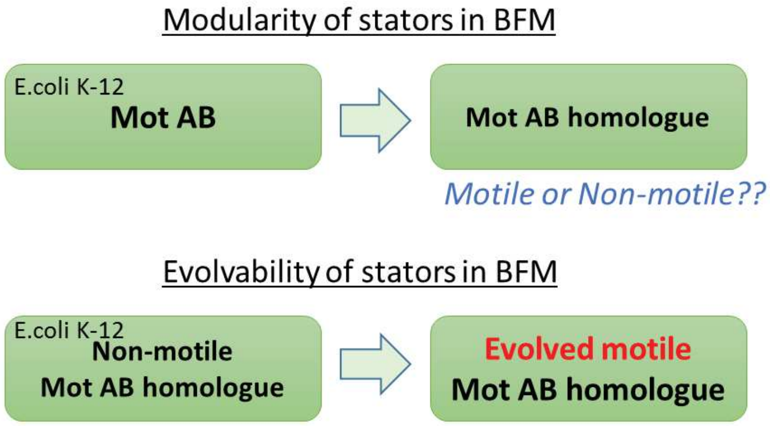


Figure 1.9 Modularity of BFM was investigated when foreign components (Mot AB homologues) are exchanged. Of those that did not show immediate compatibility (non-motile functionality), evolution experiment was performed to test their evolvability.

The cBFM experimental evolution provided an answer to the phylogenetic compatibility range, and the results are discussed in chapter 5.1. Followed up from the cBFM evolution experiment, we identified a set of foreign evolved stators that carries beneficial mutations, which described in chapter 5.2. Then, these beneficial mutations were identified and analyzed (chapter 5.3). Functional characterizations of a series of evolved cBFM and their ancestor cBFM were performed and described in chapter 5.4 and 5.5. The functional characterization of cBFM was made at the three different motility levels: population swimming in semi-solid agar, single cell swimming in liquid media and single motor rotation in liquid motility buffer. As a final point, cBFM provided an excellent model system to explore novel functionality of the BFM, since cBFM can be seen as a *E.coli* flagellar motor driven by genetically modified or non-modified foreign stators. The detailed functional characterizations of the cBFM provide the phenotypic effects caused by the foreign stators and their mutations.

References

1. Froese.T., Virgo. N., Ikegami. T., (2014). Motility at the origin of life: its characterization and a model. Artificial Life 20: 55–76
2. "What Is Biophysics?". Biophysics.org. n.p., 2016.
3. Kolomeisky, A. B. (2013). Motor proteins and molecular motors: how to operate machines at the nanoscale. Journal of Physics. Condensed Matter: An Institute of Physics Journal, 25(46), 463101.
4. Chowdhury, D. (2013). Stochastic mechano-chemical kinetics of molecular motors: A multidisciplinary enterprise from a physicist's perspective. *Physics Reports*, *529*(1), 1–197.
5. Goel, A., & Vogel, V. (2008). Harnessing biological motors to engineer systems for nanoscale transport and assembly. *Nature Nanotechnology*, *3*(8), 465–475.
6. Liu, M. S., Todd, B. D., & Sadus, R. J. (2005). Cooperativity in the motor activities of the ATP-fueled molecular motors. Biochimica et Biophysica Acta - Proteins and Proteomics, 1752(2),
7. Rothwell, P. J., & Waksman, G. (2005). Structure and mechanism of DNA polymerases. *Advances in Protein Chemistry*, *71*(4), 401–440.
8. Bebenek, K., & Kunkel, T. A. (2004). Functions of DNA polymerases. *Advances in Protein Chemistry*, *69*, 137–165.
9. Abid Ali, F., & Costa, A. (2015). The MCM Helicase Motor of the Eukaryotic Replisome. *Journal of Molecular Biology*, *428*(9), 1822–1832.
10. Pyle, A. M. (2008). Translocation and unwinding mechanisms of RNA and DNA helicases. *Annual Review of Biophysics*, *37*, 317–336.
11. Patel, S. S., Pandey, M., & Nandakumar, D. (2011). Dynamic coupling between the motors of DNA replication: Hexameric helicase, DNA polymerase, and primase. *Current Opinion in Chemical Biology*, *15*(5), 595–605.
12. Kim, D. E., Narayan, M., & Patel, S. S. (2002). T7 DNA helicase: A molecular motor that processively and unidirectionally translocates along single-stranded DNA. *Journal of Molecular Biology*, *321*(5), 807–819.
13. Rayment, I. (1996). Kinesin and myosin: molecular motors with similar engines, 501–504.
14. Lipowsky, R., Chai, Y., Klumpp, S., Liepelt, S., & Moller, M. J. I. (2006). Molecular motor traffic: From biological nanomachines to macroscopic transport. *Physica A: Statistical Mechanics and Its Applications*, *372*(1 SPEC. ISS.), 34–51.
15. Higuchi, H., & Endow, S. A. (2002). Directionality and processivity of molecular motors. *Current Opinion in Cell Biology*, *14*, 50–57.
16. Endow, S. A. (2003). Kinesin motors as molecular machines. *BioEssays*, *25*(12), 1212–1219.
17. Goel, A., & Vogel, V. (2008). Harnessing biological motors to engineer systems for nanoscale transport and assembly. *Nature Nanotechnology*, *3*(8), 465–475.
18. Mukherjee, S., Bora, R. P., & Warshel, A. (2015). Torque, chemistry and efficiency in molecular motors: a study of the rotary–chemical coupling in F1-ATPase. *Q Rev Biophys*, *48*(4), 37–54.
19. McGuire, C., Cotter, K., Stransky, L., & Forgac, M. (2015). Regulation of V-ATPase assembly and function of V-ATPases in tumor cell invasiveness. *Biochimica et Biophysica Acta - Bioenergetics*, *1857*(8), 1213–1218.
20. Isaksen, T. J., Lykke-hartmann, K., & Lykke-hartmann, K. (2016). Insights into the Pathology of the α 2 -Na + / K + -ATPase in Neurological Disorders; Lessons from Animal Models, *7*. 1–8.
21. Stewart, A. G., Laming, E. M., Sobti, M., & Stock, D. (2014). Rotary ATPases-dynamic molecular machines. *Current Opinion in Structural Biology*, *25*, 40–48.
22. H.Berg. (2004). *E.coli* in Motion. Springer.
23. Blair, D. F. (2003). Flagellar movement driven by proton translocation. *FEBS Letters*, *545*(1), 86–95.
24. Sowa, Y., & Berry, R. M. (2008). Bacterial flagellar motor. *Quarterly Reviews of Biophysics*, *41*(2), 103–32.

25. Thormann, K. M., & Paulick, A. (2010). Tuning the flagellar motor. *Microbiology (Reading, England)*, *156*(Pt 5), 1275–83.
26. van Den Heuvel, M. G. L., & Dekker, C. (2007). Motor proteins at work for nanotechnology. *Science (New York, N.Y.)*, *317*(5836), 333–6.
27. Wang, J. (2009). Can man-made nanomachines compete with nature biomotors? *ACS Nano*, *3*(1), 4–9.
28. Charvin, G., Bensimon, D., & Croquette, V. (2003). Single-molecule study of DNA unlinking by eukaryotic and prokaryotic type-II topoisomerases. *Proc. Natl. Acad. Sci. USA*, *100*(17), 9820–9825.
29. Liljas, A. (2006). Deepening Ribosomal Insights. *ACS Chemical Biology*, *1*(9), 567–569.
30. O'Donnell, M., & Li, H. (2016). The Eukaryotic Replisome Goes Under the Microscope. *Current Biology*, *26*(6), R247–R256.
31. Sun, J., Shi, Y., Georgescu, R. E., Yuan, Z., Chait, B. T., Li, H., & O'Donnell, M. E. (2015). The architecture of a eukaryotic replisome. *Nature Structural & Molecular Biology*, *22*(November), 1–9.
32. Mccarter, L. L. (2005). Multiple Modes of Motility: a Second Flagellar System in Escherichia coli, *187*(4), 10–13.
33. Berg, H. C. (2013). Cell motility: turning failure into function. *Nature Physics*, *advance on*(8), 460–461.
34. Berg, H. C. (2008). Bacterial flagellar motor. *Current Biology*, *18*(16), R689-91. Journal Article.
35. Minamino, T., Imada, K., & Namba, K. (2008). Molecular motors of the bacterial flagella. *Current Opinion in Structural Biology*, *18*(6), 693–701.
36. DeRosier, D. J. (1998). The turn of the screw: the bacterial flagellar motor. *Cell*, *93*(1), 17–20.
37. Elgeti, J., Winkler, R. G., & Gompper, G. (2014). Physics of Microswimmers - Single Particle Motion and Collective Behavior.
38. Turner, L., Ryu, W. S., & Berg, H. C. (2000). Real-time imaging of fluorescent flagellar filaments. *Journal of Bacteriology*, *182*(10), 2793–2801.
39. Kearns, D. B. (2010). A field guide to bacterial swarming motility. Nature Reviews. Microbiology, 8(9), 634–44.
40. Chaban, B., Hughes, H. V., & Beeby, M. (2015). The flagellum in bacterial pathogens: For motility and a whole lot more. *Seminars in Cell and Developmental Biology*, *46*, 91–103.
41. Wadhams, G. H., & Armitage, J. P. (2004). Making sense of it all: bacterial chemotaxis. *Nat.Rev.Mol.Cell Biol.*, *5*(12).
42. Harshey, R. M., & Partridge, J. D. (2015). Shelter in a Swarm. *Journal of Molecular Biology*, *427*(23), 3683–3694.
43. Croze, O. A., Ferguson, G. P., Cates, M. E., & Poon, W. C. K. (2011). Migration of chemotactic bacteria in soft agar: Role of gel concentration. *Biophysical Journal*, *101*(3), 525–534.
44. Kojima, S. (2015). Dynamism and regulation of the stator, the energy conversion complex of the bacterial flagellar motor. Current Opinion in Microbiology.
45. Minamino, T., & Imada, K. (2015). The bacterial flagellar motor and its structural diversity. *Trends in Microbiology*, 1–8.
46. Kojima, S., & Blair, D. F. (2004). The Bacterial Flagellar Motor: Structure and Function of a Complex Molecular Machine. In B. T.-I. R. of Cytology (Ed.) (Vol. Volume 233, pp. 93–134). CHAP, Academic Press.
47. Thormann, K. M., & Paulick, A. (2010). Tuning the flagellar motor. *Microbiology (Reading, England)*, *156*(Pt 5), 1275–83.
48. Yonekura, K., Maki-Yonekura, S., & Homma, M. (2011). Structure of the flagellar motor protein complex PomAB: implications for the torque-generating conformation. *Journal of Bacteriology*, *193*(15), 3863–70.
49. Chen, S., Beeby, M., Murphy, G. E., Leadbetter, J. R., Hendrixson, D. R., Briegel, A., … Jensen, G. J. (2011). Structural diversity of bacterial flagellar motors. *The EMBO Journal*, *30*(14), 2972–81.
50. Stock, D., Namba, K., & Lee, L. K. (2012). Nanorotors and self-assembling macromolecular machines: the torque ring of the bacterial flagellar motor. *Current Opinion in Biotechnology*, *23*(4), 545–554.
51. Morimoto, Y., & Minamino, T. (2014). Structure and Function of the Bi-Directional Bacterial Flagellar Motor. *Biomolecules*, *4*(1), 217–234.

52. Calladine, C. R. 1978. Change in waveform in bacterial flagella: the role of mechanics at the molecular level. J. Mol. Biol. 118:457–479.8.
53. Fujii, T., Kato, T., & Namba, K. (2009). Specific Arrangement of alpha-Helical Coiled Coils in the Core Domain of the Bacterial Flagellar Hook for the Universal Joint Function. *Structure*, *17*(11), 1485–1493.
54. Brown, M. T., Steel, B. C., Silvestrin, C., Wilkinson, D. a, Delalez, N. J., Lumb, C. N.,Berry, R. M. (2012). Flagellar hook flexibility is essential for bundle formation in swimming Escherichia coli cells. *Journal of Bacteriology*, *194*(13), 3495–501.
55. Partridge, J. D., Nieto, V., & Harshey, R. M. (2015). A New Player at the Flagellar Motor : FliL Controls both Motor Output and Bias, *6*(2), 1–11.
56. Fabela, S., Domenzain, C., De la Mora, J., Osorio, A., Ramirez-Cabrera, V., Poggio, S., … Camarena, L. (2013). A Distant Homologue of the FlgT Protein Interacts with MotB and FliL and Is Essential for Flagellar Rotation in Rhodobacter sphaeroides. *Journal of Bacteriology*, *195*(23), 5285–96.
57. Takekawa, N., Kojima, S., & 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. *Journal of Bacteriology*, *196*(7), 1377–85.
58. Morimoto, Y. V, Nakamura, S., Hiraoka, K. D., Namba, K., & Minamino, T. (2013). Distinct roles of highly conserved charged residues at the MotA-FliG interface in bacterial flagellar motor rotation. *Journal of Bacteriology*, *195*(3), 474–81.
59. Garza, a G., Harris-Haller, L. W., Stoebner, R. a, & Manson, M. D. (1995). Motility protein interactions in the bacterial flagellar motor. *Proceedings of the National Academy of Sciences of the United States of America*, *92*(6), 1970–4.
60. Blair, D. A. F. B. (1998). Electrostatic interactions between rotor and stator in the bacterial flagellar motor, *95*(May), 6436–6441.
61. Braun, T. F., Al-Mawsawi, L. Q., Kojima, S. & Blair, D. F. (2004). Arrangement of core membrane segments in the MotA/MotB proton- channel complex of Escherichia coli. Biochemistry 43, 35–45.
62. Sato, K. & Homma, M. (2000). Functional reconstitution of the Na+-driven polar flagellar motor component of Vibrio alginolyticus. J Biol Chem 275, 5718–5722.
63. Roujeinikova, A. (2008). Crystal structure of the cell wall anchor domain of MotB, a stator component of the bacterial flagellar motor: implications for peptidoglycan recognition. *Proc. Natl. Acad. Sci. U.S.A.*, *105*(30).
64. Paul, K., Gonzalez-Bonet, G., Bilwes, A. M., Crane, B. R., & Blair, D. (2011). Architecture of the flagellar rotor. *The EMBO Journal*, *30*(14), 2962–71.
65. Delalez, N. J., Berry, R. M., & Armitage, J. P. (2014). Stoichiometry and Turnover of the Bacterial Flagellar Switch Protein FliN. *Asm*.
66. Sircar, R., Borbat, P. P., Lynch, M. J., Bhatnagar, J., Beyersdorf, M. S., Halkides, C. J.,Crane, B. R. (2014). Assembly States of FliM and FliG within the Flagellar Switch Complex. *Journal of Molecular Biology*, (2014).
67. Delalez, N. J., Wadhams, G. H., Rosser, G., Xue, Q., Brown, M. T., Dobbie, I. M., Armitage, J. P. (2010). Signal-dependent turnover of the bacterial flagellar switch protein FliM. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(25), 11347–51.
68. Mandadapu, K. K., Nirody, J. a, Berry, R. M., & Oster, G. (2014). Mechanics of torque generation in the bacterial flagellar motor, PNAS 1–18.
69. Dean, K. M., & Palmer, A. E. (2014). Advances in fluorescence labeling strategies for dynamic cellular imaging. *Nature Chemical Biology*, *10*(7), 512–23.
70. Snapp, E. L. (2009). Fluorescent proteins: a cell biologist's user guide. *Trends in Cell Biology*, *19*(11), 649–655. Journal Article.
71. Agrawal, U., Reilly, D. T., & Schroeder, C. M. (2013). Zooming in on biological processes with fluorescence nanoscopy. *Current Opinion in Biotechnology*, *24*(4), 646–53.

72. Chandler, J.H. (2006) Characterization of the stator complexes of the *Escherichia coli* flagellar motor (Doctoral dissertation). Oxford University.
73. Greenfield, D., McEvoy, A. L., Shroff, H., Crooks, G. E., Wingreen, N. S., Betzig, E., & Liphardt, J. (2009). Self-organization of the Escherichia coli chemotaxis network imaged with super-resolution light microscopy. *PLoS Biology*, *7*(6), e1000137.
74. Beeby, M., Ribardo, D. A., Brennan, C. A., Ruby, E. G., & Jensen, G. J. (2016). Diverse high-torque bacterial flagellar motors assemble wider stator rings using a conserved protein scaffold.
75. Cordes FS, Bright JN, Sansom MSP (2002) Proline-induced distortions of transmembrane helices. J Mol Biol 323(5):951–960.
76. Tipping, M. J., Delalez, N. J., Lim, R., Berry, R. M., & Armitage, P. (2013). Load-Dependent Assembly of the Bacterial Flagellar Motor.
77. Che, Y., Nakamura, S., Morimoto, Y. V, Kami-ike, N., Namba, K., & Minamino, T. (2013). Load-sensitive coupling of proton translocation and torque generation in the bacterial flagellar motor, 1–35.
78. Lele, P. P., Hosu, B. G., & Berg, H. C. (2013). Dynamics of mechanosensing in the bacterial flagellar motor. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(29), 11839–11844. Journal Article.
79. Tipping, M. J., Steel, B. C., Delalez, N. J., Berry, R. M., & Armitage, J. P. (2013). Quantification of flagellar motor stator dynamics through in vivo proton-motive force control. *Molecular Microbiology*, *87*(2), 338–347. Journal Article.
80. Morimoto, Y. V, Nakamura, S., Kami-ike, N., Namba, K., & Minamino, T. (2010). Charged residues in the cytoplasmic loop of MotA are required for stator assembly into the bacterial flagellar motor. *Molecular Microbiology*, *78*(5), 1117–29.
81. Palmer, E., & Freeman, T. (2004). Investigation into the use of C- and N-terminal GFP fusion proteins for subcellular localization studies using reverse transfection microarrays. *Comparative and Functional Genomics*, *5*(4), 342–353.
82. Leake, M. C., Chandler, J. H., Wadhams, G. H., Bai, F., Berry, R. M., & Armitage, J. P. (2006). Stoichiometry and turnover in single, functioning membrane protein complexes. *Nature*, *443*(7109), 355–8.
83. Block, S., & Berg, H. (1984). Successive incorporation of force-generating units in the bacterial rotary motor. Nature.
84. Blair, D F & Berg, H C. (1988). Restoration of torque in defective flagellar motors. science
85. Fung, D. C., & Berg, H. C. (1995). Powering the flagellar motor of Escherichia coli with an external voltage source.
86. Ryu, W. S., Berry, R. M., & Berg, H. C. (2000). Torque-generating units of the flagellar motor of Escherichia coli have a high duty ratio. Nature
87. Sowa, Y., Rowe, A. D., Leake, M. C., Yakushi, T., Homma, M., Ishijima, A., & Berry, R. M. (2005). Direct observation of steps in rotation of the bacterial flagellar motor. Nature
88. Reid, S. W., Leake, M. C., Chandler, J. H., Lo, C.-J., Armitage, J. P., & Berry, R. M. (2006). The maximum number of torque-generating units in the flagellar motor of Escherichia coli is at least 11. Proceedings of the National Academy of Sciences of the United States of America
89. Inoue, Y., Lo, C.-J., Fukuoka, H., Takahashi, H., Sowa, Y., Pilizota, T., Ishijima, A. (2008). Torque-speed relationships of Na+-driven chimeric flagellar motors in Escherichia coli. Journal of Molecular Biology
90. Yuan, J., & Berg, H. C. (2008). Resurrection of the flagellar rotary motor near zero load. Proceedings of the National Academy of Sciences of the United States of America
91. Meacci, G., & Tu, Y. (2009). Dynamics of the bacterial flagellar motor with multiple stators. Proceedings of the National Academy of Sciences of the United States of America
92. Sowa, Y., Homma, M., Ishijima, A., & Berry, R. M. (2014). Hybrid-fuel bacterial flagellar motors in Escherichia coli.

93. Castillo, D. J., Nakamura, S., Morimoto, Y. V., Che, Y.-S., Kami-ike, N., Kudo, S., … Namba, K. (2013). The C-terminal periplasmic domain of MotB is responsible for load-dependent control of the number of stators of the bacterial flagellar motor. *Biophysics*, *9*, 173–181.
94. Paulick, A., Koerdt, A., Lassak, J., Huntley, S., Wilms, I., Narberhaus, F., & Thormann, K. M. (2009). Two different stator systems drive a single polar flagellum in Shewanella oneidensis MR-1. Molecular Microbiology, 71(4), 836–850.
95. Berg, H. C., & Tedesco, P. M. (1975). Transient response to chemotactic stimuli in Escherichia coli. *Proceedings of the National Academy of Sciences*, *72*(8), 3235–3239.
96. Block, S. M., Segall, J. E., & Berg, H. C. (1982). Impulse responses in Bacterial Chemotaxis. *Cell*, *31*(November), 215–226.
97. Paul, K., Brunstetter, D., Titen, S., & Blair, D. F. (2011). A molecular mechanism of direction switching in the flagellar motor of Escherichia coli. Proceedings of the National Academy of Sciences of the United States of America, 108(41), 17171–6.
98. Bai, F., Branch, R. W., Nicolau, D. V, Pilizota, T., Steel, B. C., Maini, P. K., & Berry, R. M. (2010). Conformational spread as a mechanism for cooperativity in the bacterial flagellar switch. *Science (New York, N.Y.)*, *327*(5966), 685–9.
99. Bai, F., Minamino, T., Wu, Z., Namba, K., & Xing, J. (2012). Coupling between switching regulation and torque generation in bacterial flagellar motor. *Physical Review Letters*, *108*(17), 1–4.
100. Togashi, F., Yamaguchi, S., Kihara, M., Aizawa, S. I., & 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. *Journal of Bacteriology*, *179*(9), 2994–3003.
101. Lee, L. K., Ginsburg, M. a, Crovace, C., Donohoe, M., & Stock, D. (2010). Structure of the torque ring of the flagellar motor and the molecular basis for rotational switching. Nature, 466(7309), 996–1000.
102. Lam, K.-H., Ip, W.-S., Lam, Y.-W., Chan, S.-O., Ling, T. K.-W., & Au, S. W.-N. (2012). Multiple conformations of the FliG C-terminal domain provide insight into flagellar motor switching. Structure (London, England: 1993), 20(2), 315–25.
103. Brown, P. N., Hill, C. P., & Blair, D. F. (2002). Crystal structure of the middle and C-terminal domains of the flagellar rotor protein FliG. *The EMBO Journal*, *21*(13), 3225–34.
104. Vartanian, A. S., Paz, A., Fortgang, E. a, Abramson, J., & Dahlquist, F. W. (2012). Structure of flagellar motor proteins in complex allows for insights into motor structure and switching. *The Journal of Biological Chemistry*, *287*(43), 35779–83.
105. Way, S. M. Van, Millas, S. G., Lee, A. H., & Manson, M. D. (2004). Rusty, Jammed, and Well-Oiled Hinges: Mutations Affecting the Interdomain Region of FliG, a Rotor Element of the Escherichia coli Flagellar Motor, 186(10), 3173–3181.
106. Karen A. Fahrner, Ryu, W. S., & Berg, H. C. (2003). Bacterial flagellar switching under load. Nature, 423.
107. Yuan, J., Fahrner, K. a., & Berg, H. C. (2009). Switching of the Bacterial Flagellar Motor Near Zero Load. Journal of Molecular Biology, 390(3), 394–400.
108. Wang, F., Yuan, J., & Berg, H. C. (2014). Switching dynamics of the bacterial flagellar motor near zero load. Proceedings of the National Academy of Sciences.
109. Lloyd, S. A. & Blair, D. F. (1997). Charged residues of the rotor protein FliG essential for torque generation in the flagellar motor of Escherichia coli. J. Mol. Biol. 266, 733–744
110. Nakamura, S., Kami-ike, N., Yokota, J. P., Minamino, T., & Namba, K. (2010). Evidence for symmetry in the elementary process of bidirectional torque generation by the bacterial flagellar motor. PNAS.
111. Pandini, A., Morcos, F., Khan, S., Ackbarow, T., Chen, X., Keten, S., Blair, D. F. (2016). The Gearbox of the Bacterial Flagellar Motor Switch. *Structure*, *0*(0), 16410–16415.
112. Fan Bai (2006). Mathematical Modeling of the Bacterial Flagellar Motor System. Oxford University. PhD thesis.
113. Duke, T. A. J., Le Novère, N. & Bray, D. (2001). Conformational spread in a ring of proteins: a stochastic approach to allostery. J. Mol. Biol. 308,541–553.

114. van Albada, S. B., Tănase-Nicola, S., & ten Wolde, P. R. (2009). The switching dynamics of the bacterial flagellar motor. *Molecular Systems Biology*, *5*(316), 316.
115. E.M. Purcell. 1977. Life at low Raynold number. AJP. 45. 3-11
116. Arber, W. (2008). Molecular mechanisms driving Darwinian evolution. *Mathematical and Computer Modelling*, 47(7–8), 666–674.
117. Rorick, M. (2012). Quantifying protein modularity and evolvability: a comparison of different techniques. Biosystems, 110(1), 22–33.
118. Finnigan, G. C., Hanson-Smith, V., Stevens, T. H., & Thornton, J. W. (2012). Evolution of increased complexity in a molecular machine. Nature, 481(7381), 360–4.
119. Watson.R., (2002). Compositional Evolution: Interdisciplinary Investigations in Evolvability, Modularity, and Symbiosis. PhD thesis.
120. Eörs Szathmáry, Ferenc Jordán, Csaba Pál (2001) Can Genes Explain Biological Complexity? Science. 1315-1316
121. Lynch, M. (2007). The frailty of adaptive hypotheses for the origins of organismal complexity. Proceedings of the National Academy of Sciences of the United States of America, 104 Suppl (Table 1), 8597–604.
122. Gray, M., Lukes, J., & Archibald, J. (2010). Irremediable complexity? Science, 920(2010), 9–10.
123. Olson-Manning, C. F., Wagner, M. R., & Mitchell-Olds, T. (2012). Adaptive evolution: evaluating empirical support for theoretical predictions. Nat Rev Genet, 13(1938), 867–877.
124. Perica, T., Marsh, J. A., Sousa, F. L., Natan, E., Colwell, L. J., Ahnert, S. E., & Teichmann, S. A. (2012). The emergence of protein complexes: quaternary structure, dynamics and allostery. Biochemical Society Transactions, 40, 475–491.
125. Lorenz, D. M., Jeng, A., & Deem, M. W. (2011). The emergence of modularity in biological systems. Physics of Life Reviews, 8(2), 129–160.
126. Watson, R and Pollack, J. (2005). Modular Interdependency in Complex Dynamical Systems. Artificial Life, 11, 4, 445-457.
127. Omer, S., Kovacs, A., Mazor, Y., & Gophna, U. (2010). Integration of a foreign gene into a native complex does not impair fitness in an experimental model of lateral gene transfer. Molecular Biology and Evolution, 27(11), 2441–2445.
128. Lind, P. A., Tobin, C., Berg, O. G., Kurland, C. G., & Andersson, D. I. (2010). Compensatory gene amplification restores fitness after inter-species gene replacements. Molecular Microbiology, 75(5), 1078–1089.
129. Paulick, A., Koerdt, A., Lassak, J., Huntley, S., Wilms, I., Narberhaus, F., & Thormann, K. M. (2009). Two different stator systems drive a single polar flagellum in Shewanella oneidensis MR-1. Molecular Microbiology, 71(4), 836–850.
130. Bubendorfer, S., Held, S., Windel, N., Paulick, A., Klingl, A., & Thormann, K. M. (2012). Specificity of motor components in the dual flagellar system of Shewanella putrefaciens CN-32. Molecular Microbiology, 83(2), 335–50.
131. Mccarter, L. L. (2005). Multiple Modes of Motility: a Second Flagellar System in Escherichia coli, 187(4), 10–13.
132. Lele, P. P., & Berg, H. C. (2015). Switching of Bacterial Flagellar Motors Is Triggered by Mutant FliG. Biophysical Journal, 108(5), 1275–1280.
133. Marsh, J. a., & Teichmann, S. a. (2015). Structure, Dynamics, Assembly, and Evolution of Protein Complexes. Annual Review of Biochemistry, 84(1), 551–575.

Chapter 2. Experimental Methods and Materials

2.1 Cells and cultures

2.1.1 *E.coli* Strains and plasmids

Strains and plasmids used in this study are shown in table 2.1, and the cBFM plasmids used for single motor tethered cell assay are shown in table 2.2. The entire list of the cBFM strains can be found in table 5.1. Briefly, RP437 derivative ΔMotAB, sticky filaments strain named JPA605, carrying pBAD33 plasmid vectors encoding MotA, FP-MotB (or MotB for wildtype) were used for most of the bead assays and tethered cell assays. For control experiments, strains that express MotA and MotB from the native promoter, RP437, MT02, MT03 and JPA804 were used. Strain MG1655 ΔMotAB was used for the cBFM chemotaxis assay, and strain RP437 ΔMotAB (JPA604) was used fusion stators chemotaxis experiments.

Table 2.1. Bacterial strains and plasmids used in this study

***E. coli* strains**	**Description**	**Source or reference**
RP437	derivatives of *Escherichia coli* K-12 strain (commonly used for *E.coli* motility assays)	Parkinson JS (1978)
MG1655	*Escherichia coli* K-12 strain	From Bertus Beaumont lab
MG1655 ΔMotAB	MG1655, ΔMotAB	From Bertus Beaumont lab
MT02	RP437, fliCst variant	A gift from Richard Berry lab
MT03	RP437, fliCst variant, ΔCheY	A gift from Richard Berry lab
JPA804	RP437, fliCst variant, 28-eGFP-MotB	A gift from Richard Berry lab
JPA605	RP437, fliCst variant, ΔMotAB	A gift from Richard Berry lab
JPA604	RP437, ΔMotAB	A gift from Richard Berry lab
JHC36	RP437, fliCst variant, ΔMotAB, ΔCheY	A gift from Richard Berry lab
Plasmids		
pBAD33	pBAD ara promoter, Cmr	ATCC® 87402™
pTrc99a	Tcr promoter, Amp	pTrc99a-rat FABPI, addgene 13577
pMMB206	Tac/lacUV5 promoter, Cmr	ATCC® 37808™
pFX40	Lac promoter, Amp	From C.Dekker lab, TU Delft
pRmHA3 3xFlag-eGFP	eGFP gene	Obtained from Anna Haagsma
pYPet-His	YPet gene	Addgene 14031
Tol2-mpx-Dendra2	Dendra2	Addgene 29574
pBAD33_MotAB	MotAB in pBAD33	From Bertus Beaumont lab
pMH_01	MotA, 28AA of MotB-eGFP-MotB in pBAD33	This study
pMH_02	MotA, eGFP-MotB in pBAD33	This study
pMH_03	MotA, Ypet-MotB in pBAD33	This study

pMH_04	MotA, Ypet-GGGGS-MotB in pBAD33	This study
pMH_05	MotA, Ypet-EAAAK-MotB in pBAD33	This study
pMH_06	MotA, Dendra2-MotB in pBAD33	This study
pMH_07	MotA, Dendra2-GGGGS-MotB in pBAD33	This study
pMH_08	MotA, Dendra2-GSGSGS-MotB in pBAD33	This study
pMH_09	MotA, Dendra2-3xGGGGS-MotB in pBAD33	This study
pMH_10	MotA, Dendra2-EAAAK-MotB in pBAD33	This study
pMH_11	MotA, Dendra2-3xEAAAK-MotB in pBAD33	This study
pMH_12	MotA, -6AA_Dendra2[1]-MotB in pBAD33	This study
pMH_13	MotA, eDendra2[2]-MotB in pBAD33	This study
pMH_14	MotAB in pTrc99a	This study
pMH_15	MotA, Dendra2-MotB in pTrc99a	This study
pMH_16	MotA, Ypet-MotB in pTrc99a	This study
pMH_17	MotAB in pMMB206	This study
pMH_18	MotA, Denra2-MotB in pMMB206	This study
pMH_19	MotA, Ypet-MotB in pMMB206	This study
pMH_20	MotAB in pFX40[3]	This study
pMH_21	MotA, Ypet-B in pFX40	This study
pMH_22	MotA, Ypet-GGGGS-MotB in pFX40	This study
pMH_23	MotA, Ypet-EAAAK-MotB in pFX40	This study
pMH_24	MotA, Dendra2-B in pFX40	This study
pMH_25	MotA, Ypet-EAAAKx3-MotB in pBAD33	This study

1. Dendra2 protein with the last six amino acids (C-terminus) deleted.
2. First 6 and last 7 AA sequence of eGFP were replaced to the first and last Dendra2 sequence. This is an optimized Dendra2 for bacterial cells, engineered in [58].
3. This is a pFX40 vector backbone without the entire 5266 bp gene inserts (YFP-MinD, MinE and LacZ and LacY).
   *Abbreviations: Cmr, chloramphenicol resistant; Amp, ampicillin resistant.

Table 2.2 cBFM-P plasmids used for single motor analysis

Plasmids	**Description**	**Bacterial species of MotAB gene**
cBFM8	Laf TU in pBAD33	Escherichia coli O111:H- str. 11128
cBFM8.1a	A134S Laf T and laf U in pBAD33	
cBFM8.1c	A134S & L284R Laf T and Laf U in pBAD33	
cBFM8.1d	A134S & E87K Laf T and Laf U in pBAD33	
cBFM11	MotPS in pBAD33	Bacillus pseudofirmus OF4
CBFM11.1i	MotP and Q52P MotS in pBAD33	
cBFM17	MotAB in pBAD33	Listeria monocytogenes EGD-e
CBFM17.1a	MotA and S39I MotB in pBAD33	
CBFM17.1b	MotA and F53V MotB in pBAD33	
CBFM17.1f	MotA and S39R MotB in pBAD33	
CBFM17.1h	MotA and F53V Mot B in pBAD33	

CBFM17.1j	18bp insertion MotA and Mot B in pBAD33	
CBFM17.1k	9bp deletion MotB	
CBFM17.1o	15bp deletion MotB	
CBFM18	MotAB putative in pBAD33	Bacillus megaterium DSM319
CBFM18.1e	Q237K MotA and MotB in pBAD33	

2.1.2 Bacterial cells glycerol stocks preparation

Strains were grown to stationary phase in LB-broth containing appropriate antibiotic(s) at 37 C (~ 16 hours) with shaking. The stationary phase cells were mixed with sterile 50 % glycerol in a 1: 1 ratio and stored at – 80 ºC freezer.

2.1.3 Media, buffers and chemical stocks preparation

The following media, buffers, antibiotics and gene inducer stocks were prepared and used for cell cultures and assays.

Tryptone Broth (1 L)

10 g Tryptone enzymatic digest from casein (95039 Sigma-Aldrich)
5g of sodium chloride
adjust pH to 0.7 with 5N NaOH
Deionized / Milli-Q water

Terrific broth (T5574 Sigma)

23.6g/L Yeast Extract
11.8g/L Tryptone
9.4g/L K2HPO4
2.2g/L KH2PO4
4ml/L Glycerol

LB media

10g tryptone
5g yeast extract
10 g sodium chloride

Antibiotics and gene inducers

Ampicillin 100 ug/ml (in distilled water)
Chloramphenicol 34 ug/ml (in 95 % ethanol)
10% w/w L-Arabinose (A3256 Sigma) (i.e.1g of Arabinose to 9ml of milliQ water)

500 mM IPTG (I6758 Sigma)
Both stocked were filter sterilized with a 0.22µm filter and stored at -20 C.

2.1.4 Cell's growth rate (OD 600) measurements

Growth rates were estimated by their optical density (OD) at a wavelength of 600 nm, which indicates the concentration of the bacterial cells in a liquid media. 96 well plates were used to measure the cell growth rates in the automated multi-mode microplate reader Cytation 3 (biotek). Each well contained 200 µl of either LB media or Tryptone broth with appropriate antibiotics and inducers as indicated. 1 µl of glycerol frozen stock cell samples were inoculated in each well, grown at 33 °C with orbital shaking. OD 600 were measured in every 20 minutes for 14~16 hours. Each strain was analyzed in triplicate to minimized the mis-pipetting errors.

2.2 Molecular biology genetic engineering

2.2.1 Plasmid vectors collection

The desired plasmid vector expression system had to meet the following conditions: low, tightly regulated but fast expression. The plasmid vectors collected are listed in Table 2.3. pBAD33 vector is well suited for the tightly controlled low expression of the gene [1]. In addition, three other vectors with IPTG inducible promoters were collected to see if they could express the genes faster than pBAD33 vectors. The four vector systems were used to express the fusion stator proteins (figure 2.1). pFX40 have a lac promoter and a medium copy number (5~10 copies per cell). This vector was collected by the C. Dekker's lab in TU Delft since they used this vector to express fusion proteins in *E.coli* [2]. pTrc99a and pMMB206 were obtained as they were used by previous BFM studies including resurrection experiments [3-4]. The sources of these vectors are listed in Table 2.1.

Table 2.3 The list of the plasmid vector expression systems collected

Plasmid vector	**Promoter**	**Regulator**	**Origin of replication**
pBAD33	Ara promoter	AraC	pACYC184/p15A
pFX40	Lac promoter	LacIq	pBR322 derivative
pTrc99a	Trc promoter	LacI	ColE1 origin
pMMB206	Tac/lacUV5	lacIq	OriV/T (IncQ)

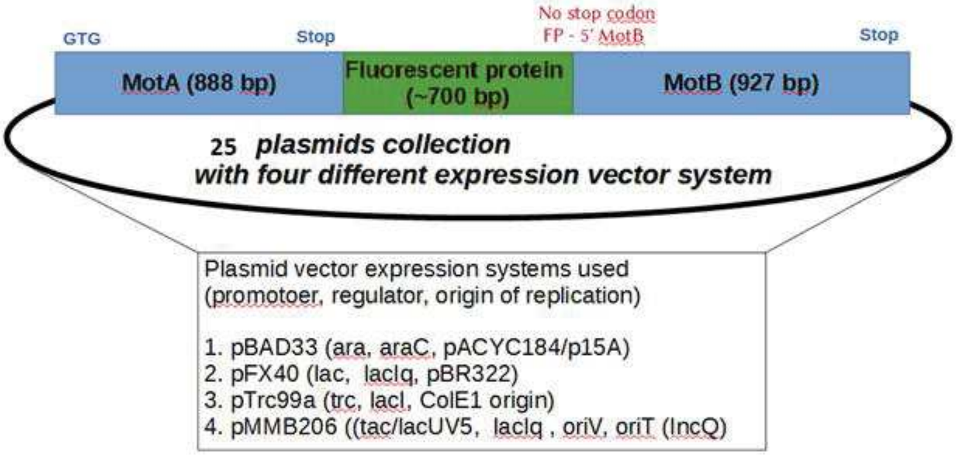


Figure 2.1 A simplified illustration of the plasmid maps with the MotA and FP-MotB gene inserts.

2.2.2 Primer design

For the standard PCR reaction, 18~25 bp of 5’ DNA sequences and the reverse complement sequence of 18~25 bp of 3’ DNA sequences were used as forward and reverse primers. For the Gibson reactions, the primers were designed to contain 25~ 40 bp overlapping linear DNA fragments. The list of PCR primers used are shown in Table 2.4.

Table 2.4 PCR primers

primer	**sequence (5' …. 3')**	**length(bp)**	**remark**
222	ggt agc ATG AAG AAT CAA GCG CAT CCG ATT ATT GTC	48	MotB initial 84bp fragment forward primer
223	GTC ATA AAG TCG GCA TAA GCA ATC TTC CAC	39	MotB initial 84bp fragment reverse primer (from 104)
224	ATG GTG AGC AAG GGC GAG GAG CTG TTC ACC	39	eGFP PCR Forward primer
225	CGC TTG ATT CTT CAT gct acc CTT GTA CAG CTC GTC CAT GCC GAG	60	eGFP PCR Reverse primer
226	GCTTATGCCGACTTTATGAC	20	for vector right arm (eGFP-MotB)
227	TCCTCGCCCTTGCTCACCATAATCTTCCACGATCCATGTG	40	for vector left arm (eGFP-MotB)
228	ATG AAG AAT CAA GCG CAT CCG	27	MotB initial 84bp fragment forward primer
229	CGG ATG CGC TTG ATT CTT CAT CTT GTA CAG CTC GTC CAT GCC GAG	59	eGFP PCR Reverse primer (no linker)
293	ATGAACACCCCGGGAATTAAC	21	Dendra 2 PCR_Forward
294	CCACACCTGGCTGGGCAG	18	Dendra 2 PCR_Reverse
295	ATGTCTAAAGGTGAAGAA	18	Ypet PCR_Forward
296	TTTGTACAATTCATTCATAC	20	Ypet PCR_Reverse
297	CCAGCCAGGTGTGGAAGAATCAAGCGCATC	30	MotB-Dendra2 PCR_Forward
298	GAATGAATTGTACAAAAAGAATCAAGCGCATCCG	34	MotB-Ypet PCR_Forward
300	ACGAATTCAAGGAGATATACATGTGCTTATCTTATTAGGT	40	MotA-Lac Promoter PCR_forward
301	TTAATTCCCGGGGTGTTCATTCATGCTTCCTCGGTTGT	38	MotA-Dendra2 PCR_Reverse
302	AATTCTTCACCTTTAGACATTCATGCTTCCTCGGTTGT	38	MotA-Ypet PCR_Reverse
303	GGTGGTgcatgcAGGCACCCCAGGCTTTACAC	32	Lac promoter PCR (SphI) Forward
304	ATGTATATCTCCTTGAATTCGTAATCATGG	30	Lac promoter PCR Reverse
305	GGTGGTAAGCTTTCACCTCGGTTCGGCT	28	MotB PCR_reverse (HindIII)
312	TCATGCTTCCTCGGTTGTCG	20	MotA End reverse
313	GGTGGTGCATGCAGGCAC	18	Lac promoter (forward) shorter primer
1	ACAACCGAGGAAGCATGAATGAACACCCCGGGAATTAAC	39	Dendra 2_Forward with MotA overlap
2	TAATCGGATGCGCTTGATTCTTCCACACCTGGCTGGGCAG	40	Dendra 2 PCR_Reverse with MotB overlap
3	ACAACCGAGGAAGCATGAATGTCTAAAGGTGAAGAA	36	Ypet PCR_Forward with MotA overlap
4	TAATCGGATGCGCTTGATTCTTTTTGTACAATTCATTCATAC	42	Ypet PCR_Reverse with MotB overlap
5	CCATGATTACGAATTCAAGGAGATATACATGTGCTTA TCTTATTAGGTTACCTG		MotA Forward (EcoRI-pFX40)
8	TTCGCTCGCGTATCGGTGATTCACCTCGGTTCGGCTGATG	40	pFX40 MotB reverse (for MotAB PCR)

9	CATCAGCCGAACCGAGGTGAATCACCGATACGCGAGCGAA	40	pFX40 MotB Forward (for pFX PCR)
12	GAATTGTACAAAggtggcggtggcagcAAGAATCAAGCGCATCCG		MotB-Ypet fusion GGGGS linker Forward primer for Gibson rxn
13	gctgccaccgccaccTTTGTACAATTCATTCATACCC		MotB-Ypet fusion GGGGS linker Reverse primer for Gibson rxn
14	GAATTGTACAAAgaagcggcagctaagAAGAATCAAGCGCATCCG		MotB-Ypet fusion EAAAK linker Forward primer for Gibson rxn
15	cttagctgccgcttcTTTGTACAATTCATTCATACCC		MotB-Ypet fusion EAAAK linker Reverse primer for Gibson rxn
16	GAATTCGAGCTCGGTACCCGAAGGATGATGTCGTGC TTATCTTATTAGGTTACC		SacI tag pcr primer for motA Forward (only for pBAD)
17	ACAGCCAAGCTTGCATGCCTGCTCACCTCGGTTCGGCTGATG		SphI tag pcr primer for motB Reverse
18	GGTGGTGAATTCGTGCTTATCTTATTAGGTTAC		EcoRI tag pcr primer for motA Forward
21	ACAGACGACAACCGAGGAAGCATGA ATGGTGAGCAA GGGCGAGGA		egfp with 25 bp motA end (for gibson) Forward
22	CAGGGCGCTTACTGGCTCAT		end part of motB reverse
23	TTGATTCTTCATgctaccgctaccgctaccCCACACCTGGCT GGGCAGGG		GSGSGS linker Dendra2 bind (reverse)
24	AGCCAGGTGTGG ggt agc ggt agc ggt agcATGAAGAATC AAGCGCATCC		GSGSGS linker MotB bind (forward)
25	TTGATTCTTCAT ctt agc tgc cgc ttc CCACACCTGG CTGGGCAGGG		EAAAK linker Dendra2 bind (reverse)
26	CAGGTGTGG gaa gcg gca gct aag ATGAAGAATCAAG CGCATCC		EAAAK linker MotB bind (forward)
27	GATGCGCTTGATTCTTCAT GGGGCTGTAGCGGGCCACGG		Dendra2 binds (without last 6AA) reverse
28	CCGTGGCCCGCTACAGCCCC ATGAAGAATCAAGCGCATC		MotB binds (Dendra2 without last 6AA) forward
29	ATGTCTAAAGGTGAAGAA CTGATCAAGGAGGACATGCGCG		Dendra2 forward with first 6 AA of eGFP seq
30	CTTGTACAATTCATCCATACC CAGGGGGCTGTAGCGGGCCACG		Dendra2 reverse with last 7 AA of eGFP seq
31	ACAGCCCCCTGGGTATGGATGAATTGTACAAG ATGAAGAATCAAGCGCATC		MotB bind (with Dendra2 and GFP ending)
32	ACAACAGACGACAACCGAGGAAGCATGA ATGAACACCCCGGGAATTAAC		Dendra 2_Forward with MotA overlap
33	CTCCTTGATCAGTTCTTCACCTTTAGACAT TCATGCTTCCTCGGTTGTCG		MotA binds reverse (with Dendra2 and GFP overhang)

2.2.3 Polymerase chain reaction (PCR) and overlap extension PCR

Polymerase chain reactions (PCR) were carried out using either Phusion® High-Fidelity DNA Polymerase (NEB) or Q5 High-Fidelity DNA Polymerase (NEB) according to their instructions. Both high fidelity polymerase PCR kits share the same protocol except the buffer and the polymerase. Overlap-extension PCR was also performed using Phusion polymerase with two DNA templates to link, instead of one DNA template. In this case, one set of each primer from the two DNA templates were used (i.e., the N-terminal primer of the first DNA template, and the C-terminal primer of the second DNA template). Multiple trouble shooting reactions were performed to obtain optimized PCR conditions.

Thermo-cycling conditions
step 1 Initial denaturation 94 ºC, 2 min
step 2 Denaturation 92 ºC, 10 seconds
step 3 Annealing 55~68 ºC[1], 10~30 seconds
step 4 Extension 72 ºC, x min[2]
step 5 to step 2 (30 cycles)
step 6 Final extension 72 ºC, 2 min

1. Annealing temperature can be determined by the Tm of primers and the sequence of the binding sites.
2. Extension time can be determined by the length of DNA sequence to be amplified (30 seconds per kb)

2.2.4 DNA electrophoresis

To check the size of the DNA fragments followed by PCR reaction or DNA restriction enzyme cut reactions, DNA electrophoresis was performed. Agarose gel was prepared at concentrations between 0.8 and 1.2 % (w/v) (depending upon the size of DNA bands expected) in 0.5 X TBE buffer (90 mM Tris-borate, 2 mM EDTA, pH 8.0). Ethidium bromide was added in the agarose gel to avoid gel staining procedure after the DNA electrophoresis. DNA fragments were separated according to their sizes while electrophoresed at between 80 and 145 V. Gels were visualized on a UV trans-illuminator and imaged using Gel Doc imager (Bio-Rad).

2.2.5 DNA purification

PCR product DNA fragments were purified using NucleoSpin Gel and PCR clean up kit by Macherey-Nagel, according to their instructions. Small-scale plasmids DNA purification (mini-prep) was carried out using NucleoSpin plasmid kit (Macherey-Nagel), according to their instructions. In brief, 3 ml of stationary phase culture was centrifuged and resuspended in resuspension buffer. Lysis buffer was added and left for 1 min to ensure full lysis reaction. Neutralization buffer was then added and the mixture was centrifuged for 5 min (~11,000 x g). The supernatant was decanted and the plasmid DNA was purified using silica spin column. Plasmid DNA was eluted in 30~70 µl elution buffer and stored at - 20 ºC until required.

2.2.6 Restriction digests

Restriction digests were carried out either to create the desired DNA fragments with given sticky endings for the T4 ligation reaction or to check success of the ligation reactions. DNA restriction digest reaction was carried out using appropriate restriction enzyme(s) (New England Biolabs), the desired DNA templates, the reaction buffer provided, 2 - 10 units of enzyme and BSA (if necessary). The 20 - 50 µl volumes of total reaction mixture was incubated for 2 h at 37 ºC, unless the instructions stated otherwise.

2.2.7 DNA ligation by T4 ligase

After the restriction digested DNA fragments with matching overhangs were generated, the DNA products were purified by column purification. Gel purification was avoided since the agarose gel substance may hinder the ligation reactions. Vector and insert DNA templates in a ratio of about 1: 3 or 1: 5 (total amount of DNA present ~1µg) were ligated in a reaction volume of 20 µl containing 2 µl T4 DNA ligase (NEB), supplied reaction buffer, and left at 16 ºC for overnight (~16 h). This ligation DNA solution was subsequently transformed into high efficiency competent cells.

2.2.8 Gibson recombination reactions

To generate the DNA vector and insert templates for the Gibson recombination reaction, PCR was performed with the primers carrying 20 ~ 40 overhang sequences. The primers were designed such that the final products can generate a minimum of 25 bp overlap sequences (ideally over 35 bp sequences) between the two or three DNA templates. See the table 2.4 for the list of Gibson reaction primers. The PCR products were purified by column purification. Gel purification was avoided since the agarose gel substance may hinder the ligation reactions. Gibson assembly master mix (NEB) was used and the reaction mixture was prepared according to their instruction. Vector and insert(s) DNA products were mixed in a ratio of 1:3 ~ 1:10 (total amount of DNA present ~ 1 µg) in a total reaction volume of 20 µl, containing 10 µl Gibson reaction master mix (NEB). Incubated in a thermocycler at 50°C for 1 hour, followed by (high efficiency) competent cell transformation with 2~5 ul of the Gibson reaction mixture.

2.2.9 Competent cells and *E.coli* transformation

Competent cells for the motilities assays were prepared using the following calcium-chloride method. 5 ml of desired *e. coli* strain was grown overnight at 37° C in LB to stationary phase. The overnight culture was diluted in 1:100 in fresh LB medium and grown at 37 °C until O.D. (600 nm) 0.4-0.6 was reached. The cells were harvested by centrifugation, resuspended carefully in 100 mM $CaCl_2$ and incubated on ice for 20 minutes. The last step was repeated twice, including a longer 1-hour incubation on ice, and small aliquots with 15 % glycerol were prepared and stored at -80 °C. For the T4 ligation or Gibson assembly reaction mixture, commercially available high efficiency competent cells NEB5 alpha or NEB10 beta were used instead.

For the heat shock transformation procedure, Ca^{2+} competent cells were thawed on ice, about 1~5ng of super-colied purified plasmids (or 10~50 ng for the ligation mix) were added and incubated on ice for 30 minutes. Cells were heat shocked at 42 °C for 1 minute, followed by 15 min on ice. Either 1ml of LB media or SOC media was added to the heat-shocked cells and incubated at 37 °C for 1 hour to allow cells to recover and express antibiotic resistance genes. (Using SOC media generated a much higher yield of colonies the next day). 200 µl of the mixture was spread onto antibiotics selective LB agar plate, then incubated at 37 °C overnight.

2.3 Bioinformatics

2.3.1 DNA sequencing

DNA was sequenced by Beckman Coulter Genomics Sequencing Service using Sanger DNA sequencing method. DNA sequences were analyzed using the NCBI blast sequence alignment and SnapGene software package. Sequencing plasmid DNA samples were supplied at a concentration of 100 ng / µl (5 µl /reaction) and primers at 5 µM (2 µl /reaction). The sequencing primers used are listed in Table 2.5.

Table 2.5 sequencing primers

primer	**sequence (5' …. 3')**	**length(bp)**	**remark**
77F	CTACCTGACGCTTTTTATCGC	21	Binds ara promoter pBAD33 forward primer
211	CGGCAACGATGCGCTTAAGCG	21	Binds end region of MotB forward primer
212	CAGCTTGCCGGTGGTGCAGATG	22	Binds beginning region of EGFP reverse primer
213	GCAGCTCGCCGACCACTACCAGC	23	Binds end region of EGFP forward primer
220	GCGTTTGGTATTGTTGCGGCTG	22	Binds mid region of MotA forward primer
230	ACGCCGAAAGCCAGAATGAG	20	Binds end region of MotB forward primer
232	GAACTTCAAGATCCGCCACA	20	Binds mid-end region of eGFP forward primer
233	TGATTACACCCAAAGCCAGG	20	Binds mid region of MotB forward primer
251	GCATGAAAACGCCGAAAGCCAG	22	Binds end region of MotB forward primer
312	TCATGCTTCCTCGGTTGTCG	20	Binds end region of MotA reverse primer
10	GCAGCTGGCACGACAGGTTTCC	22	pFX40 forward sequencing (before lac promoter)
11	GGTTCTAAACATCGGGCGATTC	22	Binding mid region of MotB reverse primer
19	AAGCGCACCCAGCTCGGCGG	20	Binds mid region of MotA reverse primer
20	CTCAAAATCGATCTGGTCCAGG	22	Binds mid region of MotB forward primer
22	CAGGGCGCTTACTGGCTCAT	20	Binds end region of MotB reverse primer

2.3.2 Sequencing and structural alignments

DNA and protein sequences of cBFM MotAB and fluorescent proteins (eGFP, Ypet and Dendra2) were aligned using multiple sequence alignment package T-coffee [5] by combining the alignment methods of Clustal, Probcons and Muscle into one unique alignment. T-Coffee can align Protein (Expresso), DNA and RNA sequences. The 3D structural alignments were done via PBD analysis tools available online at www.rcsb.org.

2.4 *E.coli* motility assays

2.4.1 Chemotaxis soft-agar plate assay

The population motility was measured by observing chemotaxis on soft agar plates. As cells metabolize nutrients in the agar, they create a gradient of chemicals and, if chemotactic, swim from the point of inoculation up the attractant gradient to the regions of higher nutrient concentration, forming a large ring. The chemotaxis assay was performed in half fold diluted LB media with appropriate antibiotics and gene inducers. The percentage of agar was strictly set to 0.3% (or less) to ensure swimming motility rather than swarming motility [6-7]. The 0.3 % agar plates were prepared in Bertus Beaumont lab for cBFM chemotaxis assays. The 0.3% agar plate for cBFM assays was composed of:

½ x LB agar (400 ml)
2 g peptone from casein, tryptic digest (Fluka 500 gr cat no:70172)
1 g yeast extract (Fluka 100 gr cat no: 70161)
2 g NaCl
1.2 g agar (0.3 %)

autoclave: 20 min 120 °C, after cooling (~50 °C)
400 µL chloramphenicol antibiotics + 4 mL 20% arabinose (0.2 %)

For the chemotaxis assays of FP-MotB fusion stators, either 0.223 or 0.25 % agar plates were prepared instead of 0.3 % agar plate (0.3% agar plates generated much harder agar plates than the ones prepared in Beaumont lab in TU Delft due to the difference in autoclave procedure in CBS, Montpellier). 27 mL of solution were added to each plate and they were dried overnight with closed lid. 5 µl aliquots of stationary phase cultures (in LB-broth, at 37 ºC) were inoculated onto the soft agar plates, with no more than three samples per plate. To prevent any evaporation, the soft agar plates were placed in boxes (three plates per box) and incubated at 37 ºC for 8~12 h, depending upon the strains being assessed. The chemotaxis speeds were measured as the distance of the forefront of the ring from the inoculation point (or diameter of the ring) after a given incubation time. Each set was repeated in triplicate.

2.4.2 Single cell swimming assay in liquid media

Swimming behavior of cBFM-P *E.coli* cells in liquid media were monitored in a bovine serum albumin (BSA) tunnel slide. The BSA tunnel slide (~100 µm height) was prepared with a microscope slide, a same size cover slide and a double sided sticky tape (figure 2.2a). Up to seven stick strips (~ 4mm x ~25mm size strips) were placed on the microscope slide, then a cover slide on top with an overhang of about 1-2 mm with respect to the microscope slide to finish flow channels. About 20 ul of 0.5% BSA solution (5 mg/ml) was added to the flow channels and dried overnight to prevent the cell bodies sticking to the surface of the slide.

E. coli was grown in 5ml of half LB + CAM (25µg/ml) at 37 °C with shaking at 250 rpm until stationary phase (overnight culture). The overnight culture was diluted 1:10 in fresh 5 ml of half LB + CAM (25µg/ml) + 0.2 % arabinose media and grown further at 37 °C for 2 hours (± 15min) to induce the stator gene expression (OD600 of ~ 0.5 or less). The culture was diluted tenfold in the same fresh half LB media. About 10~15 µl of the diluted cell culture was added to the each BSA tunnel slide. Both ends of the slides were sealed with vacuum silicone. The cells were rested in the slide for at least two minutes before recording so that most cells can swim right above the glass surface. The swimming behaviors were monitored and recorded under the dark field microscope at room temperature (~23 °C) for 30 seconds each. Three independent experiments were performed for each cBFM strain.

2.4.3 Single motor - tethered cell body and rotating bead assay

Cell preparation was the same for both tethered cell body and rotating bead assays. Cells were grown at 33 °C in 3 ml tryptone broth (Sigma Aldrich) with the appropriate antibiotics and inducers until OD 600 of 0.55~0.65. For the cBFM tethered cell assay, cells were grown at 37 °C instead. About 0.7 ml of cell suspension was passed through two syringes with 26-gauge needles (26 µm) connected by a piece of polyethylene tubing (~10 cm long, 0.58 mm inner diameter) at least 50 times. This procedure is performed on sticky FliC mutant strains (JPA605, JHC36, MT02, MT03) to shear the long filaments of *E.coli* flagellum to achieve shorter flagellum, which will be either attached to the bead or to the slide. Shortened flagella can ensure that a spinning bead is close to the cell body and does not attach itself to neighboring cells or flagella. The sheared cell suspension was collected by centrifuging at 3,615 g (3,000 rpm) for two minutes. The cell pellet was resuspended in ~400 µl of motility buffer (10 mM potassium phosphate / 0.1 mM EDTA/ 10 mM lactic acid, pH 7.0).

For tethered cell assay, the prepared cell sample in motility buffer was added to the custom-made flow chamber (Figure 2.2b), allowing the cells to rest and the sticky flagellar stubs to be attach to the glass surface by hydrophobic interaction. The flowing cells can be washed by the fresh motility buffer and Kim-wipe tissue. For bead assay, the glass cover slide of the flow channel (Figure 2.2c) was coated with poly L-lysine (PLL) before introducing the cell sample. Poly L lysine allow cells stuck onto the surface of the channel slide due to their charge interactions (positively charged PPL and negatively charged cell body). After the PLL coating, the cells were introduced and left settle for few minutes, then 1.1 µm diameter polystyrene latex beads (sigma) were attached to the filament stubs (just by introducing the beads in the flow channel), followed by washing with the same motility buffer. The rotation of the bead or rotation of the cell body were monitored for 3~5 min each, in a laser bright-field setup [59], recorded by CMOS fast camera at frame per second (FPS) of 100, 300 and 500 depending on the experiment. All experiments were carried out at room temperature 22 °C.

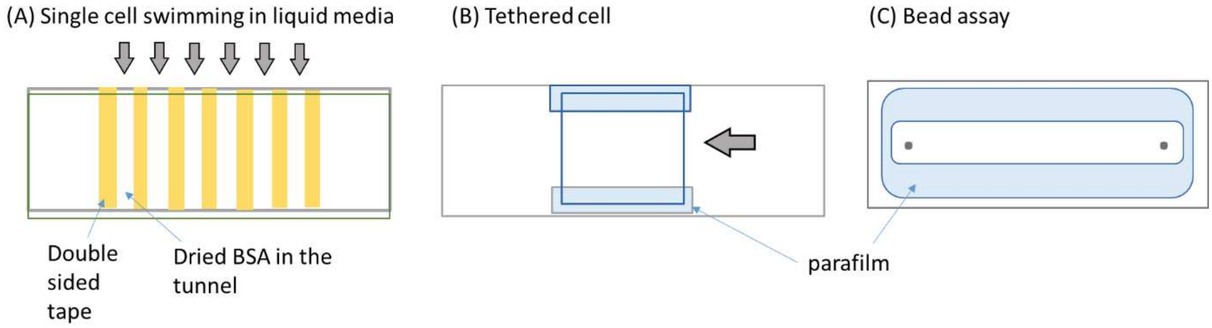
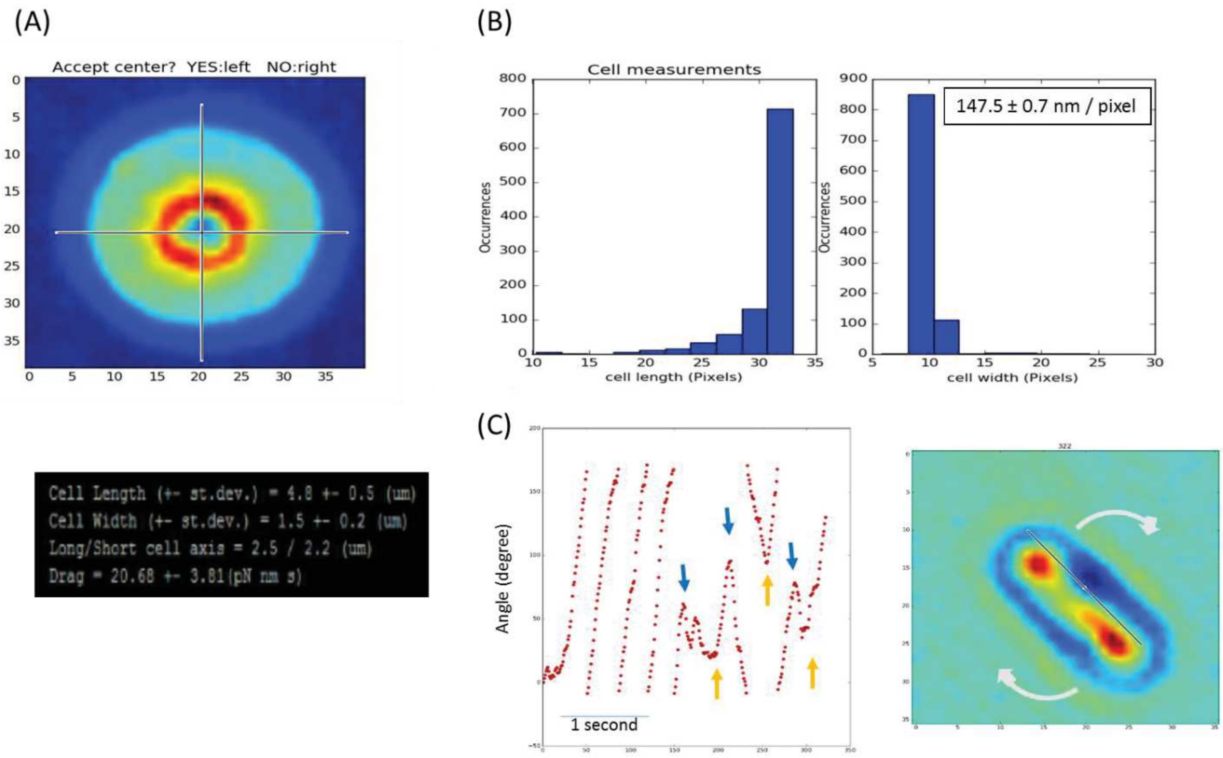


Figure 2.2 Three tunnel slides for (A) swimming assay, for (B) tethered cell assay and (C) for bead assay.

2.5 Single motor data analysis

2.5.1 Tethered cell data analysis

The recorded videos were cropped around the rotating cell bodies to reduce the file sizes. This step can reduce a file size of ~ 1 GB to 200~300 MB. The video includes stacks of the images of the rotating cell body (100 images per second, FPS=100), and the center of rotation is where the flagellar is tethered on the glass surface. The center of the rotation was identified by the standard deviation image of the recorded video (figure 2.3a). The center of the rotation was identified either by automatic 'center detection' or by a manual click on the image. The drag of the rotating cell body varied based on the size of the cell body and the position of the working motor in the cell body. The tethered cell drag was calculated based on the cell length, the distance between the center of the cell and the rotational axis, length and the width of the cell (figure 2.3b). The equations to calculate these frictional drag coefficients can be found in SI of [8]. The angle of the cell body from the center of the rotation were tracked (figure 2.3c). From this tracked angle verse time, the rotation speed and the switching events of the cell body were measured. The torque produced by the motor was calculated by the calculated drag value and the rotation speed using the software written by Dr. Pedaci (python).



```
Cell Length (+- st.dev.) = 4.8 +- 0.5 (um)
Cell Width (+- st.dev.) = 1.5 +- 0.2 (um)
Long/Short cell axis = 2.5 / 2.2 (um)
Drag = 20.68 +- 3.81(pN nm s)
```

Figures 2.3 A representative image of a standard deviation of rotating cell body (A) and a distribution of measured cell length and width of this cell (B) are shown. The cell length, cell width and the long/short cell axis for this particular cell are shown in the bottom left corner. A single frame of this rotating cell body image and its angle vs time trace are shown (C). Angle and time trace and a snap shot of the corresponding rotation cell. Switching events can be seen when the tracked angle change from positive to negative. Blue arrows represent switching events from CCW to CW, and orange arrows represent switching events from CW to CCW (FPS=100).

Switching of the tethered cell was defined as when the cell switches its direction of rotation for more than 90 degree (angle) with a minimum speed of 0.1 Hz. Consequently, the switching threshold was set as 2/3 of CCW mean rotation speed (down threshold) and a fixed rotation speed of 0.1 (up threshold). The speed that passes both thresholds is counted as a switch. The angle of rotation per frame can be changed by the rotation speed; when the cell is rotating at a low speed (<2 Hz), high number of frames are required to check whether they rotated in opposite direction for more than 90 degrees. So the filter set was adjusted accordingly. The filter is defined as replacing Y number of raw data points into a single data point with the median of neighboring raw data points (median filter), where Y being the filter set value. The rotation of the cell body was recorded as a movie with a 100 frame per second (FPS), which gives 100 tracked angles of the rotation (raw data points) per second, and this angle of the rotation per frame can be measured by an equation: (360 degrees * X Hz) / 100 FPS, where X is the speed of the rotation (Hz). Thus, the filter set was decided by the minimum number of frames needed to turn 90 degrees, which is 90 degrees divided by the angle of rotation per frame.

2.5.2 Bead assay data analysis

The flagellar motor rotation was measured by a 1.1 μm polystyrene bead on the sticky flagellar stub of the BFM (Figure 2.4_a). The center position of the bead on the rotating stub was tracked using a high-speed camera, and their time trace of the X and Y positions were plotted (Figure 2.4c). Figure 2.5 shows the bead position in x and y plane separately, and x y axis together with lines connecting each entry. In order to ensure that the video of the rotating beads was not effected by other extraneous factors, the bead traces with a right circle or a fine ellipse were selected. Normally the rotating beads were recorded for at least 3 minutes. Since this is a long enough time to cause any small drift of the stage, where the cells on the coverslip lies, a neighboring stuck bead were also recorded as a reference bead. This reference bead can tell us the stage movements during the experiment. To correct this artifact, each entry of the center position of the rotating bead was subtracted by the center position of the reference bead. The corrected time trace of the x and y positions of the bead is shown in figure 2.4d. This drift correction made the trajectory of the bead more like a single circle without much scattered dots around. Then, the algorithm fits an ellipse into this bead trajectory using the least squares fitting [9-10] (Figure 2.4e). This fitted ellipse bead trajectory was used to obtain the angle (turn) time trace, and subsequently, speed time trace as shown in figure 2.6. The sampling frequencies (FPS) of the bead assays was 300 Hz, which allows ~3.3 ms of time resolution.

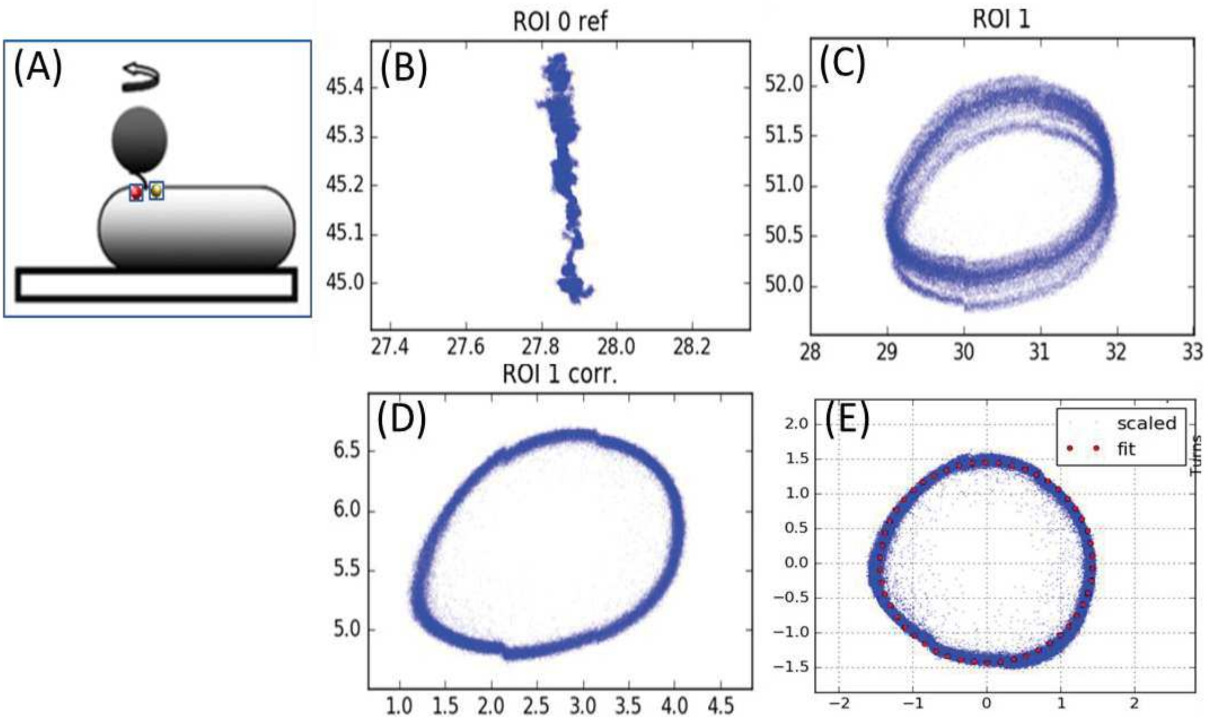


Figure 2.4 (A) A schematic image of a rotating bead attached to the flagellar stub in *E.coli* cell. (B) A reference bead stuck on the glass slide trajectory in a x-y plane. This bead drifted in Y-axis. (C) A rotating bead trajectory in a x-y plane before drift correction. (D) A rotating bead trajectory in a x-y plane after drift correction. The trajectory is ellipse (instead of circle) because the location of the flagellum that the bead was attached to was not in perfectly perpendicular. (E) A drift corrected rotating bead trajectory in a x-y plane after an ellipse fit (a fitted ellipse to the entries). The blue dots represent all the x-y plane entries of the time trace. The angle time trace and speed time trace of the corresponding bead are shown in figure 2.6.

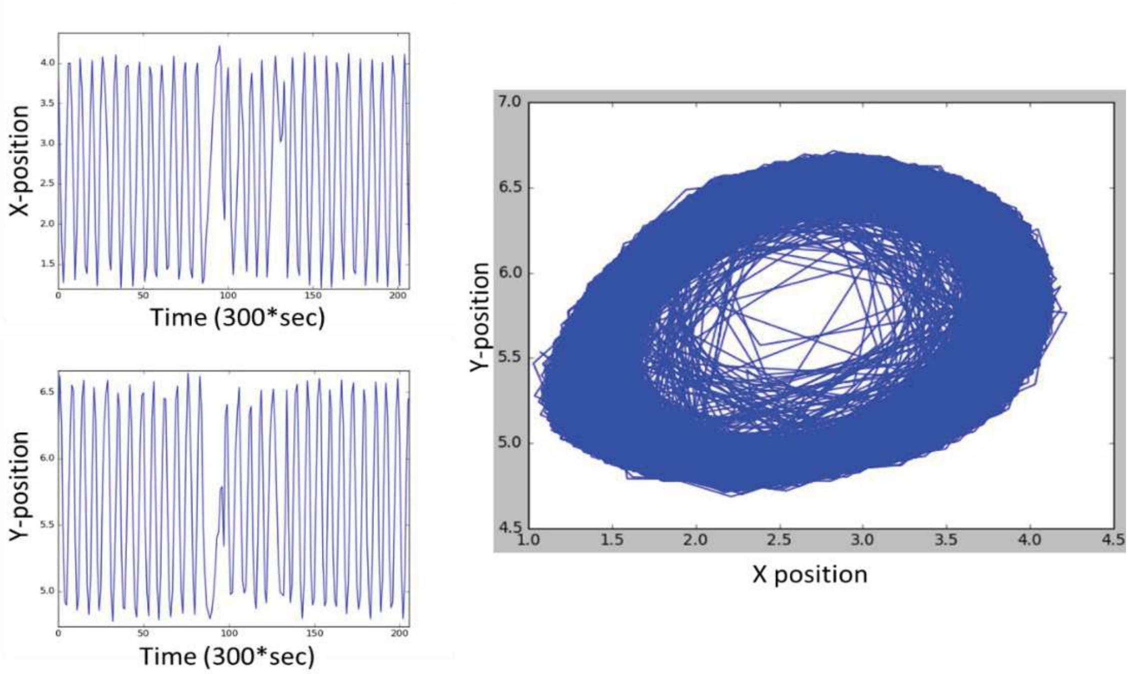


Figure 2.5 The X- and Y- axis time traces corresponding to the bead in figure 2.4 is visualized (Left). When these two x-y plane time trace is plotted together, it creates a circle as shown in the right. All the entries (dots) are connected with a line.

The angle time trace shows a simple linear regression relationship between the angle and the time. This reveals the fact that this motor rotated mostly in CCW bias rotational direction in a steady speed. When the angle time trace is zoomed in, this linear regression relationship can be more like a multiple linear regression if there were switching events (figure 2.6a). Here, switching event was defined as when the rotation speed crosses 2/3 of both mean positive speed and mean negative speed. The switching threshold for positive speed (up_th) and negative speed (dw_th) can show a great difference if the mean rotation speeds in positive and negative speed varies a lot, like seen by fusion stator motors. Therefore, to count the number of switching events in the speed-time trace, the up switching threshold was set to 2/3 of the mean positive rotation speeds (red arrow in figure 2.6b) and the down switching threshold was set to 2/3 of the mean negative rotation speed (green arrow in figure 2.6b). When the speed time trace is zoomed in, it shows rapid switching events from CCW to CW and CW to CCW.

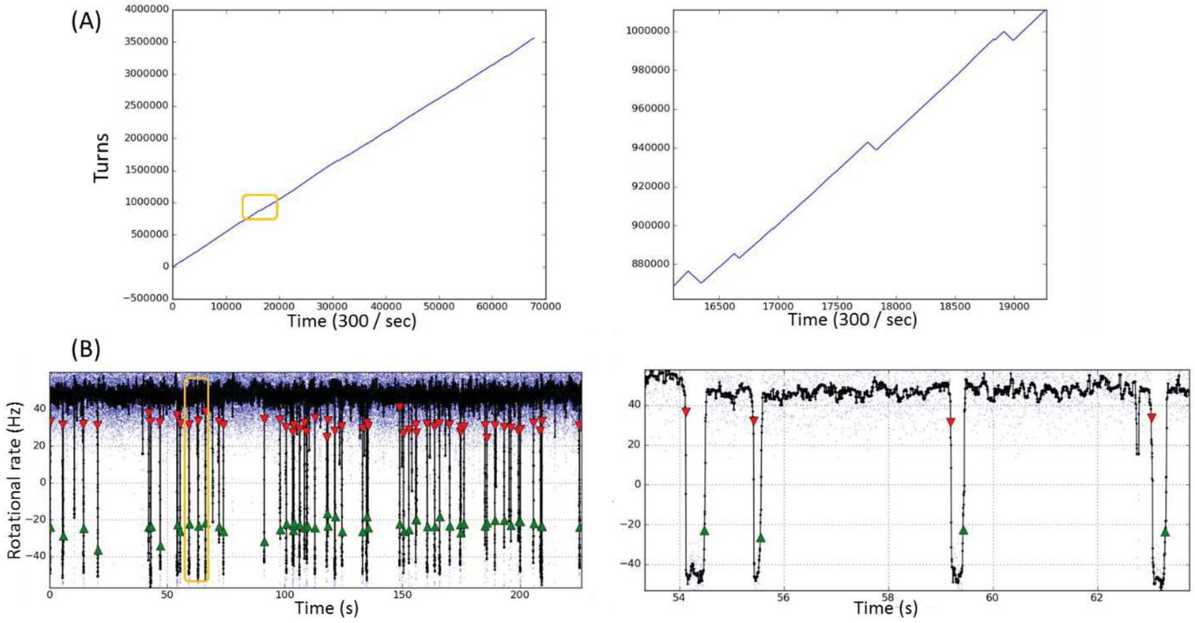


Figure 2.6 Switching events seen from angle-time trace (A) and speed time trace (B). The switching events were counted by setting a switching threshold (2/3 of mean rotation speed). The red and green arrows in the speed time trace represents up and down switching threshold, respectively.

The measured speed-time traces are intrinsically noisy (possibly due to the thermal noise influencing the position of the bead). To reduce the noise, we applied a 21-points (approximately 0.07 seconds) window median filter, which assigns to each point in the trace the median of the 21 points around it. A filter of 21 points was used for most of the cells. Figure 2.7 shows the effects of the filter in speed time traces. For the cells with a rotation speed of lower than 20 Hz, higher filter set was used to minimize higher noise level and to ensure counting all the switching events.

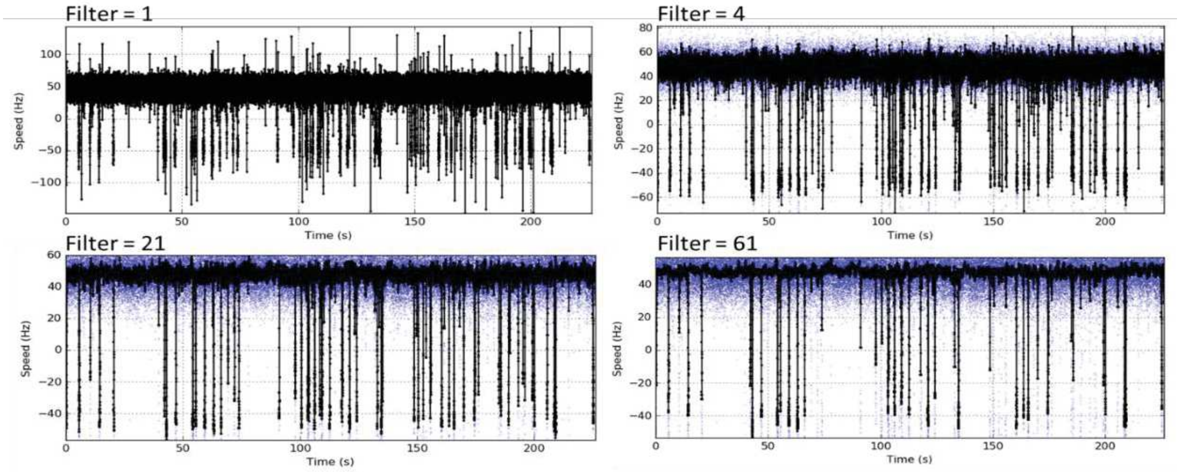


Figure 2.7 Effect of the filter window over the same speed time trace. Blue points: raw data, dark: filtered trace.

2.6 Single molecule fluorescence imaging in live *E.coli* cells

2.6.1 Cleaning the microscope slide

The coverslips (Menzel glaser) were rinsed with acetone, ethanol and filtered water, sonicated in 1M KOH for 20 min, rinsed again with water, dried quickly under air flow, passed through a flame, and then stored dry until use. This cleaned slide showed reduced amount of visible background fluorescence on the glass slide when excited by 552 nm laser. However, this cleaning procedure was not able to completely get rid of all the background fluorescence. Interestingly, the background fluorescence on the glass slides were excited mostly by 552 nm laser but not so much by 488 nm laser. So, the glass slide without the cleaning procedure above were used for all the fluorescent proteins excited by 488 nm laser.

2.6.2 Epi and TIRF fluorescence microscope

An inverted custom built fluorescence microscope (developed in Pedaci lab) was used. The microscope contains multiple illumination paths for 650nm (red), 552 nm (yellow-green), 488nm (blue), 405nm (violet) laser lights and 100 x 1.45 NA oil objective (Nikon). Both epi- and TIRF- (total internal reflection) fluorescence imaging was performed in this microscope. In a wide field epi-fluorescence imaging mode, the light travels through the objective lens to illuminate the sample, and then the light emitted from the sample travels back through the same objective to the detector (figure 2.8). In TIRF mode, a tilted excitation generates an evanescent field at the cell sample at the interface between two media having different refractive indices, which is the contact area between a specimen and a glass coverslip [11, 12]. Thus, TIRF imaging is advantageous for membrane fluorescent protein imaging. The switch between Epi and TIRF was facilitated by the dichroic mirror, mounted just before the excitation light passes through the objective (figure 2.8). The position of the dichroic mirror was adjusted to control the incident illumination angle, such that the light can pass through the objective either on axis or with the critical angle for total internal reflection. Since this adjustment was done manually, the TIRF images taken in the same day are comparable, and the direct comparison of the ones taken in different days was avoided.

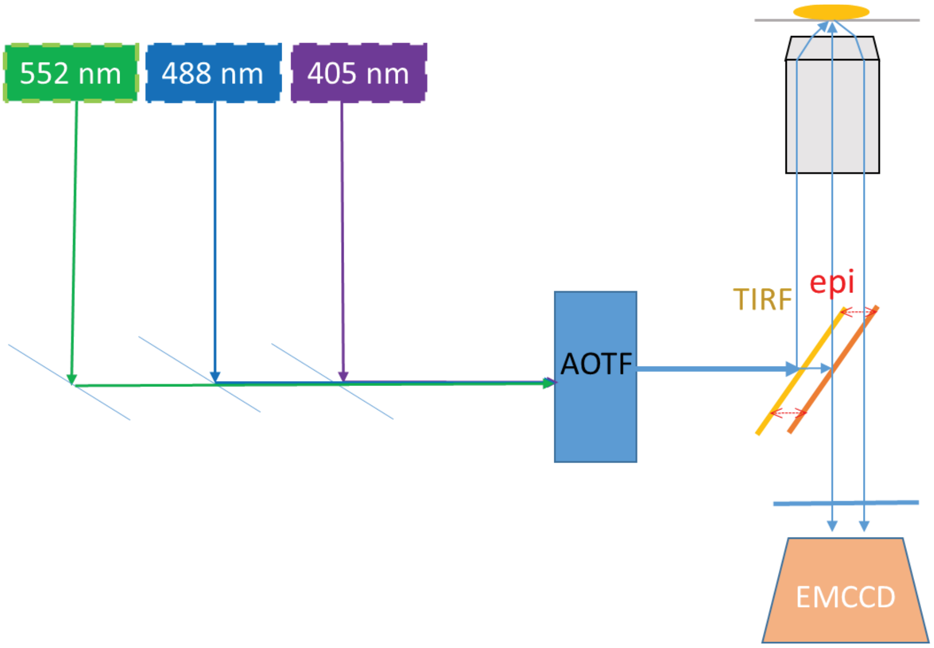


Figure 2.8 A simplified epi- and TIRF- fluorescence microscope. The dichroic mirror was moved to change the excitation illumination angle to the objective (TIFP < > epi).

650nm laser light was used for the low intensity bright-field illumination, therefore used for bead assay. 488nm laser light was used to excite eGFP, Ypet and unconverted Dendra2 fluorescent proteins. The emission light passed through a 495 nm dichroic mirror (T495lpxr, Chroma) and a 525 nm emission filter (band width = 50 nm). 405nm and 552nm laser lights were used to photo-convert and excite Dendra2, respectively, and the emission light passed through a 560 nm dichroic mirror (T560lpxr, Chroma) and a 590 nm filter (band width = 50 nm). In order to record the fluorescence signal and the rotating bead of the same motor of the *E.coli* cell, a dual recording set up was built (figure 2.9). The additional dichroic mirror (T647lpxr) splits the FP emission light and the red light scattered by the beads. The emission filter (ET700/50 nm) in front of the CMOS camera (Optronis CL600x2M) was mounted (in tilted way) to filters out most of the unnecessary background signals and efficiently record the bead rotation. An iris can be placed in between the two dichroic mirrors to localize a small region of the interest where rotating beads are.

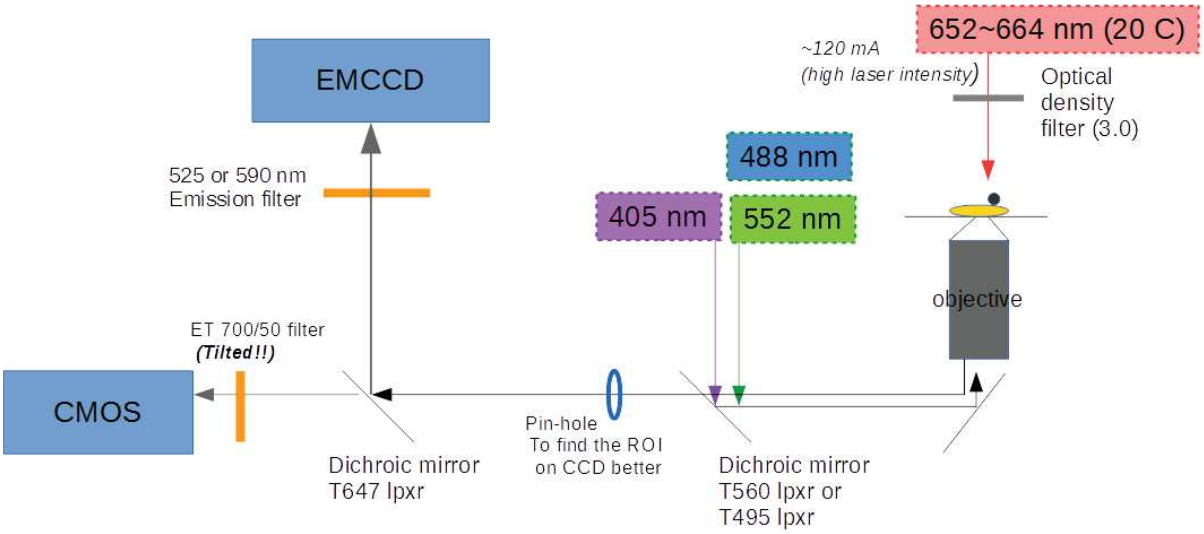


Figure 2.9 A schematic drawing of the dual recording set up. EMCCD camera can record the FP fluorescent signals from the cell samples while CMOS camera can record rotating beads or rotating cell bodies, simultaneously. The emission filter and the dichroic mirror are depending on which FP to image.

Cell sample preparation for the fluorescence imaging was the same as described in 2.4.3 (single motor bead assay and tethered cell assay). Especially, it was convenient to use the tethered cell assay flow slide since it contains both stuck cells and tethered cells. Stuck cells can be used to visualize the fluorescence intensity of the whole cell body or membrane and the tethered cells can be used to visualize the fluorescence intensity of the functioning motor as the sheared sticky filaments were tethered to the bottom window of the flow chamber. The cell sample slide was placed onto a slide holder which was then screwed into a piezoelectric nanostage on a z-nano positioning stage (PI). Total illumination area was ~5,435 μm^2 and the illumination intensity of the 488 nm laser at the sample was ~ 7 mW in epi mode and ~ 0.3 mW in semi TIRF mode (when the laser intensity setting was 20 mW). Each pixel size of the images (512x512) from the EMCCD camera was 144 ± 1 nm and from the CMOS camera was 148 ± 1 nm. All experiments were performed at 20 °C.

2.6.3 Image acquisition and data analysis

Images (512 x 512 pixels) were acquired using a back-illuminated cooled (-80 °C) Electron Multiplying Charge Coupled Device (EMCCD) camera (iXon Ultra 897, Andor). An acousto-optic tunable filter (AOTF) was used to control and select a specific wavelength from the laser sources. The image acquisition was controlled by a custom written software (Labview 8, National Instruments). EM (electron multiplying) gain and the exposure time were varied, but, normally EM gain of 50 ~ 132 and exposure time of 100~500 ms (in frame transfer mode) were used for the unconverted three FPs. The image acquisition for the photo-converted Dendra2 is discussed in chapter 5.4. Images from the EMCDD camera were recorded in .tif format and imported into Python or ImageJ for further analysis. The details of image analysis will be discussed in Chapter 4.

References

1. Guzman, L. M., Belin, D., Carson, M. J., Beckwith, J., Guzman, L., Belin, D., & Carson, M. J. (1995). Tight regulation , modulation , and high-level expression by vectors containing the arabinose PBAD promoter, 177(14).
2. Wu, F., Rijn, E. Van, Schie, B. G. C. Van, Keymer, J. E., & Dekker, C. (2015). Multi-color imaging of the bacterial nucleoid and division proteins with, 6(June), 1–15.
3. Muramoto, K., & Macnab, R. M. (1998). Deletion analysis of MotA and MotB, components of the force-generating unit in the flagellar motor of Salmonella. Molecular Microbiology, 29(5), 1191–202.
4. Sowa, Y., Rowe, A. D., Leake, M. C., Yakushi, T., Homma, M., Ishijima, A., & Berry, R. M. (2005). Direct observation of steps in rotation of the bacterial flagellar motor. Nature, 437(7060), 916–9.
5. Ridgeway, T., & Nw, L. (2000). T-Coffee : A Novel Method for Fast and Accurate Multiple Sequence Alignment.
6. Kearns, D. B. (2010). A field guide to bacterial swarming motility. Nature Reviews. Microbiology, 8(9), 634–44
7. Croze, O. A., Ferguson, G. P., Cates, M. E., & Poon, W. C. K. (2011). Migration of chemotactic bacteria in soft agar: Role of gel concentration. Biophysical Journal, 101(3), 525–534.
8. Che, Y. S., Nakamura, S., Kojima, S., Kami-ike, N., Namba, K., & Minamino, T. (2008). Suppressor analysis of the MotB(D33E) mutation to probe bacterial flagellar motor dynamics coupled with proton translocation. Journal of Bacteriology, 190(20), 6660–6667.
9. D. Umbach and K. N. Jones, 2003. A few methods for fitting circles to data, IEEE Transactions on Instrumentation and Measurement, vol. 52, no. 6, pp. 1881–1885.
10. A. Fitzgibbon, M. Pilu, and R. B. Fisher, Direct least square fitting of ellipses, IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 21, no. 5, pp. 476–480, 1999.
11. http://www.nikon-instruments.jp/eng/page/products/wtirf.aspx
12. http://www.microscopyu.com/articles/fluorescence/tirf/tirfintro.html

Chapter 3. Characterizing the flagellar motors driven by fluorescent protein tagged stators

3.1 Rationale and aims

Stator proteins have been labeled by a fluorescent protein (FP), mainly GFP, by several groups to monitor their localization and understand their dynamics owing to their central role in torque generation [1-6]. Labeling the stators by a FP is the most commonly used method to visualize them, because they are in the inner membrane of the cell and have no accessible part from the extracellular side, making standard immuno-fluorescence or other membrane protein labeling methods unfeasible. According to the putative topology of MotA and MotB, both the N-terminus and C-terminus of MotA and the N-terminus of MotB are in the cytoplasmic side. The C-terminus of MotB is in the periplasm, where the stator complexes are anchored. When a FP is attached to MotA both to the N-terminus and the C-terminus, the motor becomes non-functional and no fluorescent spots are visible [7]. When a FP is attached to the N-terminus of MotB, however, the motor is functional. Hence, the FP-MotB is used by many studies to monitor stators by fluorescence. It has been reported that, however, while the motors with the FP-MotB stator were functional, their motility is reduced [1,4]. It is only speculated that the FP tag on MotB interferes somehow with the power-stroke mechanism, but no studies are available to understand how the FP tag on the stator affects the torque generation as well as the switching mechanism.

In this study, MotB fusion proteins with the three different fluorescent proteins (eGFP, YPet, Dendra2) were constructed, initially, to learn more about the stator dynamics in the functioning motor. Our functional characterization of the motors carrying the FP tags consistently showed unusual behaviors, mainly related to switching. Considering the proximity of the fluorescent tags to the stator-rotor interface, this impaired function of the motor is not surprising. Therefore, the goal of the project was re-oriented and started to investigate the detailed activities of the FP tagged motors. Using single molecule biophysics approaches, the following aims and questions were addressed.

Aims of this chapter

1. Quantitatively analyze population chemotaxis and single motor speed, and the effects of the FPs on them.
2. Identify whether the reduced motility caused by FPs tag is due to the less number of stators bound to the motor or whether each stator produces less torque.
3. Identify how the various aspects of the switching activities are affected by the FP tag.
4. Examine if the impaired functions of the motor can be restored by having a linker between FP-MotB.
5. Explore possible explanations for how different FP tags (even though they were similar in size, structure and location) cause different effects on the motor.
6. Propose possible explanations that could explain the impaired motor functions caused by the FP tag on the stators.

3.2 FP-MotB fusion stator constructions

Enhanced GFP (eGFP) is a mutant variant of wild type GFP [8]. eGFP is the most widely used, well characterized, stable, bright, monomer and previously used FP for *E.coli* motility proteins. Hence, eGFP-MotB fusion construct was tried at first. The construct previously made by other groups presents - a N-terminus fragment (~28 codons) of motB which was added in front of eGFP-, following the hypothesis that the N-terminus of motB could contain a putative membrane protein trafficking motif sequence [7]. Accordingly, we constructed the 28 AA-eGFP-MotB fusion protein by linking the first 28 codons sequence of motB to eGFP, followed by a ligation into pBAD33_motA and motB vector (Figure 3.1). The vector with this FP-MotB fusion protein was named pBAD33_MotA, 28AA-eGFP-MotB (or pMH_01).

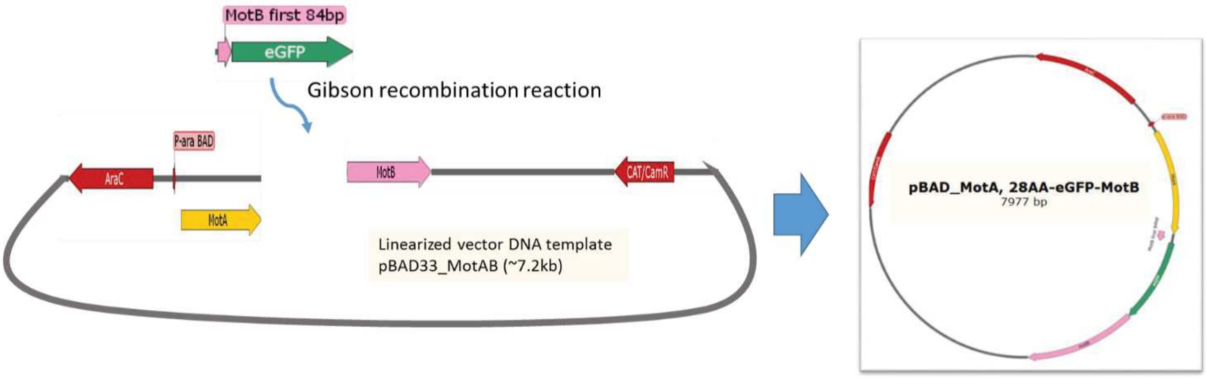


Figure 3.1. 28AA-eGFP-MotB fusion protein construction. The N-terminus 84 bp of MotB sequence was fused to eGFP using overlap extension PCR. Then, this fragment was ligated into the linear pBAD_MotAB vector DNA by Gibson recombination reaction. The resulting construct is called pBAD_MotA, 28AA-eGFP-MotB (or pMH_01).

In addition, eGFP-MotB fusion stator, which do not have the additional 28 AA sequence of motB in front of eGFP, was constructed to test whether the additional 28 AA were necessary. In addition to the eGFP fusion stator, two more fluorescent proteins – YPet and Dendra2- were fused to the N-terminus of motB. YPet is a Venus mutant variant of YFP, made for FRET study paring with CyPet [9]. Dendra2 is a monomeric green to red photo-convertible protein, derived from octocoral *Dendronephthya sp* [10, 11]. YPet was chosen since it is supposed to be almost twice as bright as eGFP [12]. Dendra2 was chosen since it is a photo-convertible FP which can be used for PALM super resolution microscopy. All the three FPs, including unconverted form of Dendra2, can be excited at 488 nm and emit around 510 nm, making it especially well-suited for detection by commonly used microscopes. It is reported that eGFP and Dendra2 show reasonable folding efficiencies and fast maturation rates at 37 °C in both bacteria and mammalian cells [12-13]. However, YPet folding efficiency and maturation rate are less clearly known. YPet shows a good photostability and a better resistance to acidic environments than other yellow FPs [14], though eGFP and Dendra2 provide higher photostability. In total, four FP-MotB fusion proteins with the three different fluorescent proteins were constructed (Table 3.1). All the constructs did not have any linker sequences between the FP and motB.

Table 3.1. Photophysical properties of the FPs fused to MotB [10~14]

MotB-fusions	**Excitation (nm)**	**Emission (nm)**	**Brightness$^{(a)}$**	**Photostability$^{(b)}$**	**pKa**	**Oligomerization**
eGFP	488	507	100	174	6.0	monomer or weak dimer
YPet	517	530	238	49	5.6	monomer or weak dimer
Dendra2 (Green)	490	507	67	-	7.5	monomer
Dendra2 (red)	553	573	57	3.3* (photostablity of Dendra [11])	-	monomer

(a) Relative brightness (% of eGFP). (b) Time for bleaching from an initial emission rate of 1,000 photons/s down to 500 photons/s (t1/2; for comparison, fluorescein at pH 8.4 has t1/2 of 5.2 s); data are not indicative of photostability under focused laser illumination.

motA and FP-motB were expressed under the control of the arabinose promoter by pBAD33 vector to ensure a tight regulation of the expression in motAB knockout RP437 derivative *E.coli* cells [30]. Gene replacement on the chromosome was not tried, because we wanted to have a control over the gene expression using different inducer concentrations. The gene expression control was intended to perform the classic resurrection experiment [22-23], which requires cells to express MotAB proteins gradually. The advantage of having a gene expressed by inducible vector system is being able to control the expression level by the different concentrations of inducer molecules and simpler step to construct the strain. The disadvantages of having gene expressed by inducible vector system can be 1) higher cell to cell variations, 2) different growth conditions as compared to the cells do not carry any plasmid vectors, 3) lower or higher gene expression than the native expression level, and 4) over expression of the gene of interest may interfere with other endogenous gene expression mechanisms.

Despite the fact that the structure of the stator complex is not known, based on the existing structural information, we can predict the following: 1) one stator complex carries two FPs linked directly to the N-terminus MotB, likely forming a dimer, because the known stoichiometry of MotA and MotB trans-membrane proteins is 4 MotA and 2 MotB [15], 2) the FP tag is placed in the center of the stator complex, because the two MotBs are likely located at the center of the stator complex [16], 3) the FP tag may directly or indirectly interact with the cytoplasmic region of MotA, because MotA proteins have two long cytoplasmic loops around the short N-terminus stretches of MotB (Figure 3.2a); 4) the FP tag creates an extended stretch below the stator complex, because the FP is probably longer than the length of the cytoplasmic complex of MotA. The cytoplasmic part of MotA can be predicted to be about 3.5 nm long (personal communication with Dr. Gilles Labesse) and the FP has a shape of a cylinder with a length of 4.2 nm and diameter of 2.4 nm [8], 5) the FP tags may interact with FliG proteins in the C-ring complex, because FliG proteins interact with the torque generating cytoplasmic domain of MotA [17]. Therefore, considering all these topological constrains, it is likely that the FP tag interacts with the cytoplasmic loops of MotA and FliG in the C-ring (Figure 3.2).

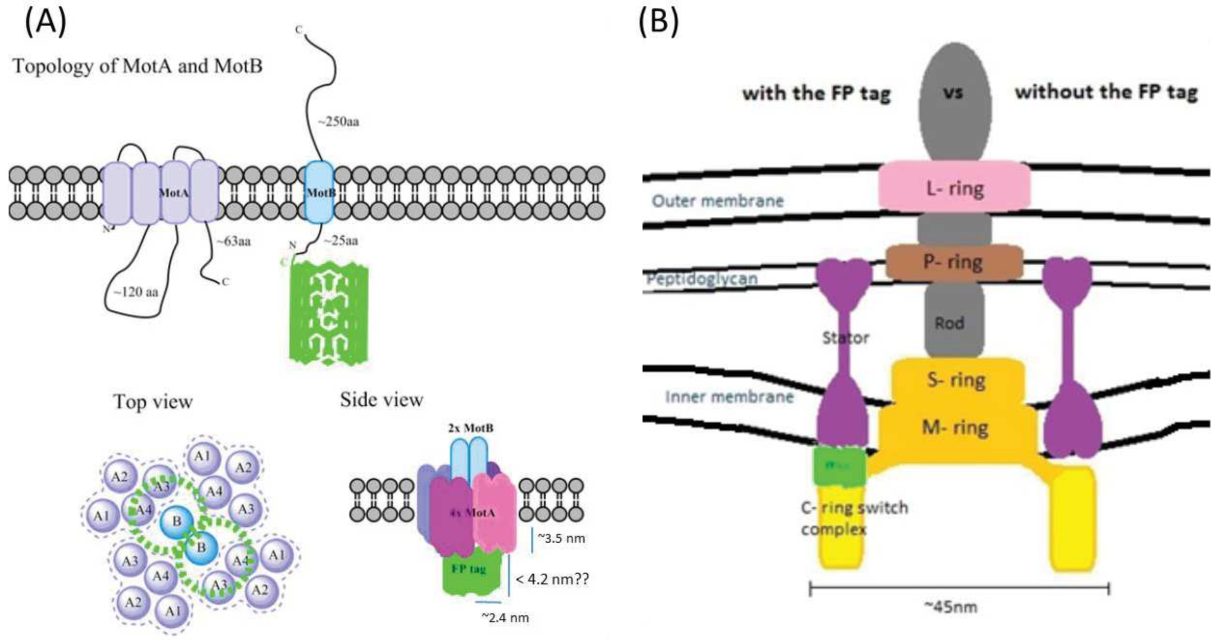


Figure 3.2. (A) Topology of the MotA and FP-MotB fusion protein and the top and side view of the schematic illustration of the fusion stator are shown. The AA lengths of each cytoplasmic loops and peptidoglycan domain are shown. MotA has four transmembrane domains and MotB has one. The C-terminus end of the FP is linked directly to the N-terminus of MotB and the FP tag is located at the center of the stator complex with an extended stretch below the stator complex. The amino acid (AA) sequence of MotA, MotB, FP are 333, 309, 239 (eGFP), 230 (Dendra2), 238 (YPet), respectively. (B) A schematic illustration of the BFM complex with FP-tagged stator or WT stator (no FP tag).

3.3 Reduced chemotaxis motility by fusion stator motors

In order to check the success of the FP fusion stator constructs, their motility and fluorescence signals were checked prior to the single motor analysis. Fluorescence microscopy data is described in chapter 4. The motility of the strains with the fusion stators were checked by chemotaxis assays. JPA604 strain (MotAB knockout of RP437 strain) carrying appropriate plasmid vectors (pBAD33) expressing MotA and FP tagged MotB was used. As a positive control, wildtype (WT) strain expressing MotA and MotB by the same vector was prepared. Another wildtype (WT_np, RP437) strain expressing MotA and MotB by their native promoter was also prepared. The soft agar plates contained appropriate antibiotics and indicated concentration of inducers, except for RP437 strain which express stator proteins by the native promoter. Chemotaxis performance include the abilities of the motors to rotate and to switch properly. For simplicity, the strains of *E.coli* used in this work are referred to as YPet, eGFP, Dendra2, WT and WT_np strains. When the chemotaxis speed of the two eGFP strains were compared, both were motile, but the eGFP-MotB showed a larger diameter of the colony than the 28AA eGFP-MotB fusion strain (Figure 3.3). This indicates that this additional 28 AA MotB N-terminal part was not necessary and caused a more reduced motility. Thus, eGFP-MotB fusion, without the 28AA, were used as eGFP stator throughout the study.

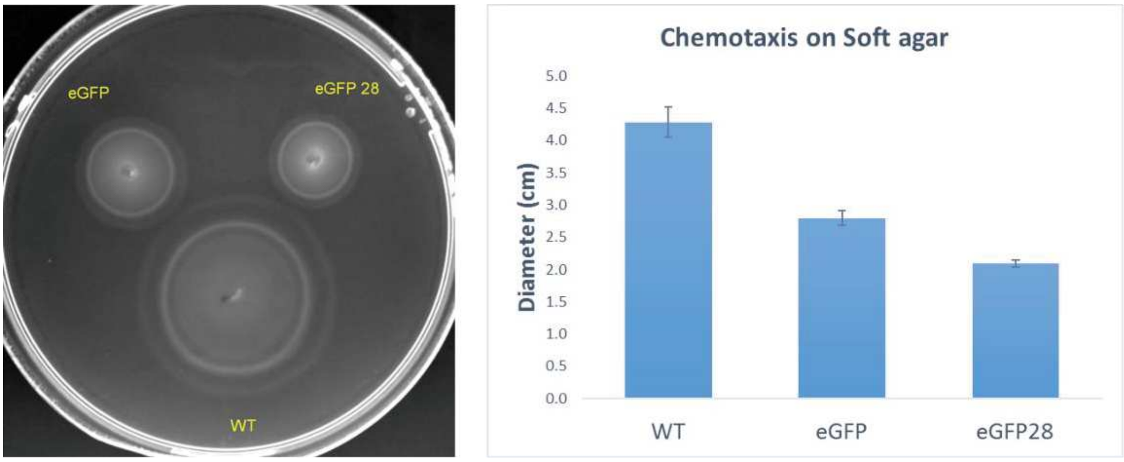


Figure 3.3. Left: a representative chemotaxis soft agar plate with WT, eGFP and 28 eGFP strains. Right: chemotaxis motility of the same strains for 8 hours on the 0.25% soft agar in the presence of 0.01% arabinose. Five replicates of the same assay were performed and the error bars indicate the standard deviations of the values.

When the motilities of the three fusion stator strains were compared, YPet and eGFP showed a similar chemotaxis speed (~50% speed of WT), with YPet being slightly faster than eGFP. Dendra2, on the other hand, showed much lower chemotaxis speed than YPet or eGFP (30-40 % speed of WT) (Figure 3.4). The chemotaxis speeds were increased as the L-arabinose (gene inducer) concentrations increased, with a maximum speed at 0.1% arabinose as expected by [19] (Figure 3.4). When the chemotaxis motility of the two wildtype strains - WT_np (RP437) and the WT (JPA604 carrying pBAD33_MotAB)- were also compared, the WT_np chemotaxis at a similar speed to the WT in soft agar plate with arabinose 0.325 mM ~ 0.65 mM (equivalent to 0.05~0.01% of [ara]). However, a direct comparison of these two data should be avoided, because WT and WT_np have 1) different growth conditions since no antibiotic and L-arabinose was required for RP437 strain, which does not carry any plasmid vectors, and 2) different in cell to cell variations; more homogenous gene expression is expected in WT_np than in WT. In addition, RP437 showed a broader range of chemotaxis speeds based on the softness of the agar plates for unknown reason.

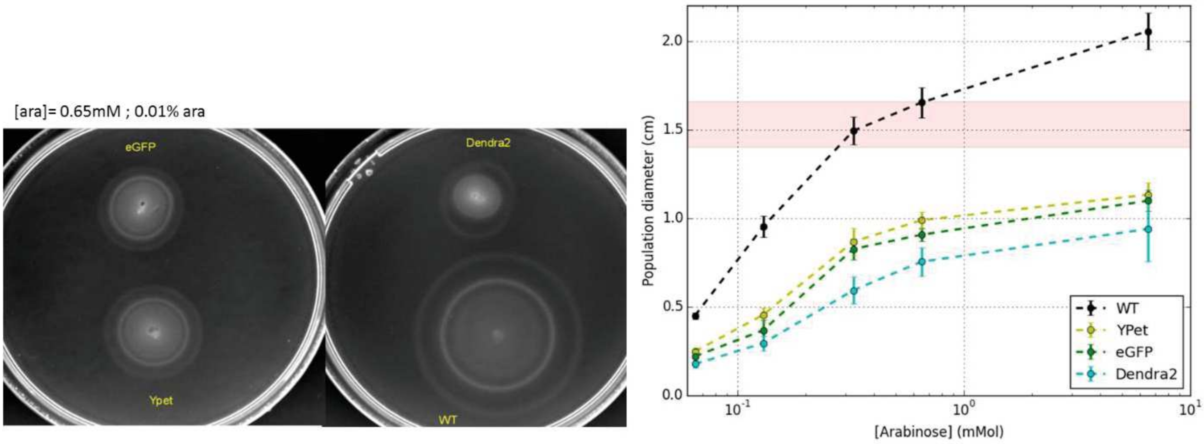


Figure 3.4 Left: Representative chemotaxis assay (at 0.65 mM ara) showing the four strains. Right: Chemotaxis motility of the four strains by different inducer concentrations. L-arabinose concentrations of 0.001%, 0.002%, 0.05%, 0.01%, 0.1% are equivalent to 0.065 mM, 0.13 mM, 0.325 mM, 0.65 mM, 6.5 mM, respectively. The area marked pink represents the RP437 strain (WT_np) chemotaxis motility.

The YPet and Dendra2 fusion stators were also expressed by three other vector expression systems than pBAD33, and their motilities on soft agar were tested (Figure 3.5). The chemotaxis speeds varied (~30%) depending on the expression vectors used, but the relative chemotaxis speeds of the three strains (WT, YPet, Dendra2) remained constant. pMMB206 and pBAD33 vectors showed a tight protein expression control, while pFX40 and pTrc99a [20] vectors did not. Interestingly, YPet fusion by pTrc99a vector did not swim well when IPTG (50 μM) inducer was present in the media, but swim better when no inducer was present. The OD600 growth curve of the YPet-MotB fusion stator on pTrc99a strain also clearly revealed that the cell cannot grow well when the IPTG inducer was added to the media (Appendix A). The underlying reason behind this reduced growth rate is not known. One hypothesis is that the gene expression by the pTcr99a vector is higher than other vectors, and over expression of the YPet-MotB fusion protein inhibits the cell growth. Protein expression level measurements will be required to answer these questions. In this study, only pBAD33 vectors were used for the experiments.

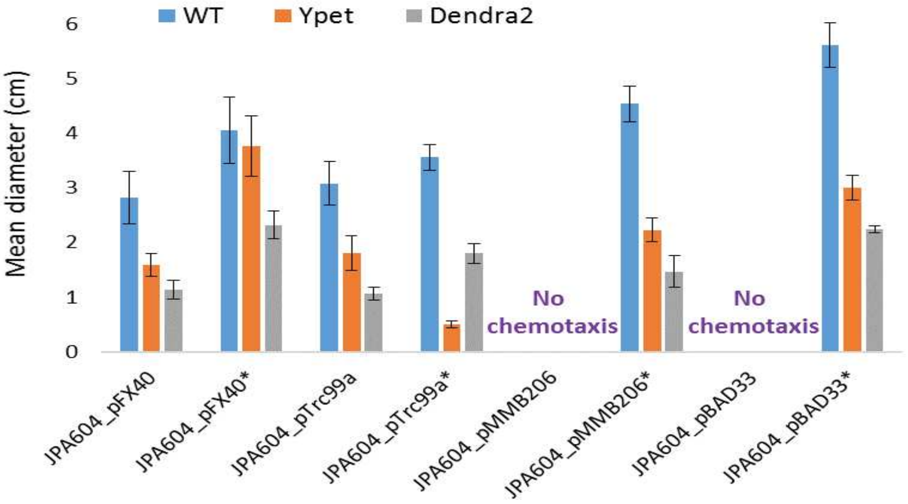


Figure 3.5. Motility tests of the WT, YPet and Dendra2 fusion stators expressed by four different vectors. The graph shows the mean diameters (in cm) of the indicated strains, after ~8 h on 1/2 LB soft agar plates. The semi-solid agar plates were without or with (second columns with *) inducers. The inducer used was either 50 μM IPTG or 0.2% arabinose. The antibiotics ampicillin (100 μg/ml) were used for pFX40 and pTrc99a, and chloramphenicol 10 μg/ml and 33 μg/ml were used for pMMB206 and pBAD33, respectively. The strains carrying pMMB206 and pBAD33 vectors did not chemotaxis when no inducer was present, suggesting a much better control over the protein expression than pFX40 and pTrc99a. The data is from the three to four independent experiments, and the error bars represent standard deviations.

3.4 Single motor rotation speed by fusion stator motors

In order to examine the functions of the motors carrying the three fusion stators, 1.1 μm diameter beads were attached to the truncated hydrophobic flagella filament of the motor and their rotation was monitored by a fast CMOS camera at 300 frames per sec (FPS). In figure 3.6, we show the distributions of the speeds recorded from the single-motors of the four strains. The rotation speeds of WT varied from ~20 Hz to ~ 70 Hz with an average rotation rate of ~45 Hz (number of motors = 56). The wild type motors (WT_np), expressing the stator genes by the native promoter in the genome, also showed similar rotational speeds, suggesting the growth condition (presence of antibiotics and arabinose in the media) and the source of the stator expression do not affect the motor rotation speeds in this experiment condition. Despite the big difference in the chemotaxis speed, the average rotation speed of YPet fusion stators (~43 Hz, n.motors= 59) was close to the WT rotation speed. On the other hand, motors with the stators tagged by eGFP and Dendra2 produced 58% and 33% average rotation speeds with respect to WT (~ 26 Hz, n.motors= 39 and ~15 Hz, n.motors= 54, respectively).

This assay was performed in both high (0.1% L-arabinose = 6.5 mMol) and low induction conditions (0.002% L-arabinose = 0.13 mMol). The 0.1% of L-arabinose was selected for the high induction growth condition as pBAD expression vector induces maximum expression around 0.1 ~ 0.2 % of arabinose concentration [19]. The 0.002% was selected for the low induction growth condition as the immune-blot assay performed in [4] suggests that 0.002% arabinose expresses MotAB proteins close to the stator gene expression level by the native promoter. Although a slightly reduced average rotation speeds were observed from the low induction growth condition cells, possibly due to less number of stator bound to the motor in average, the average rotation speeds of the cells grown in these two different conditions were similar. This suggests that the number of stators expressed by the low induction condition (0.002% ara) was high enough to generate a motor with full (or high enough number of) stators. Another difference between these two induction conditions was the yield of finding the rotating beads. In general, a higher yield of rotating beads was seen from the cells grown in the high induction condition than from the cells grown in the low induction condition. This may indicate that the cells grown in the low induction condition contain higher number of empty motors (motors without any stators bound). The major difference between the cells grown in these two induction conditions was their switching frequencies, which will be discussed in section 3.7.

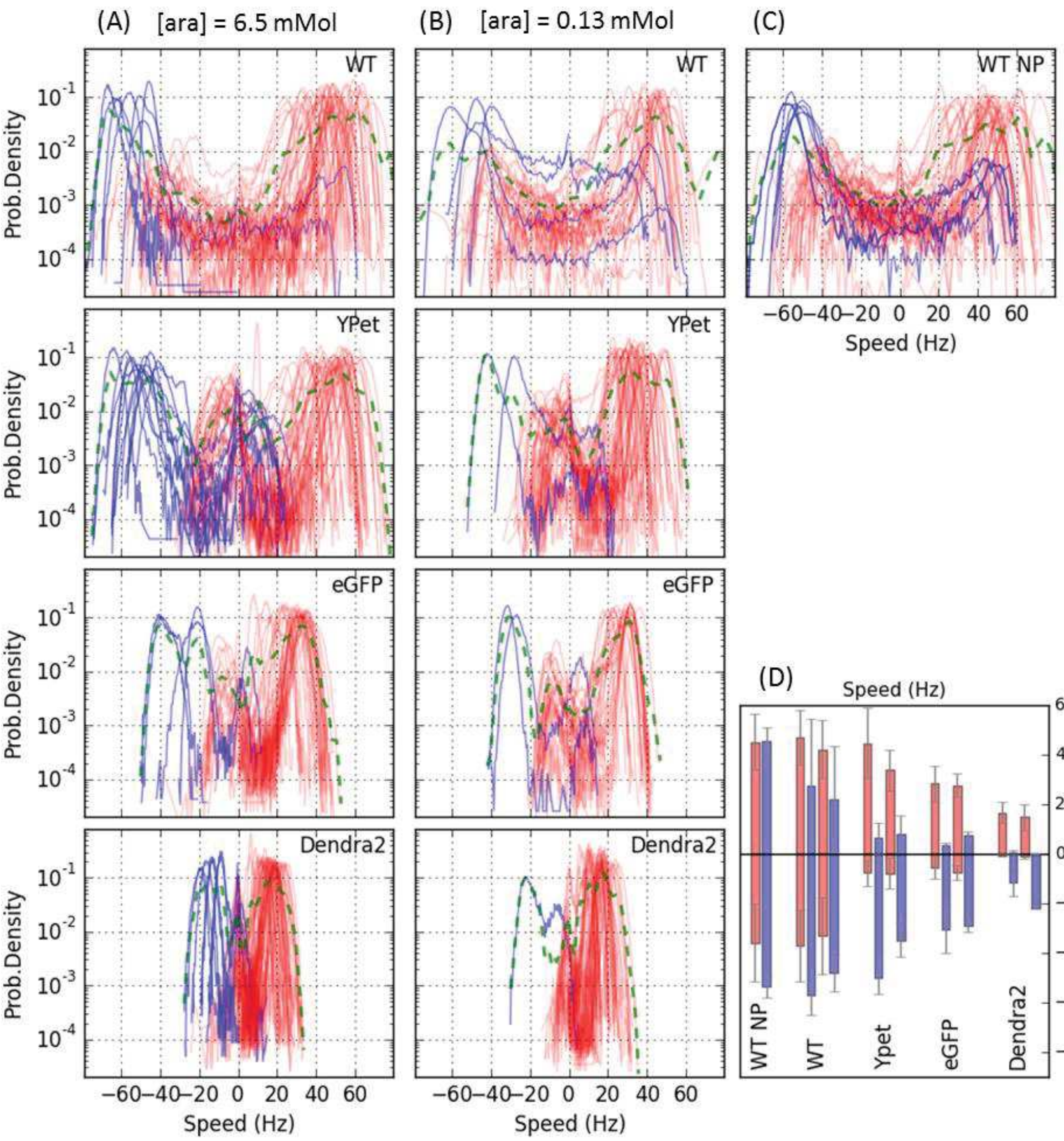


Figure 3.6. Speed distributions for the four strains (wild type MotB, YPet-MotB, eGFP-MotB, Dendra2-MotB fusion stators) at (A) high [ara] = 6.5mMol and (B) low [ara] = 0.13 mMol induction and (C) WT_np motor (MT02, stator induced by native promoter). Red indicates CCW biased cells and blue indicates CW biased cells. The number of motors measured were 56, 59, 39, 54 for high induction and 34, 47, 27, 44 for low induction (for WT, YPet, eGFP and Dendra2, respectively). The green dashed lines represents the averaged distribution of all motors. (D) the mean rotation speeds of the speed distributions of A~C. The red bars indicate CCW biased cells and the blue bars indicate CW biased cells. The first two bars (red and blue) are from the high induction growth cells and the other next two bars (red and blue) are from the low induction growth cells. WT_NP has only two bars instead of four bars, since no inducer was needed for this strain. The error bar represents the standard deviation. Each cell was measured for 3~5 minutes at 300 FPS (time resolution of ~10 ms).

The motors that rotate in counter-clock wise (CCW) for more than 50% of the total recording time (approx. 3min) are defined as CCW biased cells, and vice versa for CW biased cells. When CCW biased cells do not switch at all, the CCW bias is 1. While most of the motors were CCW biased cell, occasionally some CW biased cells (below 0.5 of the CCW bias) were observed from all four strains (Figure 3.7a). These CCW and CW biased cells are represented by red and blue curves in figure 3.6. Despite the big differences in switching frequencies (chapter 3.7), there was no big difference in the rotational bias between the fusion stator motors and WT motors with an overall mean of about 86 % of time spent rotating in CCW, except the high induction (0.1% ara) YPet motors and the low induction Dendra2 motors (figure 3.7b).

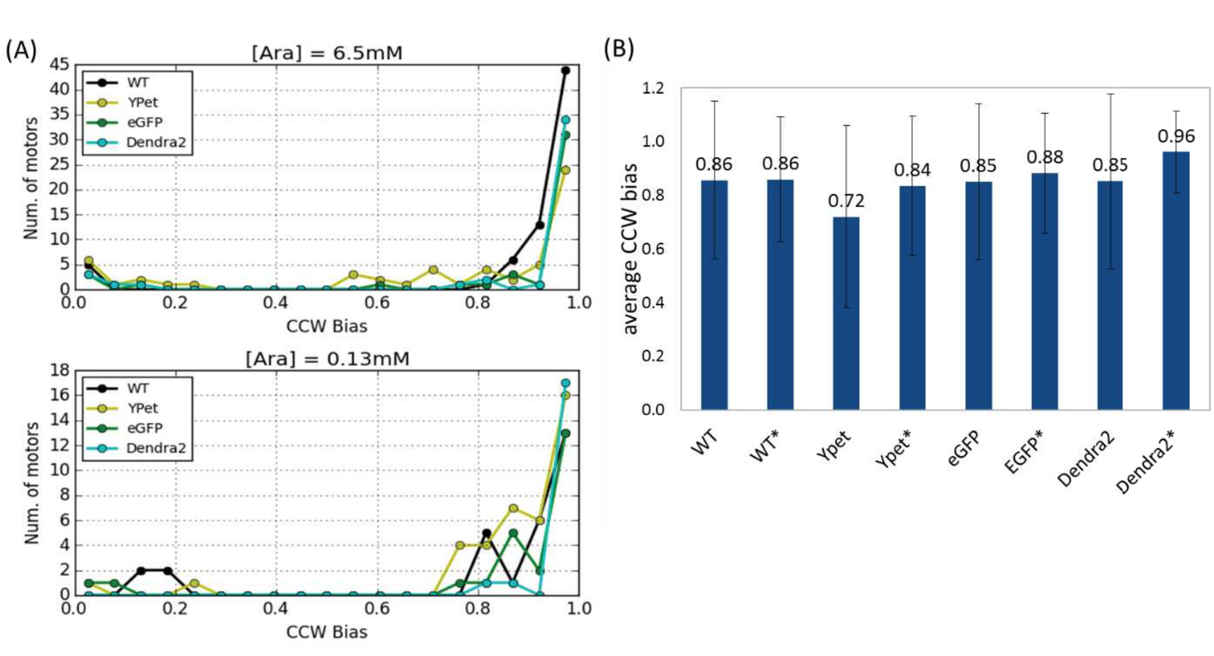


Figure 3.7. (A) Number of the motors plotted against CCW bias. When CCW biased motors do not switch at all, the CCW bias is 1. When CW biased motors do not switch at all, the CCW bias is 0. (B) Averaged CCW bias of the total cells. The x axis labeled with * (asterisk) indicate the motors from the [Ara]=0.13mM and without it are the motors from [Ara] = 6.5 mM.

Both CCW and CW biased cells from all four strains showed similar mean rotation speeds, which are represented by the two main peaks (green curves summing the blue and red curves) at opposite speeds in figure 3.6. Therefore, the reduced average rotation speeds by the eGFP and Dendra2 stators affects symmetrically both CW biased and CCW biased cells. The peaks of the blue and red curves of the tagged motor decreased by roughly 4% (YPet), 40% (eGFP) and 68% (Dendra2) as compared to WT motor. We call this effect “symmetric torque reduction (STR)”. When a single WT motor switched between CW and CCW states, their rotation speeds were rather symmetric. In other words, when switching, WT motors reach the same absolute value of the speed in the opposite direction. Interestingly, when fusion stators motors switch, this symmetric switching speeds was not observed (figure 3.6, 3,8). For example, when CCW biased YPet motor switches from CCW to CW, it reaches a fraction of its previous speed in the CW direction (and the same is true for CW biased cells switching from CW to CCW). This was true for all three tagged motors. We call this hindered switching as “asymmetric switching” (ASW). Therefore, in summary, tagged motors exhibit symmetric torque reduction (STR) by showing the reduced speed in both CW and CCW biased rotation direction, and asymmetric switching (ASW) by showing a lower speed in the less-biased rotation direction when they switch.

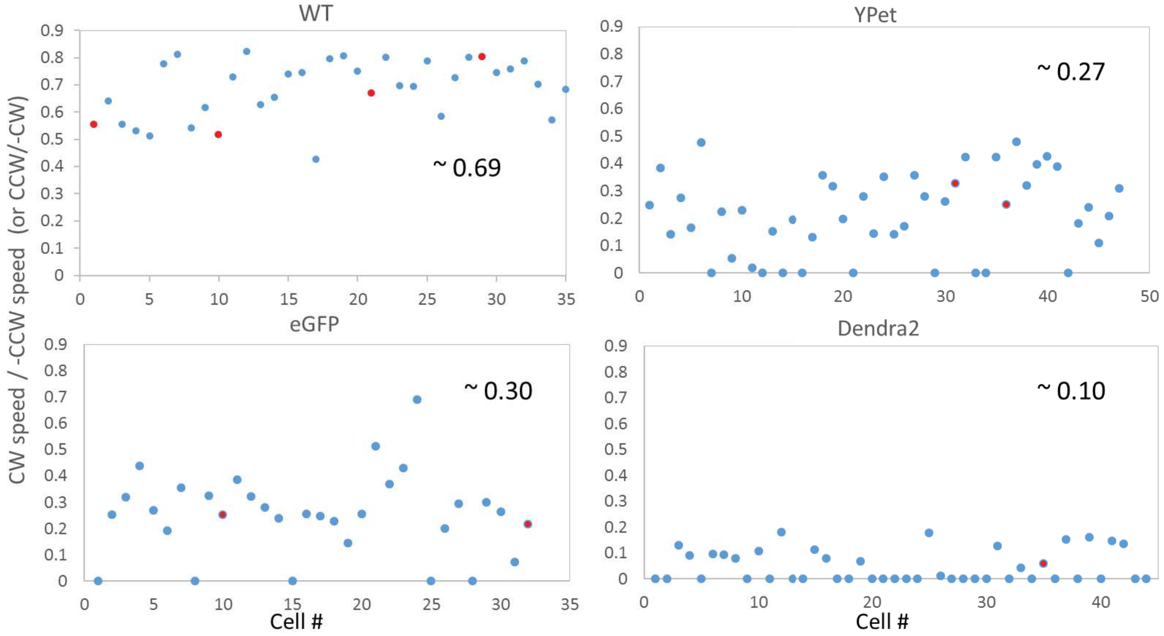


Figure 3.8. Ratio of the CW mean rotation speeds to the CCW mean rotation speeds (blue dots), and ratio of the CCW mean rotation speeds to the CW mean rotation speeds (red dots); less-bias direction over bias direction. Each dot represents individual motors. Average of the mean ratios (excluding the zero ratios) are shown on the graphs. The data is from the low induction condition cells.

3.5 Different torque generated by single fusion stators

Regardless of their bias, the eGFP and Dendra2 fusion stator motors showed reduced absolute rotation speeds compared to the YPet fusion and WT motor (in terms of the STR, defined above, of the motors with the eGFP and Dendra2 tags). Since torque and rotation speed of the motor are proportional to the number of stators bound to the motor [22, 23], and the torque generated by a single stator can also vary depending on the load, temperature and pH [2, 24, 25], the observed symmetric speed reduction could be the result of either lower number of active fusion stators bound to the rotor complex, or of a lower torque produced by each fusion stator. In order to identify the underlying cause for such reduction, the torque generation per single stator among the four strains was compared. This comparison was made by measuring the step-wise speed increments (or decrements) during the bead assays, which corresponds to the stator association and dissociation to the rotor complex [23].

Normally, the rotation speed of the motor was constant during the 3~5 min of the recorded period. However, we occasionally observed cells showing spontaneous step-wise increments and decrements of the rotation speeds (figure 3.9). Since the stators actively engage and disengage around the rotor [1], it is assumed that these naturally occurring steps in torque are due to the dynamic interactions between the stator and the rotor (termed "stator turnover" in [1]). In addition to this subset of motors showing spontaneous step-wise speed changes in the switching strain (JPA605), additional bead assays with the non-switching ΔCheY strain (JHC36) was performed, because switching events may hinder accurate analysis of the step-wise speed increments.

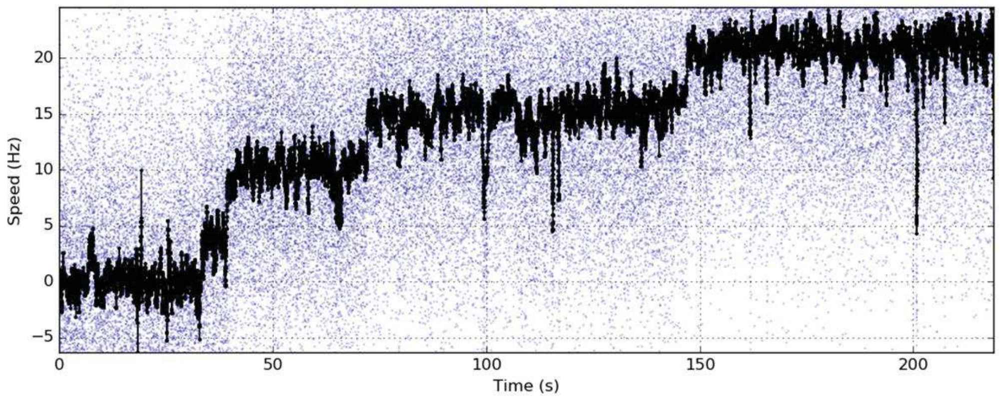


Figure 3.9. Representative speed-time traces of a WT motor showing spontaneous speed increments during bead assay.

Torque values, instead of speed, were compared to estimate an accurate average torque generated by single stators, because torque values can make more accurate comparisons as it takes into account the difference in the lengths of the flagellar stubs where rotating beads are attached to. The difference in flagellar stubs lengths lead to different drag coefficients of each motor, which affect the motor rotation speeds at a given generated torque. Torque was calculated from the estimated drag coefficients and the speed measurements. At low Reynolds number, torque is given by $\tau = \gamma * \omega$, where ω is the angular speed and γ is the drag coefficient caused by a bead rotating on a circular trajectory with a radius r. The drag coefficient is estimated to be $\gamma = 6\,\pi\,\eta\,r$, where η is a viscosity of fluid media, and the equation is based on the Stokes formula [21]. The average diameter of the circular trajectory, which is proportional to 2 x stub length, was 1.344 ± 0.2 μm (measured by the eGFP motors). Figure 3.10 shows representative WT and eGFP motors with discrete torque changes and their torque histograms with multiple Gaussian distributions.

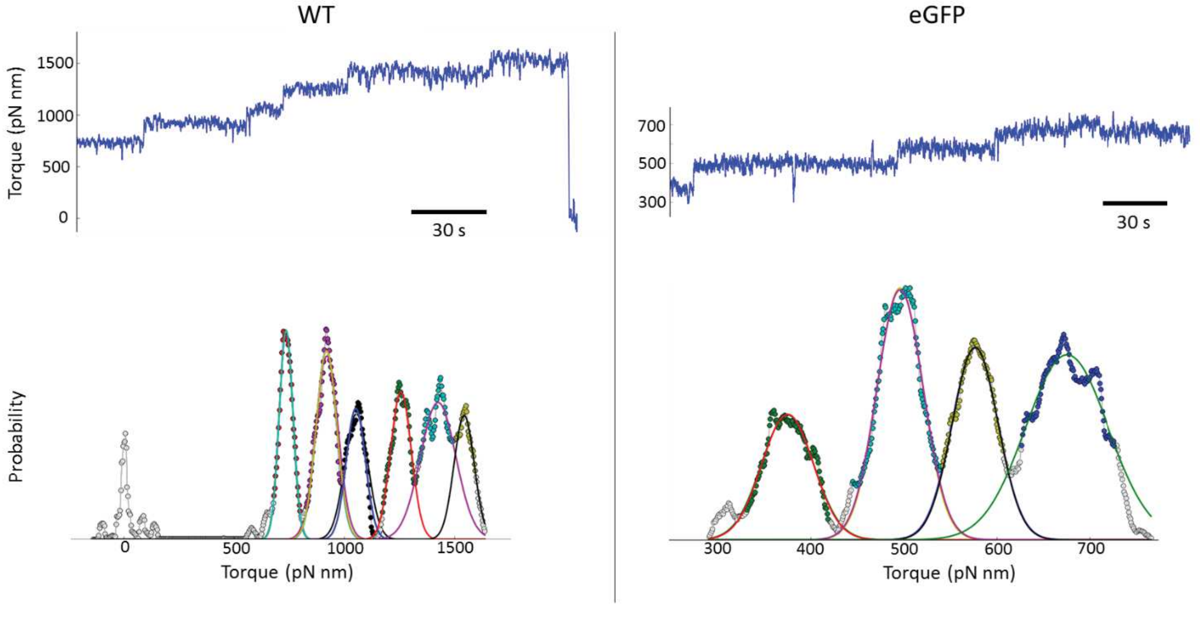


Figure 3.10 Torque-time traces and their torque histograms of the representative WT and eGFP motors. The torque-time traces show discrete torque changes over time. The torque levels were determined by fitting multiple Gaussian distributions to the histograms (the algorithm which draw the histograms artificially amplifies the discrete torque levels by summing local histograms along the time trace). The Gaussian curves, determined by selecting two positions around the peak manually, are shown in different colors in the histogram. The torque peaks were separated by roughly equal intervals, corresponding to different numbers of stators bound to the motor. Here, WT motor shows torque intervals of roughly ~180 pN nm, while eGPF motor shows torque intervals of ~ 100 pN nm.

Typically, the measured torque values were separated by roughly equal intervals, with some variations in torque increments within the same motor and among different motors. The measured torque intervals from individual motors (the total number of 17~18 cells per strain) are shown in figure 3.11a, and the distribution of the pair-wise distances between measured torque steps is shown in figure 3.11b. These graphs describe the probability distributions of all the distance between the values of the torque plateaus found in given traces. From the first peak of the pair-wise distance distributions of torque levels, we obtained a good estimation of the torque generated by a single stator. The torque generated by WT single stators is estimated to be 162 ± 48 pN nm, which agrees with the previous two studies, 158 ± 49 pN nm by [24] and 146 ± 35 pN nm by [22]. YPet motors generated similar torque per stator to the WT motors but only slightly reduced with a wider distribution, with a peak value of 149 ± 73 pN nm. eGFP and Dendra2 fusion motor showed reduced torque per stator; 96 ± 35 pN nm and 65 ± 38 pN nm, respectively. In general, it was more difficult to observe steps from the Dendra2 strains than other strains, probably due to their torque intervals were close to the noise level in the speed time traces. In summary, the torque intervals by YPet motors do not decrease much with respect to WT, while eGFP and Dendra2 motors do. In addition, since YPet motors were able to generate similar torque per stator (torque intervals) to the WT motor, while their switching ability was significantly impaired (Asymmetric Switch) compared to the WT motor, this result may imply that torque generation and switching ability can be decoupled.

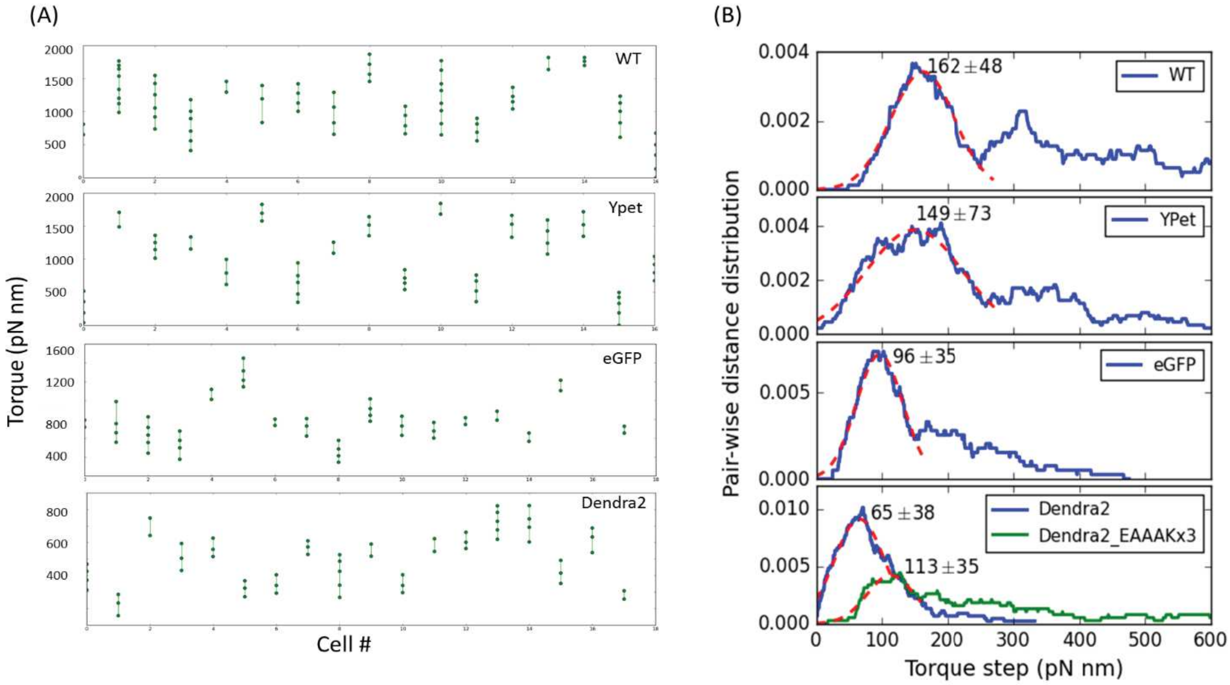


Figure 3.11. Stepwise torque changes of the four strains. The data was collected by normal bead assays with JPA605 strains and with JHC36 (non-switching) strains. (A) Torque changes per individual motors. # of cells measured: WT=17, YPet=17, eGFP=17 and Dendra2=18. (B) Pair-wise distance distribution of the measured torque steps of the five strains. The peak values of the distribution are shown, which are determined by Gaussian fit in the first peak of each distribution.

Table 3.2 summarize the activity of the BFM measured at three different levels: population chemotaxis, single motor, single stator. Except for the chemotaxis motility of YPet, the order of the three motility measurements of the four strains is consistent: WT shows the best motility followed in order by YPet, eGPF and Dendra2. The incomplete switching ability of the YPet motors explains the reduced chemotaxis motility of the YPet strains with respect to the WT strains, despite their similar average torque in single motor assay. The average rotational speed of WT motors was 46 Hz, which is equivalent to ~1,600 pN nm. Since the torque generated by a single WT stator was estimated to be 162 ± 48 pN nm, the average number of stators in the motor can be estimated to be 10 ± 2 (1,600 divided by 162). In the same manner, the estimated average number of the FP-tagged stators in the motor are 10 ± 3 for YPet motor, 10 ± 3 for eGFP motor and 8 ± 3 for Dendra2 motor. Therefore, except Dendra2 motor, all three motor had about 10 stators bound to the motor, suggesting that the mechano-sensing (the ability to recruit more stators when experiencing a high load on the motor) of the YPet and eGFP fusion stators was similar to WT motor on relatively high load (~1.1 μm bead was attached to the flagellar stub). Since the number of stators when fully bound in the motor is proposed to be 11 [22], this result also eliminates the possibility of having MotA alone stators (stator complex without MotB) bound to the motor and disturb the torque generation and the switching events.

Table 3.2 Summary of the three motilities of the four strains

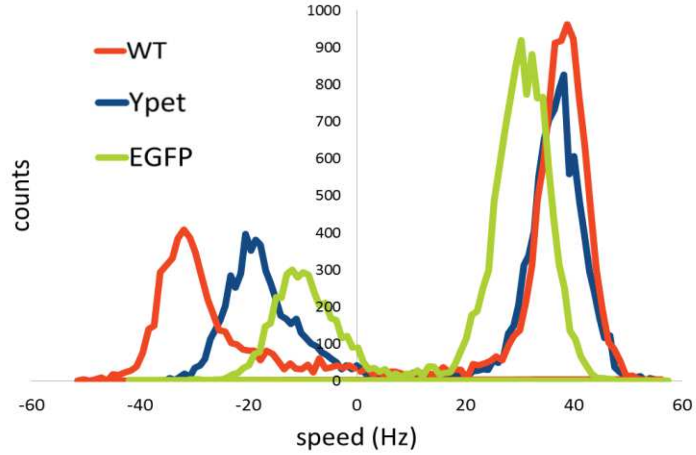
	Chemotaxis (cm) *	Single motor (Hz)*	Single stator (pN nm)	Estimated average number of stators bound to the motor
WT (MotB)	5.2 ± 0.2	45.7 ± 10.8	162 ± 48	9.9 ± 2.5
YPet-MotB	2.9 ± 0.2	43.3 ± 11.7	149 ± 73	10.2 ± 2.7
eGFP-MotB	2.8 ± 0.1	26.6 ± 7.1	96 ± 35	9.7 ± 2.6
Dendra2-MotB	2.4 ± 0.4	15.6 ± 5.1	65 ± 38	8.4 ± 2.7

*The chemotaxis values are from [ara] = 6.5mM and the single motor speed values are from [ara] = 6.5mM CCW bias cell.

3.6 Impaired switching abilities by fusion stator motors

3.6.1 Asymmetric switching of the tagged motor

As shown in figure 3.6 and 3.8, the most apparent difference in WT and FP fusion motors was the different ratio of the mean speeds in CW and CCW rotation in a switching motor; WT motors rotate at similar speeds in both directions (symmetric switching), while FP fusion motors rotate in different speeds from one direction to another (asymmetric switching). Asymmetric switching (ASW) behavior by tagged motors was consistently observed throughout the whole bead assays. The degree of ASW varied by different fusion motors as well as within the same fusion motors (figure 3.8). The ASW was most apparent in Dendra2 tagged motors, which pause rather than rotating in the opposite direction. The speed histograms shown in figure 3.12 illustrate the asymmetric switching of the YPet and eGFP motors (lower speeds in CW rotation with respect to their CCW), and the symmetric switching of WT motor (similar speeds in both CW and CCW rotations). In this histogram, WT and YPet shows similar time spent in CW direction of rotation (~34% of time), although WT motor switched much more frequently, indicating the difference in time spent in CW rotational states as well as a possibility of different switching time/duration. These two differences in switching events will be discussed in the following two sections.



	WT	Ypet	eGFP
FPS	300	300	300
filter_win	4	4	4
tot_time	49.99	49.99	49.99
bias_CW	**0.34**	**0.36**	**0.26**
bias_CCW	0.66	0.64	0.74
numb_switches	134.00	78.00	38.00
freq_switch	**2.68**	**1.56**	**0.76**
CCW mean time sec	0.47	0.80	1.73
CCW st.dev sec	0.33	0.66	2.81
CW mean time sec	0.27	0.48	0.73
CW st.dev sec	0.16	0.43	0.49

Figure 3.12. Speed histograms of the three CCW-biased WT, YPet and eGFP switching cells, showing ASW of YPet and eGFP motors. The histogram is plotted by the selected time course of 50 seconds when they showed most frequent switching events (WT, Ypet, eGFP switched 134 times, 84 times and 43 times, respectively). When YPet and eGFP tagged motors change their direction of rotation, the rotation speed could not reach their original rotation speed: YPet motor went from ~38 Hz to ~ -17 Hz and eGFP motor went from ~30 Hz to -11 Hz (WT motor: ~35 Hz to -33 Hz). Statistical data are shown as a table on the right.

3.6.2 Extended CW resident time by the tagged motors

The time spent between two consecutive switching events is defined as resident time, which also refer to the time remained at each CW or CCW rotational state: CW resident time or CCW resident time. In case of the motors rotating a 1.1 μm bead, the distributions of the CW and CCW resident time were exponential with a long tail (non-exponential) (figure 3.13). The main difference of the WT motors and fusion stator motors was the extended CW resident time of CCW biased motors, and vice versa for the CW biased motors. WT motors remained at CW rotational states for no more than ~1.3 s, while the FP tagged motors often showed the CW resident time of longer than 1.3 s, exhibiting a wider distribution with a longer tail than WT motors (figure 3.13). For example, the average CW resident time of the WT motors was 0.3 ± 0.1 s, while the YPet motors showed an average CW resident time of ~ 2.7 ± 3.7 s (data taken from the [ara] = 6.5 mM). This observation was true for the measurements from both [ara] concentrations. The difference in resident time between the WT motors and the fusion stator motors may explain the similar CCW bias (~0.85) of the WT, YPet and eGFP motors in average, shown in figure 3.7 [Ara]=0.13mM, even though their switching frequencies were different by a factor of two or four. The switching frequencies are discussed in section 3.7.

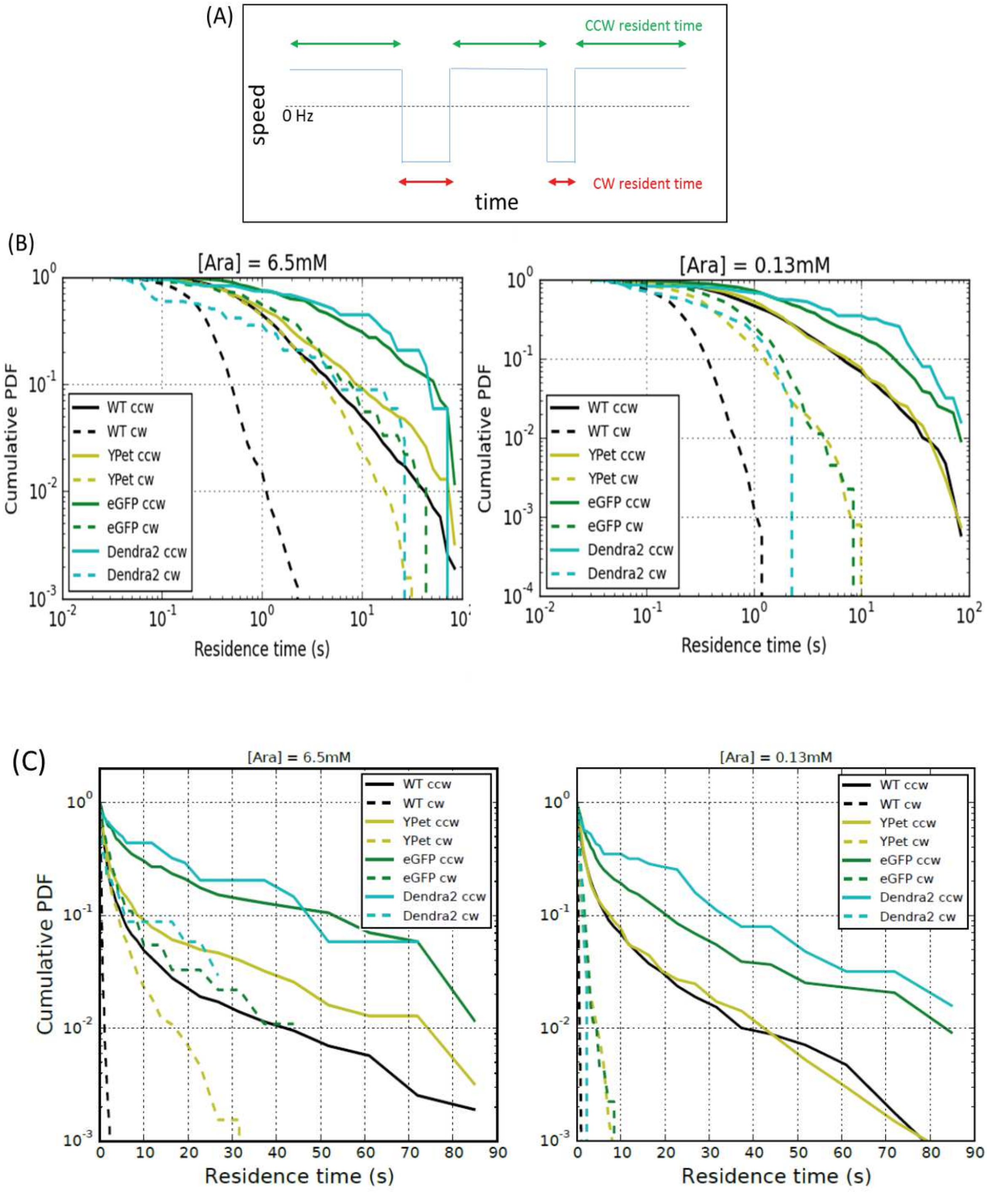


Figure 3.13. (A) Illustration of how to define CCW and CW resident times. (B) Resident time of the CCW bias cells in both induction conditions [ara] = 6.5mM (left) and [ara] = 0.13mM (right). The CW resident time of WT (black dashed lines) shows distinctly shorter tails than the other fusion stator motors in both [ara]. The CW resident time of [ara]=0.13 mM were lower than that of [ara] = 6.5 mM in average. In [ara]=0.13 mM, the CCW resident time of the WT and YPet (black and yellow solid lines) did not show a huge difference while eGFP and Dendra2 tend to have longer resident time. (C) is showing the same graph in (B) in second in x-axis, instead of logarithm values.

3.6.3 Extended switching duration by the fusion stator motors

As an initial attempt to characterize in more detail the difference in the switching process of WT and FP tagged motor, single switching events were detected and zoomed in. In order to check any difference in switching duration in the four strains, all the switching events of a single motor were overlaid and their speed histograms were compared (Figure 3.14). If the tagged motors spend more time to switch, we expect to see higher counts between the positive and negative mean speeds in the overlaid speed histogram during switch. Although certain differences in the switching event were observed in these histograms, it was hard to obtain accurate quantitative analysis. Mainly due to the lack of high temporal resolution of the data and the algorithms to make accurate comparisons.

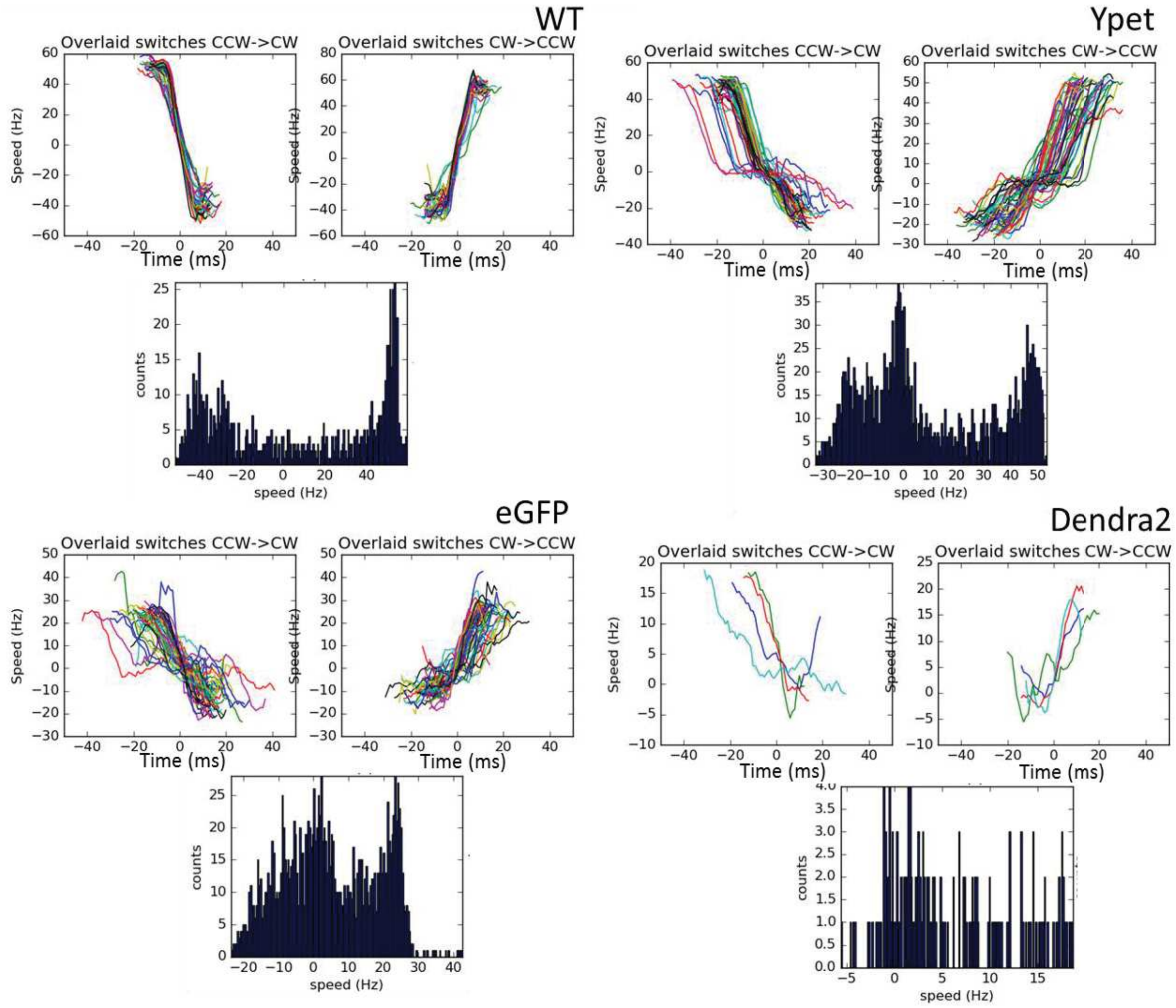


Figure 3.14 Examples of the overlaid switching events of the four strains' motors. We found complete switch events that cross both CCW and CW thresholds (2/3 of mean speed) in the same direction in the median filter set =11. Then the overlaid speed histogram was plotted by these switching events at 2/3 of mean speed ±100 (unfiltered) points in each side. WT motors show two distinct peaks which represent mean speed in CCW and CW, and low counts between these two peaks (i.e short time spent during the switch event). YPet motors here show three peaks; two on each side representing the mean rotation speeds and another peak at 0 Hz. eGFP motors show two peaks including one at 0 Hz. Therefore, the YPet and eGFP motors shown here tend to remain at 0 Hz during switching. Switching duration analysis for Dendra2 was difficult due to their low rotation speed and less frequent switching events.

In addition to the overlaid switches histograms, switching durations of the four strains were measured. Switching duration is defined as the time took from 2/3 of the mean positive speed to 2/3 of the mean negative speed (and from negative speed to positive speed as well). The switching duration distribution of the three strains (WT, Ypet and eGFP) showed peaks between 20 ~ 30 ms (Figure 3.15a). Since switching duration analysis of Dendra2 was difficult due to the low speeds and less frequent switching events, Dendra2 motors were not considered here. According to [27], the WT motor switching duration distribution have a peak at 8 ms with a standard deviation of 15 ms. A possible explanation for such a discrepancy with our measurement is the different time resolution and filter set used. The running window median filter is defined as replacing each point with the median of X points around it, where X is the filter set value. Figure 3.15b shows how different filter sets can shift the switching duration peaks noticeably. The three FP fusion strains also showed the similar shift. In the data from current study, the switching duration can be reduced to be closer to 8 ms if the filter set is reduced from 11 to 4 as well. However, a filter set of 11 was used instead of 4, because having the latter potentially causes errors in counting the switching events (due to the low time resolution of ~10 ms).

Interestingly, although the three strains showed the same peak around 25 ms, YPet and eGFP tagged motors showed wider switching duration distributions compared to WT motor. This difference can be seen by the average and standard deviation of the switch durations: for WT it is 33 ± 26 ms, for eGFP 49 ± 40 ms and for YPet 56 ± 50 ms. This is ~ 33% and ~ 60 % decrease in switching durations from WT. Therefore, together with the fact that the WT rotation speeds in both CCW and CW were higher than that of YPet and Dendra2, the FP tags caused longer switching duration, potentially due to their less efficient switching process. This result may help us understand how the switching mechanism by FP-tags was hindered. However, further analysis of switching duration should be performed at a higher time resolution, and perhaps at a lower load as well, because the data recorded here (300 FPS, 1.1 μm beads) do not have likely a resolution high enough for an accurate analysis of the cellular events happening within 10 ms.

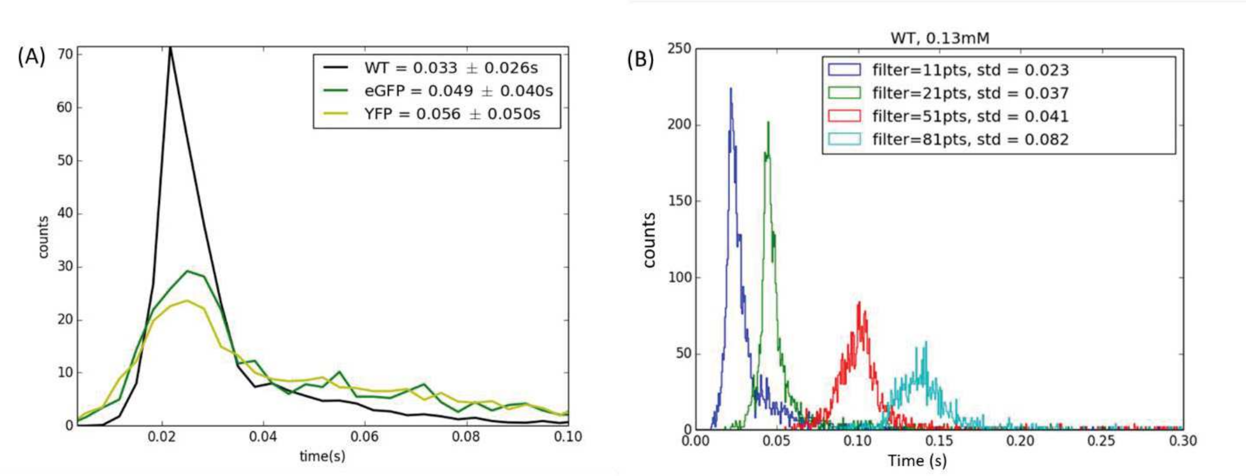


Figure 3.15. (A) switching duration of the three strains, [ara]=0.13 mM, filter set =11. Average switch duration values and standard deviations are shown in inset. (B) WT motor switching durations at different filter sets.

3.7 Switching frequency was highly affected by fusion stators

The switching frequencies of the fusion stator motors were also affected such that the three motors with FP-tags switched much less frequently than WT cells (Figure 3.16). YPet fusion stator showed less altered switching frequencies than eGFP or Dendra2, while switching frequency of the Dendra2 was most affected. This result follows somehow the same trend observed in the torque generation of single stators (YPet with 92 %, eGFP with 60 % and Dendra2 with 40 % of the WT torque generation). The influence of stators on switching frequencies/dynamics has been suggested before in [28]. In this study, torque dependent switching rate was suggested with a hypothesis that a stator can affects the switching frequency by accelerating the FliG conformation flipping rates when torque is applied to the FliG units. Since the interactions between the stators and the rotor component FliG is vital for both torque generation and switching, it can be expected that a distortion of the torque generating domain of the stator influencing the switching dynamics of the motor.

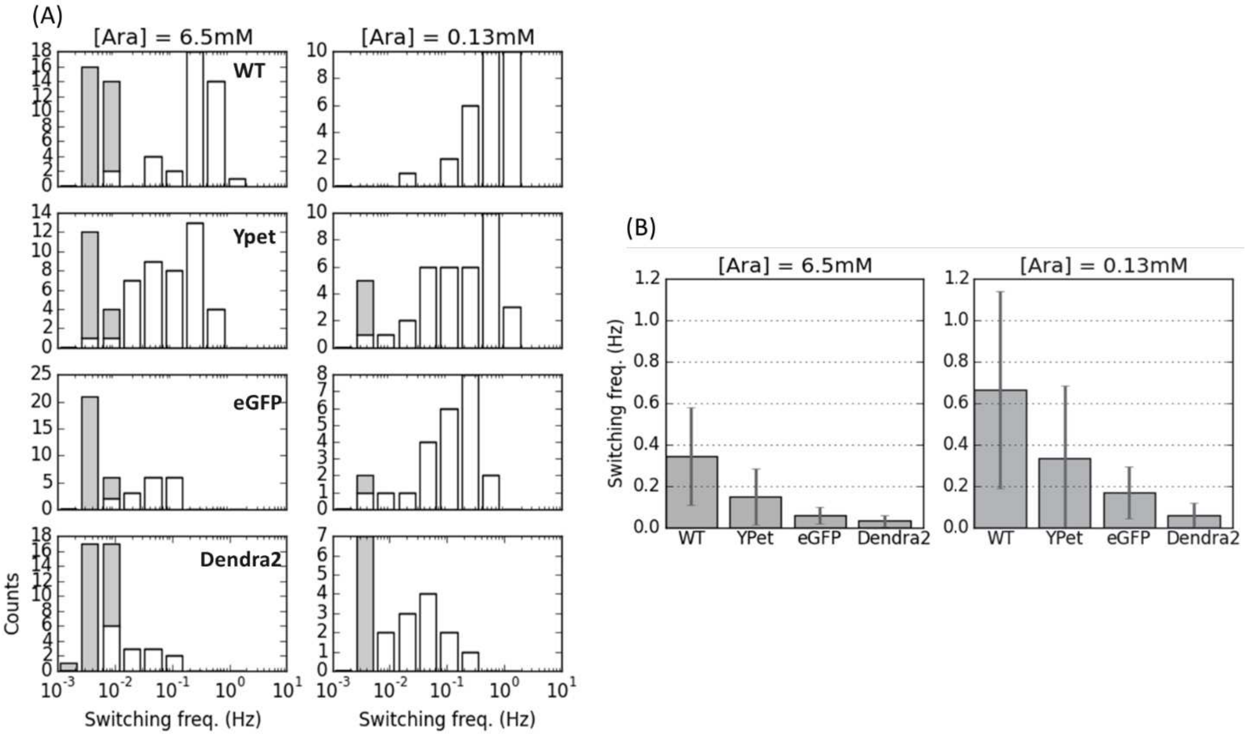


Figure 3.16. Switching frequencies of the four strains in ara = [6.5 mM] and [0.13 mM] induction conditions are shown as distribution in (A) and as mean with STD in (B). The colored bars in (A) represent cells that did not switch during the recorded time (~3 min).

Interestingly, the switching frequencies of the motors grown in the two different induction conditions were different by a factor of two (figure 3.16b). In high induction condition, WT motors showed a switching frequency of ~0.32 switches/second, including more non-switching cells during the recorded period (colored bars in figure 3.16a). In lower induction condition, the switching frequency of the WT motor increased to ~0.66 switches/second, close to the switching frequency of the WT_np (MT02, the strain that expresses stators by their native promoter) (MT02 data can be found in appendix B). The increased switching frequencies at lower induction condition were consistently observed by the fusion stator motors as well. This reduced switching frequencies of the cells over-expressing motAB is in line with [31]. In [31], the stator gene over-expressing motors showed lower fluctuations of CCW bias to some extent, but no quantitative data are available (Note that here we measured the cell rotation for about 3~5 minutes, whereas in [31] the authors measured for 30 minutes, and CCW bias fluctuations, not switching frequencies). Previously, the correlations of the switching frequencies to load, torque (speed), mechanical feedback of the flagellum have been studied [28-30], but no extensive study has been performed so far with respect to the stator expression level. Although we have observed a clear difference of the switching frequencies from the motors grown in different gene induction conditions, in order to validate a direct correlation between the stator expression level and the switching frequency, further investigations in the expression levels of the switching regulating signaling proteins (such as CheY, CheW) in both low and high induction conditions will be required, as over-expression may hinder expression levels of the other signaling molecules.

3.8 Having a linker between the FP and MotB improves motor functions

The previous results were based on the direct fusion between the FP and MotB. In this section, we studied how a linker (of 5AA and 15 AA) between the FP and the N-terminus of MotB influences the torque-generation and the switching ability of the BFM. Four types of linkers were tested: short and flexible (GGGGS), short and rigid (EAAAK), and two longer versions of them (GGGGS x3 and EAAAK x3) [32]. The functional effects of the linker were primarily focused on Dendra2 fusion stator, because they exhibited most impaired functions in both torque generation and switching ability. Dendra2 fusion stators with four different linkers were constructed and their motilities were investigated. Functional effects of the linkers on YPet motors were also investigated since they showed the most WT-like motor functions. The list of the fusion stators with a linker that we studied are: Dendra2-EAAAK-MotB, Dendra2-GGGGS-MotB, Dendra2-EAAAKx3-MotB, Dendra2-GGGGSx3-MotB, YPet-EAAAKx3-MotB, YPet-EAAAK-MotB, YPet-GGGGS-MotB.

Chemotaxis motilities of the fusion stators with a linker were compared to the fusion stators without a linker (figure 3.17). In Dendra2 fusion group, Dendra2 with a long and rigid linker (EAAAKx3) showed the largest chemotaxis diameter, followed in order by Dendra2-GGGGSx3, Dendra2-EAAAK, Dendra2-GGGGS and Dendra2. The fusions with a linker showed improved chemotaxis motilities and the two longer linkers (3x) improved chemotaxis more than the shorter versions (1x). In YPet fusion group, YPet with the rigid linker (EAAAK) also showed the largest chemotaxis diameter compared to the flexible linker (GGGGS). In both fusion groups, having a linker improved chemotaxis and the rigid linkers (EAAAK) always improved better than the flexible linkers (GGGGS).

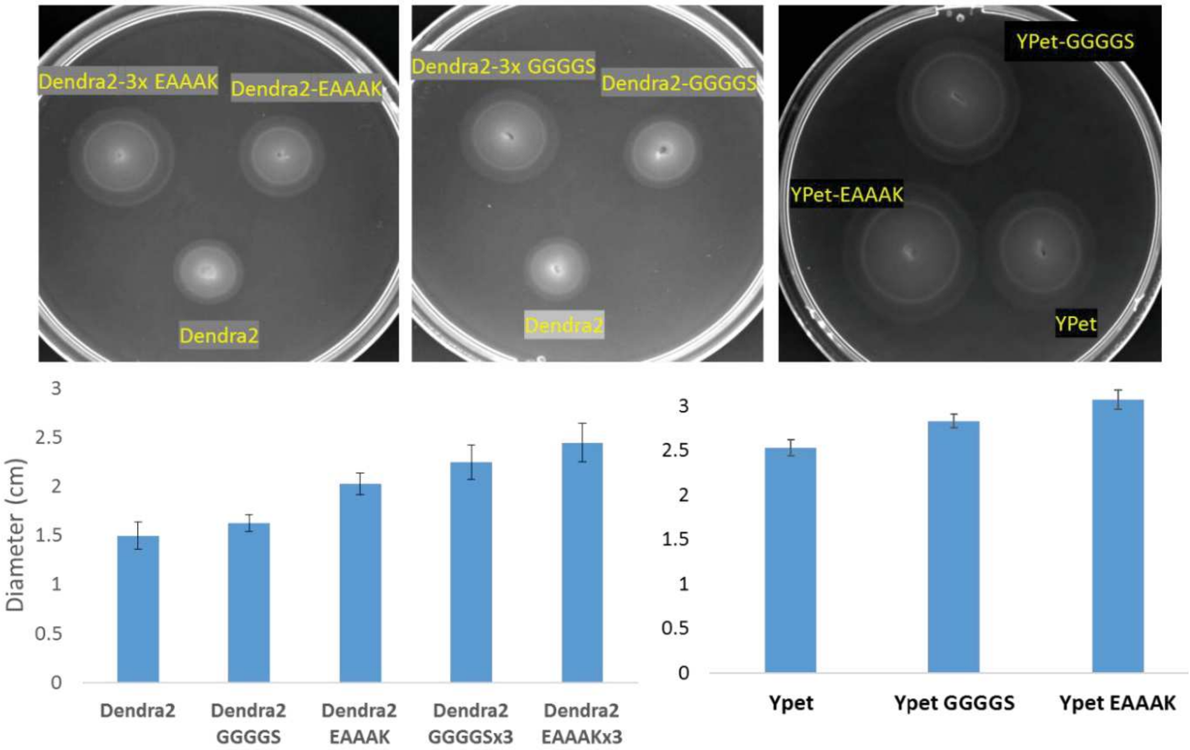


Figure 3.17. Chemotaxis speed comparisons of the Dendra2 and YPet fusion stators with linkers. Soft agar plate images are shown on the top and the averaged chemotaxis diameters of the three replicates experiments are shown in the bottom. The chemotaxis of the YPet fusion with a long and rigid linker was not tested since this fusion construct was made later. The chemotaxis assay was performed on the 0.25~0.223 % agar plates with 0.01% arabinose induction condition for 8 hours at 37 C. Error bars are standard deviations.

The improved chemotaxis ability by the presence of a linker was also demonstrated in the single motor bead assays: the four linker Dendra2 fusions increased rotation speeds (figure 3.18). Especially, the Dendra2 with a longer and rigid linker (EAAAKx3) showed distinctively improved motor rotation speeds, nearly by a factor of two in average. The fastest cell observed by this strain showed a rotation speed of ~55 Hz, demonstrating that they can generate high enough torque to the wild-type motor (~46 Hz in average). This torque generation is also reflected on the steps in the torque signal of this fusion motor (figure 3.19). In comparison to the Dendra2 fusion without a linker, the torque generation by single stator improved nearly by a factor of two in average as well: from 65 ± 38 pN nm to 113 ± 35 pN nm.

The Dendra2 motors with linkers also improved the degree of ASW. The switching events observed in Dendra2 without a linker were rather ‘pauses’ (i.e. switch to a speed close to zero) than actual reversal rotations. However, the Dendra2 with linkers could rotate in both directions; in other words, the asymmetric switch (defined above) becomes less severe. These improved asymmetric switches can be seen by the two bleu peaks in the speed histograms in figure 3.18a. Although Dendra2 showed improved asymmetric switches, the cells still could not fully restore the symmetric switches. The Dendra2 fusion with the most improved asymmetric switches was EAAAKx3 linker fusion, but its ratio of the mean CCW and CW rotation speeds were 0.18, meaning that the motors rotate at CW direction with the 18% mean rotation speed of CCW direction in average.

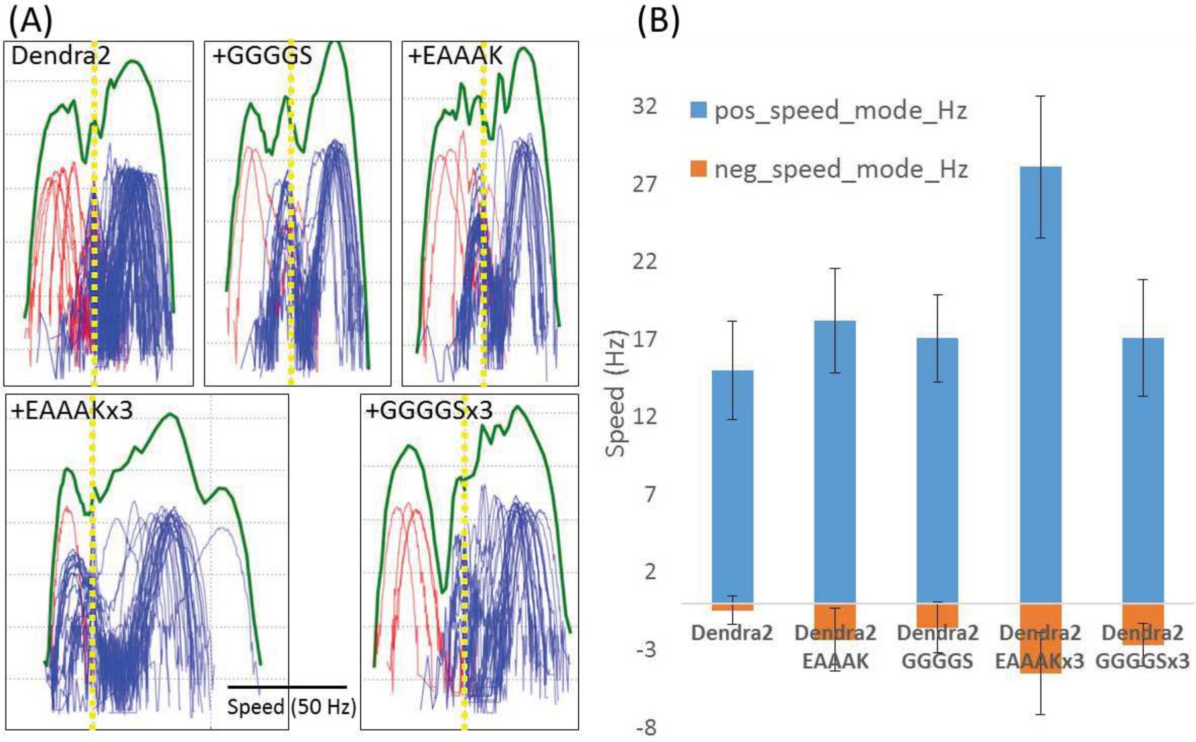


Figure 3.18. Histograms of rotation speeds of the Dendra2 motors and the Dendra2 motors with linkers (A) and their mean rotation speeds in a bar graph (B). The zero speed is indicated by yellow dotted lines. The Dendra2 motors without a linker and with CCW bias (blue) do not rotate in negative speed much (<5%). The Dendra2 motors with linkers can rotate in negative speeds (which can be seen by the small peaks in negative speed at 2~5 Hz in the histograms and the orange bars in the bar graph). The fusion with EAAAKx3 linker improved both positive and negative speed the most (with speeds of 5 Hz in CW and in 28 Hz in CCW). The number of cells measured were 54, 23, 20, 30, 30 in the right to left order in (B).

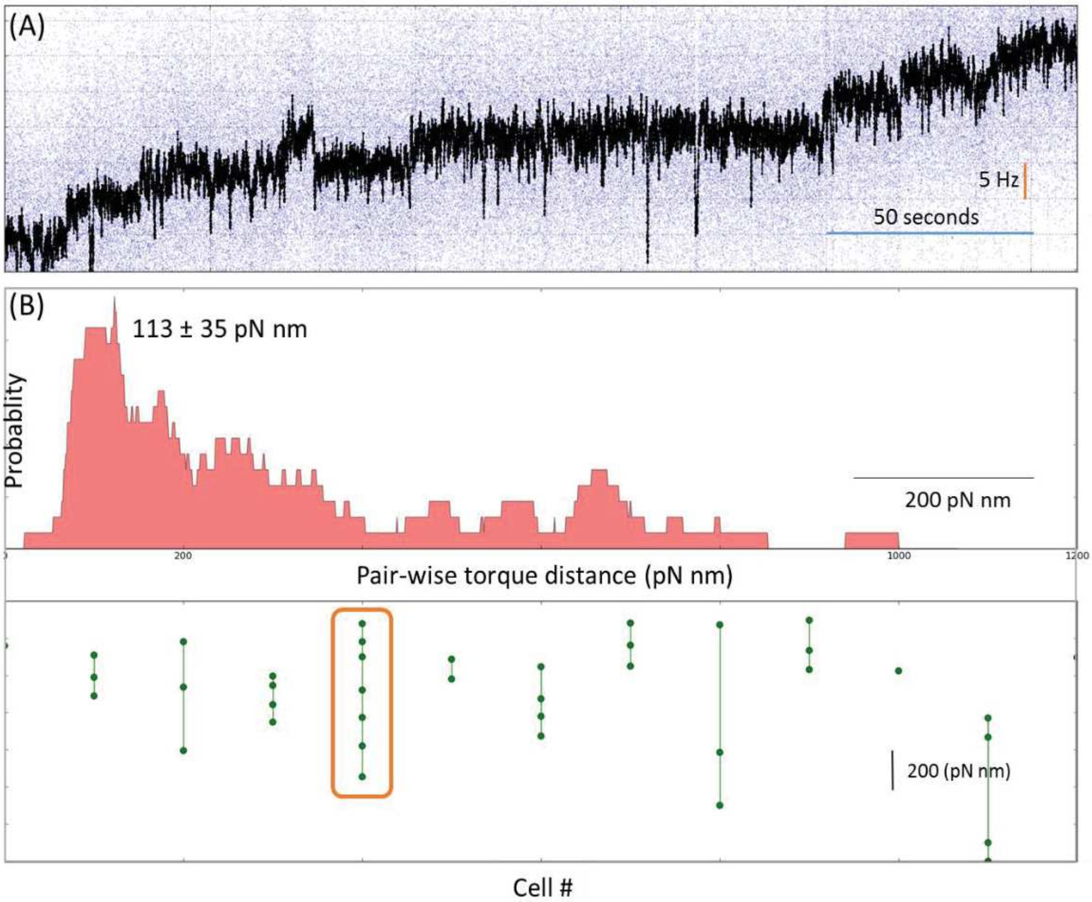


Figure 3.19. Speed per single stator of Dendra2 3x EAAAK fusion motors. (A) a step wise speed increment by a Dendra2 3x EAAAK linker motor, rotating a 1.1 µm bead. The maximum rotation speed was around 45 Hz. The speed increments for the trace in (A) are indicated in the range box in (B). (B) pair-wise (single stator) torque distance from the 10 individual Dendra2 EAAAKx3 motors.

The Dendra2 fusion motors with linkers increased their switching frequencies as well, except for the Dendra2 stator with a long and flexible linker (GGGGSx3) (figure 20). The two rigid EAAAK linkers improved the switching frequencies better than the GGGGS linkers. Consistently, Dendra2- EAAAK x3 showed the most improved switching frequency.

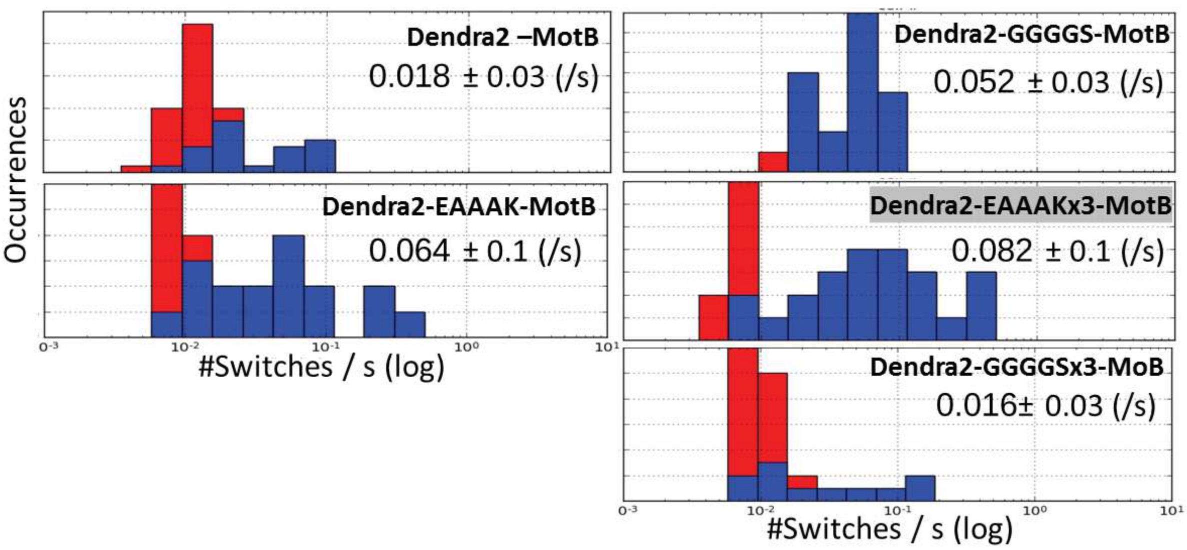


Figure 3.20 Switching frequencies of the Dendra2 fusion stators. The histograms show the distributions of the switching frequencies for each Dendra2 fusion motors, and the mean switching frequencies are shown in the histogram. The EAAAK x3 showed the highest switching frequency while GGGGS x3 showed the lowest. The red bar indicates the cells did not show switching events during the time course of recording.

The YPet fusion stators with linkers did not show a significant improvement in rotation speeds (~43 Hz in average) since the YPet motors without any linker already showed a torque similar to WT. As seen above by the Dendra2 stators with linkers, however, the long and rigid linker (EAAAKx3) consistently exhibited a positive effect on the switching mechanism. YPet-EAAAKx3-MotB fusion improved, but not fully, asymmetric switch (figure 3.21). YPet fusion without a linker showed ~27 % symmetrical rotation speeds in both directions, while YPet-EAAAKx3 fusion improved to ~ 40 %. The other two YPet fusions with short linkers (GGGGS and EAAAK) did not show such improved asymmetric switches (data not shown).

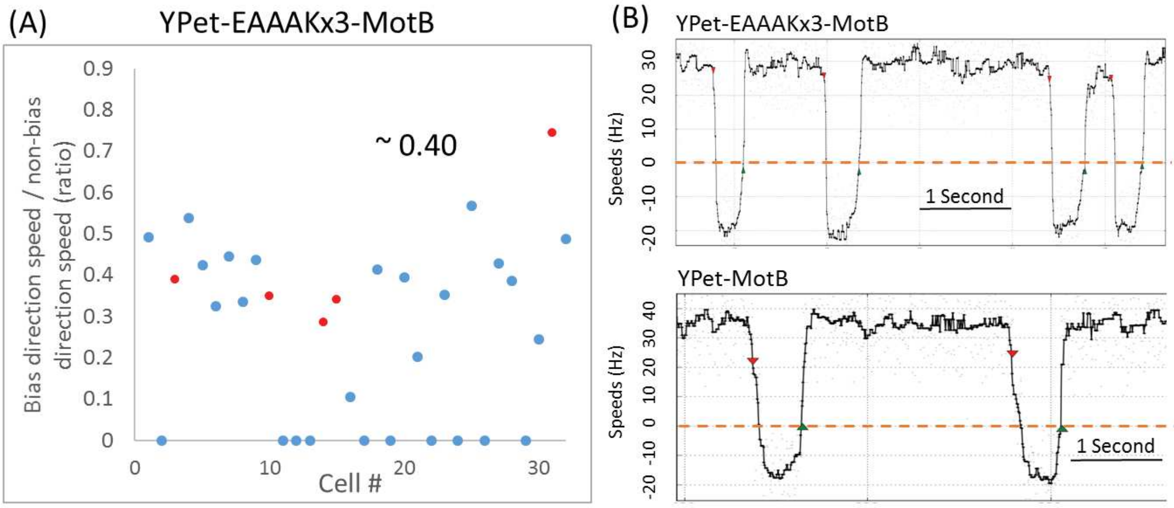


Figure 3.21 The EAAAKx3 linker improved asymmetric switching. (A) Ratio of the CW mean rotation speeds to the CCW mean rotation speeds (blue dots), and ratio of the CCW mean rotation speeds to the CW mean rotation speeds (red dots); less-bias direction over bias direction. Each dot represents individual motors. Average of the mean ratios (excluding the zero ratios) are shown on the graphs. The data is from the high induction condition cells. The same speed ratio plot for the YPet-MotB fusion is shown in fig 3.8. Having the linker EAAAKx3 improved asymmetric switches (ASW); change of the mean ratio of 0.27 to 0.40. (B) Example speed-time traces of the YPet-3xEAAAK fusion and the YPet fusion. The fusion with EAAAKx3 switched from +30 Hz to -20 Hz, while the same YPet fusion without a linker switched from +35 Hz to -15 Hz.

In addition to the improved asymmetric switching, the dwell time (resident time) in CW and CCW states of the YPet-EAAAKx3-MotB fusion motors were comparable to the WT motors. The other three YPet fusion motors (YPet-MotB, YPet-EAAAK-MotB, YPet-GGGGS-MotB) often exhibited extended CW or CCW resident time (for CCW bias or CW bias cell, respectively), whereas the YPet-EAAAKx3-MotB fusion motors did not show this extended resident time anymore (fig 3.22a). The switching frequencies of all the YPet fusions were lower than that of the WT motors. In general, having the EAAAK linkers showed a higher switching frequency than having a GGGGS linker (fig 3.22b). Taking all together, the YPet-EAAAKx3 linker motor functions like WT motor, except their not fully recovered ASW performance and reduced switching frequency. In conclusion, having a linker can improve 1) torque generation per stator, 2) asymmetric switches (but not fully) 3) switching frequencies and 4) the CW and CCW resident time.

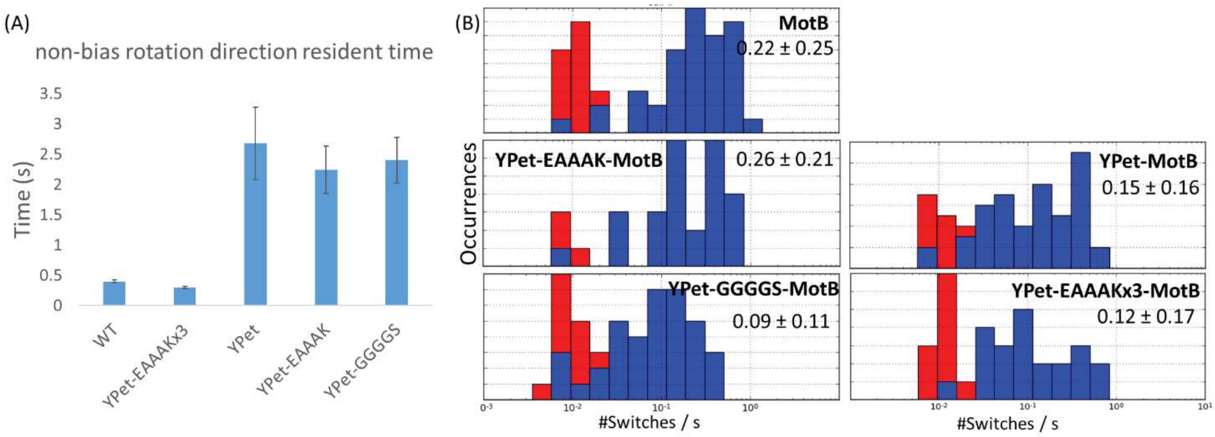


Figure 3.22 (A) The mean (non-bias rotation direction) resident time of the WT and the YPet fusion motors are shown. Non-bias rotation direction indicates either CW rotation for the CCW bias motors or CCW rotations for the CW bias motors. Its resident time is the time spent in the least visited (non-bias) direction of rotation. Unlike the other YPet fusion motors, the EAAAKx3 linker fusion of the YPet motor showed the mean non-bias rotation direction resident time of 0.3 s, which is comparable to WT. (B) The distributions of switching frequencies for each strain are shown. The numbers on the distributions indicate the mean switching frequencies. The number of cells used were: YPet (59 cells), YPet-EAAAKx3 (32 cells), YPet-GGGGS (49 cells), YPet-EAAAK (30 cells) and WT (56 cells).

3.9 Structural alignment of eGFP and YPet and Dendra2 proteins

As an initial attempt to achieve the structural understanding of how these three FP tags cause such differences in functionalities, a multiple sequence alignment of the three FPs using their structural information is performed and shown in figure 3.23a. They share a highly similar structure with 99% structural similarity. All three FPs have a typical GFP like β-barrel structure with rather flexible N- or C-terminus (figure 3.23b). The least similar sites are the N- and C-terminus regions, which both are exposed on the top surface of the folded β sheet structure, where they encounter (and may interact with) the stator proteins MotA and MotB.

```
 *
 BAD AVG GOOD
 *
eGFP          :  98
Dendra2       :  99
Ypet          :  98
cons          :   9

eGFP          MVSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTT-GKLPVPWPTLVTTLTYGVQC
Dendra2       MNTPGINLIKEDMRVKVHMEGNVNGHAFVIEGEGKGKPYEGTQTANLTVKEGAPLPFSYDILTTAVHYGNRV
Ypet          -MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKLLCTT-GKLPVPWPTLVTTLGYGVQC

cons            : * :*:.  : : *.::*:**** * :.***:*..  *. * ::  .  . **..:  *.*:: ** :


eGFP          FSRYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLVNRIELKGIDFKEDGNILGHKLE
Dendra2       FTKYPEDI--PDYFKQSFPEGYSWERTMTFEDKGICTIRSDISLEGDCFFQNVRFKGTNFPPNGPVMQKKTL
Ypet          FARYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLVNRIELKGIDFKEDGNILGHKLE

cons          *::**:.:   *:**.::****  ***: *:*.*  . *:::.:*** :.:.:.:** :*  :* :: :*


eGFP          YNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGSVQLADHYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEK
Dendra2       -KWEPSTEKLHVRD--GLLVGNINMALLLEGGGHYLCDFKTTYKAKK--VVQLPDAHFVDHRIEILGNDS-D
Ypet          YNYNSHNVYITADKQKNGIKANFKIRHNIEDGGVQLADHYQQNTPIGDGPVLLPDNHYLSYQSALFKDPNEK

cons           :::. .  : . .  . :  *:::   :*.*.  *.*.    ..     * *** *::. :  :  : . .


eGFP          RDHMVLLEFVTAAGITLGMDELYK
Dendra2       YNKVKLYEHAVARYSPLPSQVW--
Ypet          RDHMVLLEFLTAAGITEGMNELYK

cons           ::: * *. .*   .   :
```

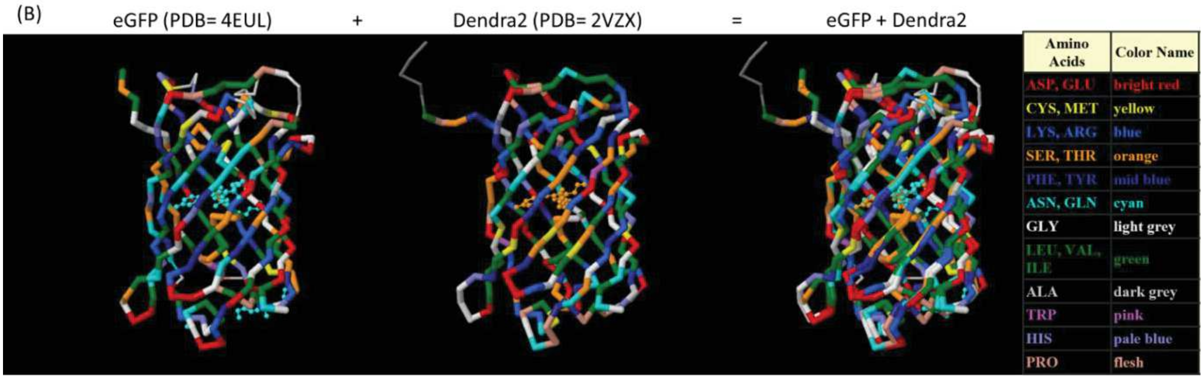


Figure 3.23. (A) Multiple sequence alignment (MSA) of eGFP, YPet and Dendra2 FP using structural information (Espresso alignment) [33]. Red portions have a high reliability and are expected to be more accurate than the rest. Consistency is estimated from a library computed using structure template-based sequence alignment. The structural templates assigned to each protein are: eGFP: 4xvpC, Ypet: 1mywA, Dendra2: 2vzxH. In this MSA, they shared 99% of the sequence similarity. (B) Structure of eGFP [8] and Dendra2 [34], and their structure alignments using the FATCAT structure alignment tool on the RCSB PDB website. The colors represent each amino acids and the table of the color scheme is shown on the right. C-terminus sites of both proteins are more flexible.

The FP-MotB fusions have such a structure that the C-terminus of FP (end of FP) is directly connected to the N-terminus of MotB (beginning of MotB) (Figure 3.2a). The sequence alignment of the fusion site (the short stretches of C-terminus endings together with the N-terminus of MotB) show that higher similarity between eGFP and YPet than Dendra2, suggesting a possible reason for the functional similarity found between eGFP and YPet fusion than that of Dendra2 fusion (figure 3.24). Several studies [35-37] reported that having a N-terminus tag or a C-terminus tag can cause functional differences of the protein of interest.

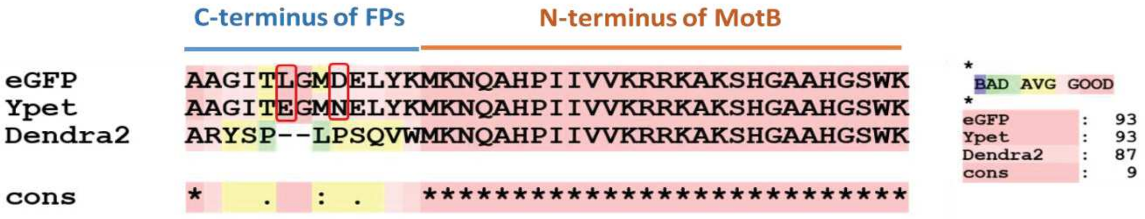


Figure 3.24. Sequence alignments of the sites connecting FP and MotB protein. eGFP and Ypet shows the identical sequences except the two columns (red boxes), Dendra2, on the other hand, shows less similarity.

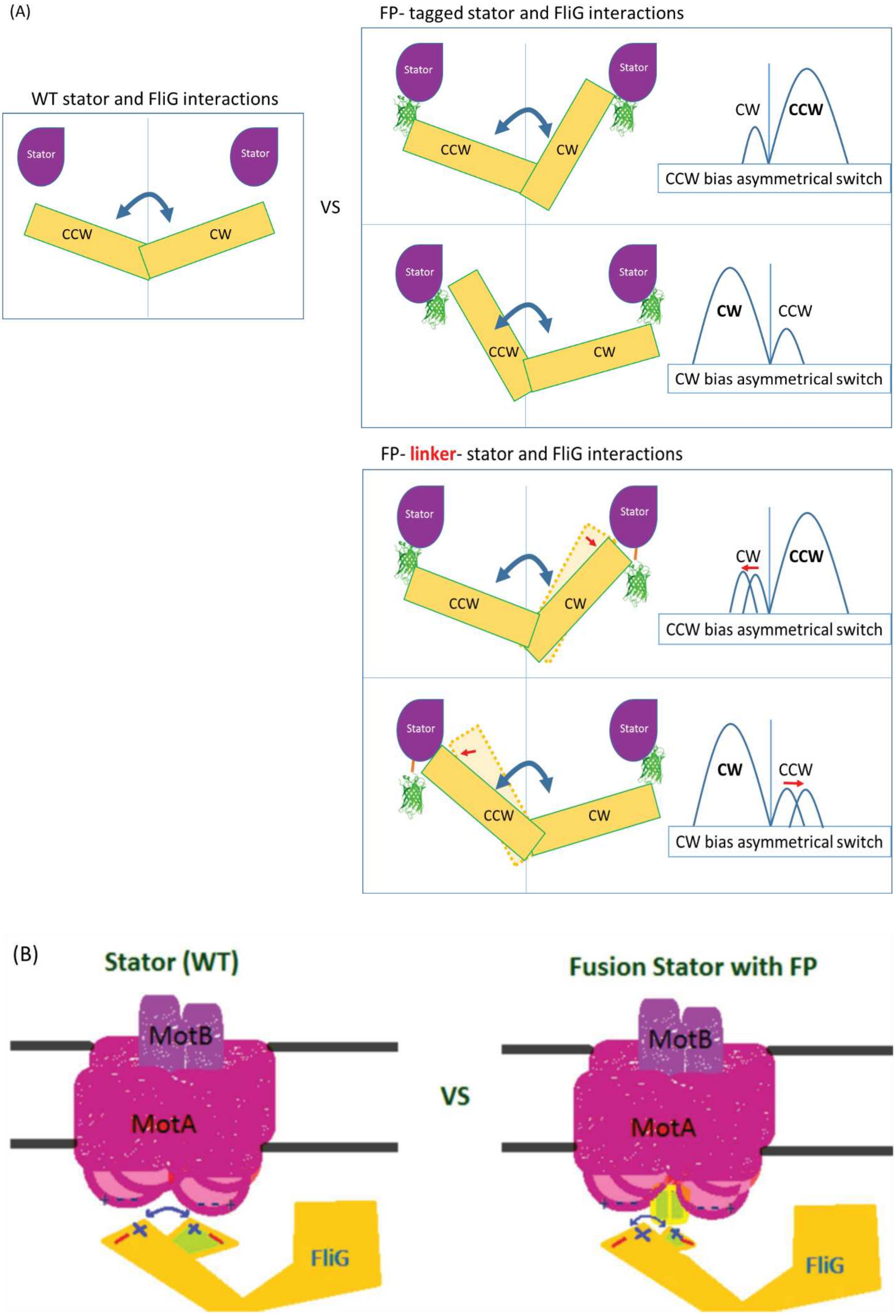
3.10 Discussion

In order to elucidate the underlining mechanism of the impaired functions in presence of the fusion stators, we can conceive different possible scenarios, which are based on well accepted two underlining mechanisms [38-40]: first, torque generation is governed by the interaction between the charged-residues in both the rotor protein FliG (helix C5 of the C-terminus) and the cytoplasmic loops of MotA [38, 39], second, a switch from CCW to CW rotation is induced by the distinct conformational change of the FliG protein, so called closed and open states of FliG protein [40-43].

The first observation was the symmetric torque reduction (STR) due to the reduced torque generation by single fusion stators. Possible mechanisms which can explain STR include 1) eGFP and Dendra2 tags (proteins with electrostatic potentials) interfere with the torque generating cytoplasmic region of MotA so that the stator and rotor interactions (torque generation mechanism) became less efficient. YPet tag however did not interfere, as significantly as the other two tags, with this region of MotA. 2) the FP tags interfere with the torque generating region of FliG (helix C5 of C-terminus domain) more significantly, rather than with MotA. 3) the FP tags interfere with the proton translocation mechanism (although this is less likely since the FP tag is located away from the putative pore site).

A second observation was the inability to achieve fully symmetric switching in tagged motors (i.e. the observation of ASW), which is observed both in the CW and the CCW biased motors. Possible causes of such asymmetric switch include 1) the FP tags directly block or interfere with the FliG (~180 °C) conformational change, which is required to go from the CCW (open) state to CW (closed) state, and vice versa for the change from CW to CCW. This possibility is based on the several FliG structure models that the switch is induced by the FliG conformational change [40-43]. In this first scenario that the FP tags prevent the FliG conformational change, we propose that the FliG can remain in between the fully CW or CCW state. Such an intermediate conformation of FliG is functional and leads to a reduced speed with respect to the fully open or close conformations, creating ASW (figure 3.24). 2) Following [17], it is conceivable that only two out of the four MotA units are specifically responsible for the CCW rotation (while the other two units apply CW torque uniquely). In this scenario, the presence of two FPs could cause significant structural perturbations for only one of the MotA pairs in a stator, while the FliG switching conformational change can be considered not impaired. For example, when the CCW state of the FliG ring is interacting with the less impaired MotA pair, the motor works properly, but when the FliG ring switches to the CW state, it interacts with the impaired MotA pair, which results in reduced torque generation in CW rotation. The scenario in 2) is more complex than that in 1), because, in this scenario, the stator should have a preferred and always identical binding orientation around the rotor, to affect systematically the speed in one direction more than in the other. In addition, this second scenario seems to be less likely, because ASW was observed from both CCW and CW based motors, affecting symmetrically. In case that stators can bind in any orientation around the rotor, for this scenario to be true, we should observe *also* another kind of asymmetric switches (i.e. CCW biased motors rotating at low CCW speed and switching to high CW speed), which we did not observe.

In summary, we can propose the following mechanisms to explain our observations (this is an initial attempt to explain the impaired motor functions observed in presence of the FP fusion stators, which requires further experiments to validate). First, to explain the symmetric torque reduction (STR), the FP tags (which is composed of charged molecules creating electrostatic potentials) interfere with the torque generating stator-rotor interface, so that the stator and rotor interactions became less efficient. The interface here is in between the charged residues in the cytoplasmic loops of MotA and the helix C5 of the C-terminus domain of FliG. Alternatively, the proton translocation mechanism could be impaired by the tag. To account for the asymmetric switching (ASW), the FP tags interfere with the conformational change of the FliG unit in contact with the stator (either by directly blocking this process or by indirectly interfering with other components in the C-ring switch complex) to prevent the complete switches between the fully open and closed conformations (Figure 3.25). A set of intermediate and functional conformations between the fully open and fully closed FliG state is expected in this proposed mechanism to explain ASW.


(A)
WT stator and FliG interactions
Stator
Stator
CCW
CW
VS
FP- tagged stator and FliG interactions
Stator
Stator
CCW
CW
CW
CCW
CCW bias asymmetrical switch
Stator
Stator
CCW
CW
CW
CCW
CW bias asymmetrical switch
FP- linker- stator and FliG interactions
Stator
Stator
CCW
CW
CW
CCW
CCW bias asymmetrical switch
Stator
Stator
CCW
CW
CW
CCW
CW bias asymmetrical switch
(B)
Stator (WT)
MotB
MotA
FliG
VS
Fusion Stator with FP
MotB
MotA
FliG

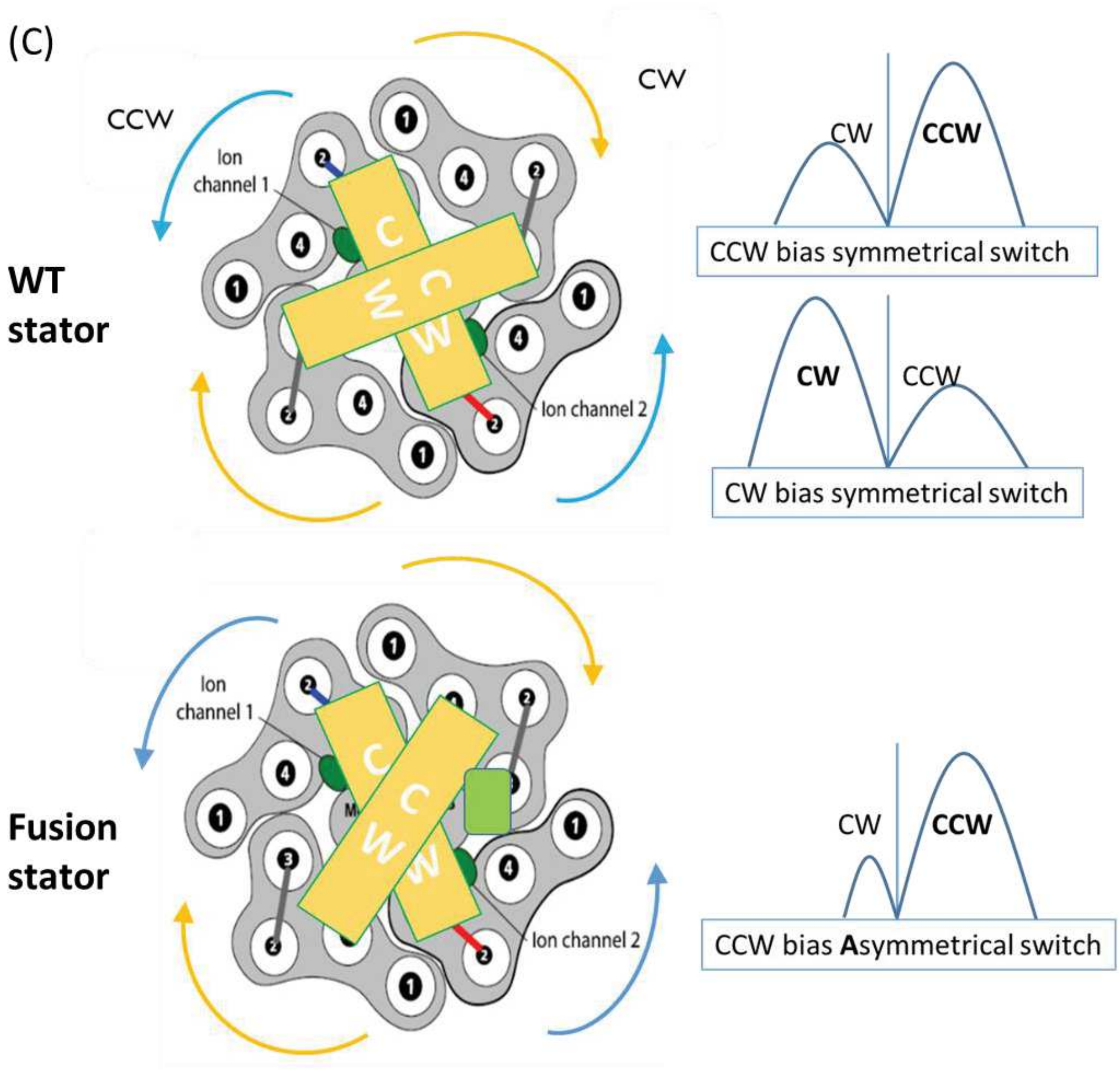


Figure 3.25 An illustration of the interaction between the fusion stator and FliG switching complex. When a FP is tagged to a stator complex, the FP tag inhibits the FliG switching complex to transit from CCW state to CW (or vice versa), this interference can be a possible cause of the asymmetric switches. (A) and (B) illustrate the same scenario. (C) illustrates this scenario in coupled with the mechanics of the torque generation [17] model. The green rectangle represents the FP tag.

Lowered switching frequencies were also observed in presence of the FP tag. Interestingly, these lowered switching frequencies showed a correlation with the reduced torque generation: YPet fusion showed the highest (among the three fusion stators) average torque generated by a single stator, followed by eGFP and Dendra2. This order in torque reduction is also observed in the switching frequencies; YPet fusion (~0.32/sec) switched most frequently among the three fusions, followed by eGFP (~0.14 /sec) and Dendra2 (~0.02 /sec). In other words, lowered torque is accompanied by lowered switching frequency. This correlation between the reduced torque and reduce switching frequencies may suggest that the stator and rotor interaction can play an important role in the switching dynamics, as it has been suggested by [28]. In addition, FP tags caused extended resident time (>1s) in less-biased rotational direction (in 1.1 μm bead assay condition), and such long resident time was not observed by the WT motors. This implies that how FP tag near the stator and FliG interface can influence all functional aspects of the switching dynamics.

The fusion stators with a linker sequence between the FP and the MotB were also constructed and tested. Having a linker improved the chemotaxis motility as well as the various aspects of the single motor functions. Interestingly, rigid linkers (EAAAK) always improved the impaired functions of the motor better than flexible linkers (GGGGS), especially the longer version of EAAAK linker (EAAAKx3 three times repeated EAAAK) improved all aspects of the functions of the motor: torque generation per single stator, asymmetric switches, less-bias resident time and switching frequency. The end to end length of the EAAAK linker is shorter than that of the GGGGS linker, forming a coiled α-helix, which is a rigid and stable structure. It is speculated that this type of linkers serves as rigid spacers to effectively separate protein domains, and to reduce their unfavorable interactions [32]. The GGGGS linkers, on the other hand, allow certain degree of flexibility, movement and interactions of the protein domains [32]. Consequently, the FP tag with the flexible linker probably remains in the vicinity and interfere with the stators more than the rigid linker might. The long and rigid linker (EAAAKx3) may increase the space between the stator and the FP tag, decreasing interference in interactions. Taken together, the improved functions of the motor by the rigid linker may provide an insight in the localization of the stators in the motor and their interactions with the FliG rotor components. In addition, these observations underline how the location and flexibility of the FP tag around the stator-FliG interface area can influence on the various functional dynamics of the motor.

References

1. Leake, M. C., Chandler, J. H., Wadhams, G. H., Bai, F., Berry, R. M., & Armitage, J. P. (2006). Stoichiometry and turnover in single, functioning membrane protein complexes. Nature, 443(7109).

2. Lele, P. P., Hosu, B. G., & Berg, H. C. (2013). Dynamics of mechanosensing in the bacterial flagellar motor. Proceedings of the National Academy of Sciences of the United States of America, 110(29).

3. Tipping, M. J., Delalez, N. J., Lim, R., Berry, R. M., & Armitage, P. (2013). Load-Dependent Assembly of the Bacterial Flagellar Motor. Mbio.

4. Morimoto, Y. V, Nakamura, S., Kami-ike, N., Namba, K., & Minamino, T. (2010). Charged residues in the cytoplasmic loop of MotA are required for stator assembly into the bacterial flagellar motor. Molecular Microbiology, 78(5), 1117–29.

5. Tipping, M. J., Steel, B. C., Delalez, N. J., Berry, R. M., & Armitage, J. P. (2013). Quantification of flagellar motor stator dynamics through in vivo proton-motive force control. Molecular Microbiology, 87(2), 338–347.

6. Paulick, A., Koerdt, A., Lassak, J., Huntley, S., Wilms, I., Narberhaus, F., & Thormann, K. M. (2009). Two different stator systems drive a single polar flagellum in Shewanella oneidensis MR-1. Molecular Microbiology, 71(4), 836–850.

7. Chandler, J.H. (2006) Characterization of the stator complexes of the *Escherichia coli* flagellar motor (Doctoral dissertation). Oxford University.
8. Arpino, J. A. J., Rizkallah, P. J., & Jones, D. D. (2012). Crystal Structure of Enhanced Green Fluorescent Protein to 1.35 Resolution Reveals Alternative Conformations for Glu222. PLoS ONE, 7(10).
9. Nguyen, A. W., & Daugherty, P. S. (2005). Evolutionary optimization of fluorescent proteins for intracellular FRET. *Nature Biotechnology*, *23*(3), 355–60.

10. Chudakov, D. M., Lukyanov, S., & Lukyanov, K. a. (2007). Using photoactivatable fluorescent protein Dendra2 to track protein movement. *BioTechniques*, *42*(5), 553–565.

11. Gurskaya, N. G., Verkhusha, V. V, Shcheglov, A. S., Staroverov, D. B., Chepurnykh, T. V, Fradkov, A. F., Lukyanov, K. a. (2006). Engineering of a monomeric green-to-red photoactivatable fluorescent protein induced by blue light. *Nature Biotechnology*, *24*(4), 461–5.

12. Shaner, N. C., Steinbach, P. A., & Tsien, R. Y. (2005). A guide to choosing fluorescent proteins. Nature Methods, 2(12).

13. Dean, K. M., & Palmer, A. E. (2014). Advances in fluorescence labeling strategies for dynamic cellular imaging. *Nature Chemical Biology*, *10*(7), 512–23.

14. Daya, R. N., & Davidson, M. W. (2010). The fluorescent protein palette: tools for cellular imaging. Chem Soc Rev, 38(10), 2887–2921.
15. Kojima S1, Blair DF. Solubilization and purification of the MotA/MotB complex of Escherichia coli.Biochemistry. 2004 Jan 13;43(1):26-34.

16. Timothy F. Braun , Laith Q. Al-Mawsawi , Seiji Kojima , and David F. Blair (2004). Arrangement of Core Membrane Segments in the MotA/MotB Proton-Channel Complex of Escherichia coli. Biochemistry. 43(1)
17. Mandadapu, K. K., Nirody, J. A., Berry, R. M., & Oster, G. (2015). Mechanics of torque generation in the bacterial flagellar motor. *Proceedings of the National Academy of Sciences*, *112*(32), E4381–E4389.

18. Parkinson, J. S. (1978). Complementation analysis and deletion mapping of Escherichia coli mutants defective in chemotaxis. *135*(1), 45–53.

19. Guzman, L. M., Belin, D., Carson, M. J., Beckwith, J., Guzman, L., Belin, D., & Carson, M. J. (1995). Tight regulation , modulation , and high-level expression by vectors containing the arabinose PBAD promoter, *177*(14).

20. Amann, E., Ochs, B., & Abel, K. J. (1988). Tightly regulated tac promoter vectors useful for the expression of unfused and fused proteins in Escherichia coli. *Gene*, *69*, 301–315.
21. Happel, J. & Brenner, H. Low Reynolds Number Hydrodynamics (Kluwer Academic, Drodrecht, 1991).
22. Reid, S. W., Leake, M. C., Chandler, J. H., Lo, C.-J., Armitage, J. P., & Berry, R. M. (2006). The maximum number of torque-generating units in the flagellar motor of Escherichia coli is at least 11. Proceedings of the National Academy of Sciences of the United States of America, 103(21), 8066–71.
23. Block, S., & Berg, H. (1984). Successive incorporation of force-generating units in the bacterial rotary motor. *Nature*.
24. Inoue, Y., Baker, M. a B., Fukuoka, H., Takahashi, H., Berry, R. M., & Ishijima, A. (2013). Temperature dependences of torque generation and membrane voltage in the bacterial flagellar motor. *Biophysical Journal*, *105*(12), 2801–10.
25. Gabel, C. V, & Berg, H. C. (2003). The speed of the flagellar rotary motor of Escherichia coli varies linearly with protonmotive force. *Proceedings of the National Academy of Sciences of the United States of America*, *100*, 8748–8751.
26. Korobkova, E. A., Emonet, T., Park, H., & Cluzel, P. (2006). Hidden stochastic nature of a single bacterial motor. *Physical Review Letters*, *96*(5), 1–4.
27. Bai, F., Branch, R. W., Nicolau, D. V, Pilizota, T., Steel, B. C., Maini, P. K., & Berry, R. M. (2010). Conformational spread as a mechanism for cooperativity in the bacterial flagellar switch. *Science (New York, N.Y.)*, *327*(5966), 685–9.
28. Bai, F., Minamino, T., Wu, Z., Namba, K., & Xing, J. (2012). Coupling between switching regulation and torque generation in bacterial flagellar motor. Physical Review Letters, 108(17), 1–4.
29. Van Albada, S. B., Tănase-Nicola, S., & ten Wolde, P. R. (2009). The switching dynamics of the bacterial flagellar motor. *Molecular Systems Biology*, *5*(316), 316.
30. Bacterial flagellar switching under load. (2003). *Nature*, *423*(June), 2003.
31. Wang, F., Yuan, J., & Berg, H. C. (2014). Switching dynamics of the bacterial flagellar motor near zero load. *Proceedings of the National Academy of Sciences*.
32. Chen, X., Zaro, J. L., & Shen, W.-C. (2013). Fusion protein linkers: property, design and functionality. *Advanced Drug Delivery Reviews*, *65*(10), 1357–69.
33. Ridgeway, T., & Nw, L. (2000). T-Coffee : A Novel Method for Fast and Accurate Multiple Sequence Alignment.
34. Adam, V., Nienhaus, K., Bourgeois, D., & Nienhaus, G. U. (2009). Structural basis of enhanced photoconversion yield in green fluorescent protein-like protein Dendra2. *Biochemistry*, *48*(22), 4905–4915.
35. Stadler C1, Rexhepaj E, Singan VR, Murphy RF, Pepperkok R, Uhlén M, Simpson JC, Lundberg E. (2013) Immunofluorescence and fluorescent-protein tagging show high correlation for protein localization in mammalian cells. Nature methods.
36. Palmer, E., & Freeman, T. (2004). Investigation into the use of C- and N-terminal GFP fusion proteins for subcellular localization studies using reverse transfection microarrays. *Comparative and Functional Genomics*, *5*(4), 342–353.
37. Sourjik, V., & Berg, H. C. (2000). Localization of components of the chemotaxis machinery of Escherichia coli using fluorescent protein fusions. Molecular Microbiology, 37(4), 740–751.
38. Lair, D. A. F. B. (1998). Electrostatic interactions between rotor and stator in the bacterial flagellar motor, *95*(May), 6436–6441.
39. Morimoto, Y. V, Nakamura, S., Hiraoka, K. D., Namba, K., & Minamino, T. (2013). Distinct roles of highly conserved charged residues at the MotA-FliG interface in bacterial flagellar motor rotation. *Journal of Bacteriology*, *195*(3), 474–81.
40. Lee, L. K., Ginsburg, M. a, Crovace, C., Donohoe, M., & Stock, D. (2010). Structure of the torque ring of the flagellar motor and the molecular basis for rotational switching. *Nature*, *466*(7309), 996–1000.

41. Lam, K.-H., Ip, W.-S., Lam, Y.-W., Chan, S.-O., Ling, T. K.-W., & Au, S. W.-N. (2012). Multiple conformations of the FliG C-terminal domain provide insight into flagellar motor switching. Structure (London, England: 1993), 20(2), 315–25.
42. Brown, P. N., Hill, C. P., & Blair, D. F. (2002). Crystal structure of the middle and C-terminal domains of the flagellar rotor protein FliG. *The EMBO Journal*, *21*(13), 3225–34.
43. Vartanian, A. S., Paz, A., Fortgang, E. a, Abramson, J., & Dahlquist, F. W. (2012). Structure of flagellar motor proteins in complex allows for insights into motor structure and switching. *The Journal of Biological Chemistry*, *287*(43), 35779–83.
44. Yonekura, K., Maki-Yonekura, S., & Homma, M. (2011). Structure of the flagellar motor protein complex PomAB: implications for the torque-generating conformation. *Journal of Bacteriology*, *193*(15), 3863–70.

Chapter 4. Single molecule fluorescence microscopy

4.1 Rationale and aims

The stators of the BFM (each a protein complex formed by 4 MotA and 2 MotB) freely diffuse in the inner membrane, and bind and unbind to the motor complex. Understanding this dynamical movement of the stator complexes remains one of the main interests in the field. One promising approach to understand this phenomenon is single-molecule fluorescence techniques, as it enables us to probe various processes on the level of individual molecules in live bacterial cells [1], such as individual stator complexes in the motor. For example, in super-resolution microscopy, being able to detect single molecule fluorescence in live cell can provide high spatial (~30 nm) resolution information of the stators localized around the motors [2]. In addition, the fluorescence signals from the three FP fusion stator motors were imaged and compared quantitatively in both Epi and TIRF microscopies.

Therefore, in this study, we aim to

1. visualize the three FPs attached to the stator protein MotB in live *E.coli* cells in both Epi illumination and TIRF illumination, and to check and compare the fluorescence signals from the three FP-MotB fusion protein constructs.
2. quantitatively compare the fluorescence signals of the cells grown in two different gene inducers [L-arabinose] concentrations.
3. image photo-converted fluorescent protein Dendra2 fused to MotB protein (by photo-activated localization microscopy, PALM).
4. establish an experimental-set up that can simultaneously record the fluorescence signals from the cell and the bead rotation (simultaneous dual recording).

4.2 Localization of the FP tagged stators

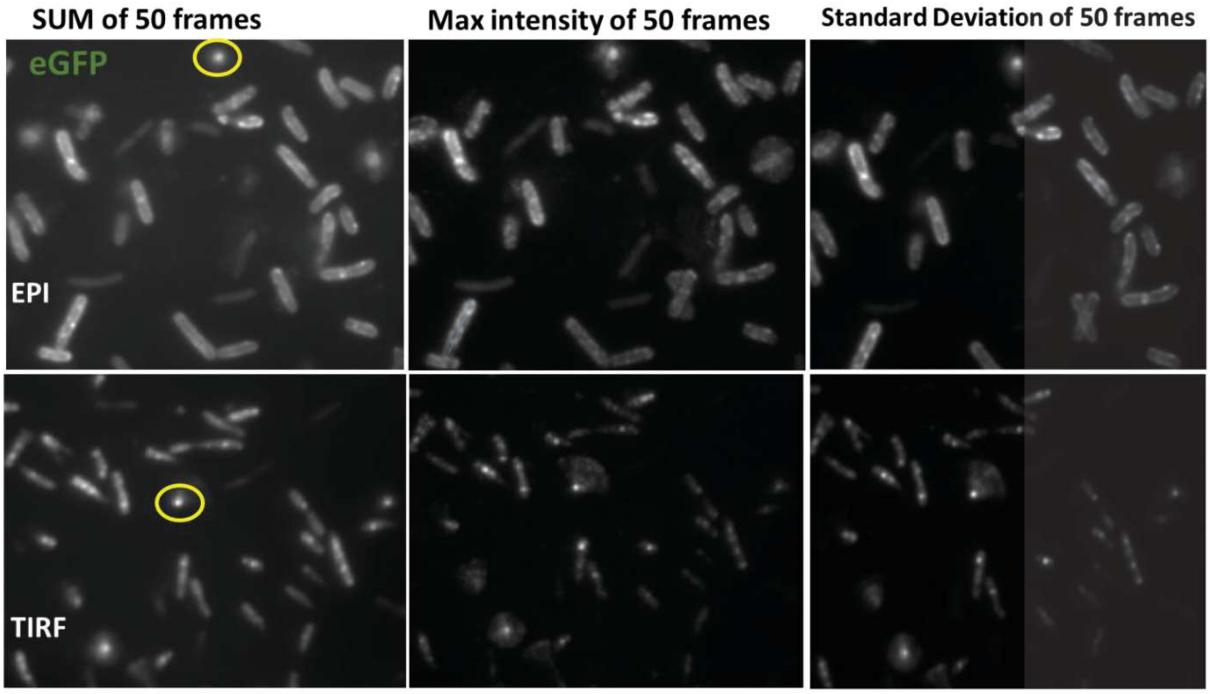
The fluorescence signals of both immobilized cells and rotating tethered cells, expressing the FP (eGFP, YPet and unconverted Dendra2)-MotB fusion proteins, were imaged by Epi- and TIRF 488 nm laser illuminations using the microscope described in chapter 2.5. The negative control cells express the wild-type (WT) MotB protein (without the FP) by the same expression vector (pBAD33). The cells were grown in the same condition as for the bead assays that described in chapter 2.4.3. The data from the Andor EMCCD camera was saved as a stack of 512x512 pixel tiff images, then the fluorescence intensities from the raw image files were computed in the open-source biological image analysis software Fiji imageJ [3].

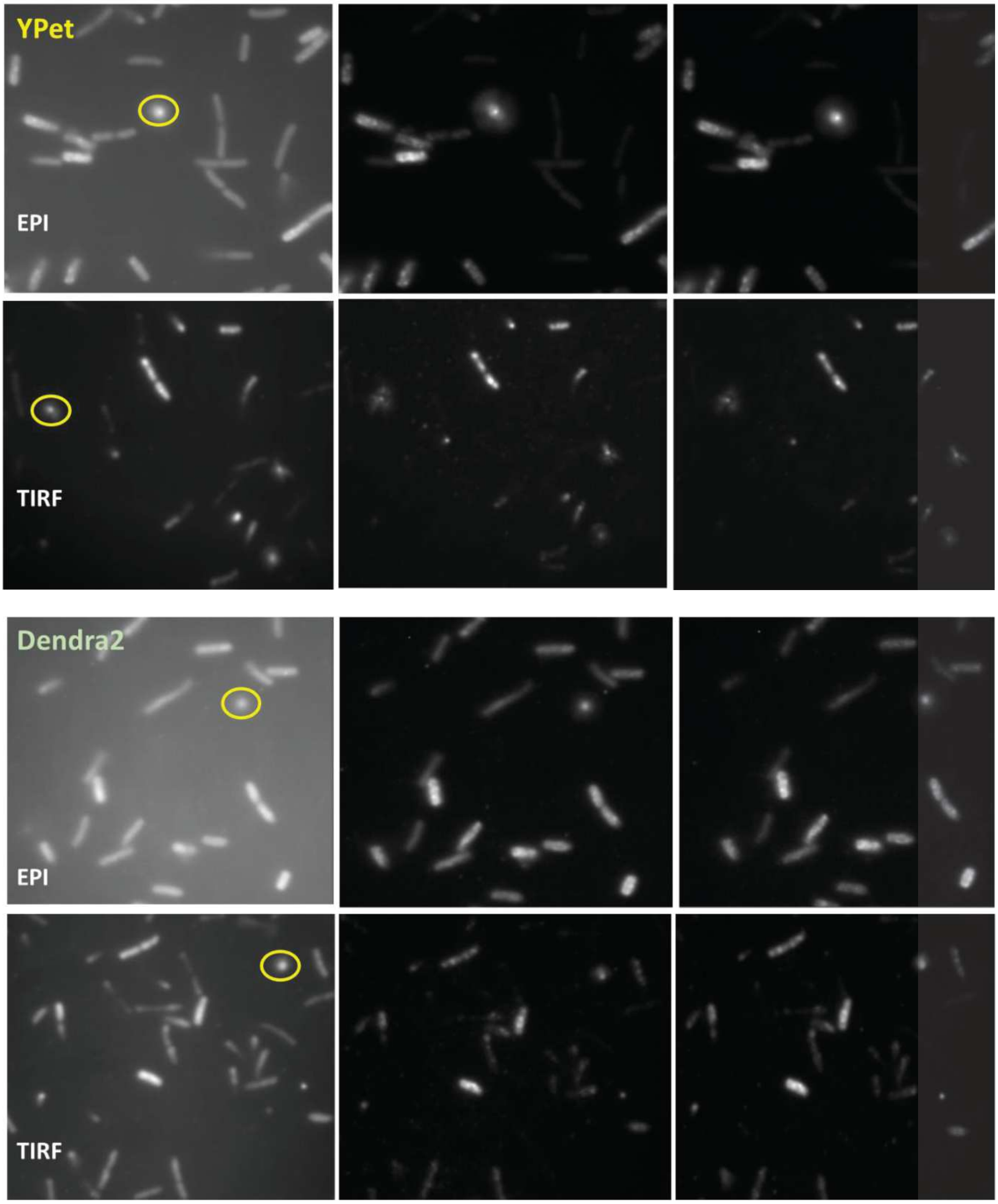
Two internal fluorescence intensities were measured: cell body fluorescence intensity and motor spot fluorescence intensity. The cell body fluorescence intensity, which includes auto-fluorescence of the cell and the fluorescence from the diffusing FPs, was estimated by taking the mean intensity of the 3x3 (pixel) region of interest (ROI) where no clusters of fluorescent spot appear in the cell. Three ROIs per cell were considered to measure an accurate cell body fluorescence intensity to average out the inhomogeneity in fluorescence. The motor spot fluorescence intensity was estimated by taking the maximum intensity of the 3x3 ROI, where they show a cluster of the bright fluorescence intensity. The fluorescence intensity of the background of the slide (instrumental noise) was low and similar between each sample. All the fluorescence intensity and image comparisons were strictly made by the images taken on the same day to ensure that the alignment and microscope setting were identical.

4.2.1 Fluorescence images of the four strains in Epi and TIRF illumination

The fluorescence signals excited by the 488 nm laser from the cells stuck on the slides showed both concentrated clusters in the membrane as expected for the motor (1~6 clusters per cells) and a more homogeneous fluorescence over the entire cell body, likely due to the diffusing population of stators in the membrane. The percentage of finding spinners per field of view were all identically high (>15%) for all four strains. In order to confirm that the fluorescence signals were actually attributable to well folded functional FP-MotB in the motor, the fluorescent spots in the rotating cell body (tethered cell assay) of the three fusion stator strains were imaged and compared to their negative control cells which holds the same condition but expresses MotB protein without the FP tag. A bright fluorescence signal was found in the center of the rotating cell body from all three FP tagged MotB proteins expressing cells, while negative control cells produce a much lower fluorescence (figure 4.1). This result supports that those spots were actually clustered FP-MotB fusion proteins around the motor.

These fluorescence signals at the center of the rotation from the tethered cells were detected in both Epi and TIRF illumination. TIRF mode (having the light coming through the objective with a glancing angle of incidence) improved visualizing the fluorescence spots in the cells. The figure 4.1 shows examples fluorescence images of the four strains in both Epi and TIRF illumination. The images are the sum. max and std of the stacks of 50 frames (100 ms/frame). In those images, the cells visualized in TIRF appear smaller cells than in Epi mode, since their peripheral membrane sides were not illuminated. The cells in these images were expressed by 0.002% arabinose (low gene induction condition).




YPet
EPI
TIRF
Dendra2
EPI
TIRF

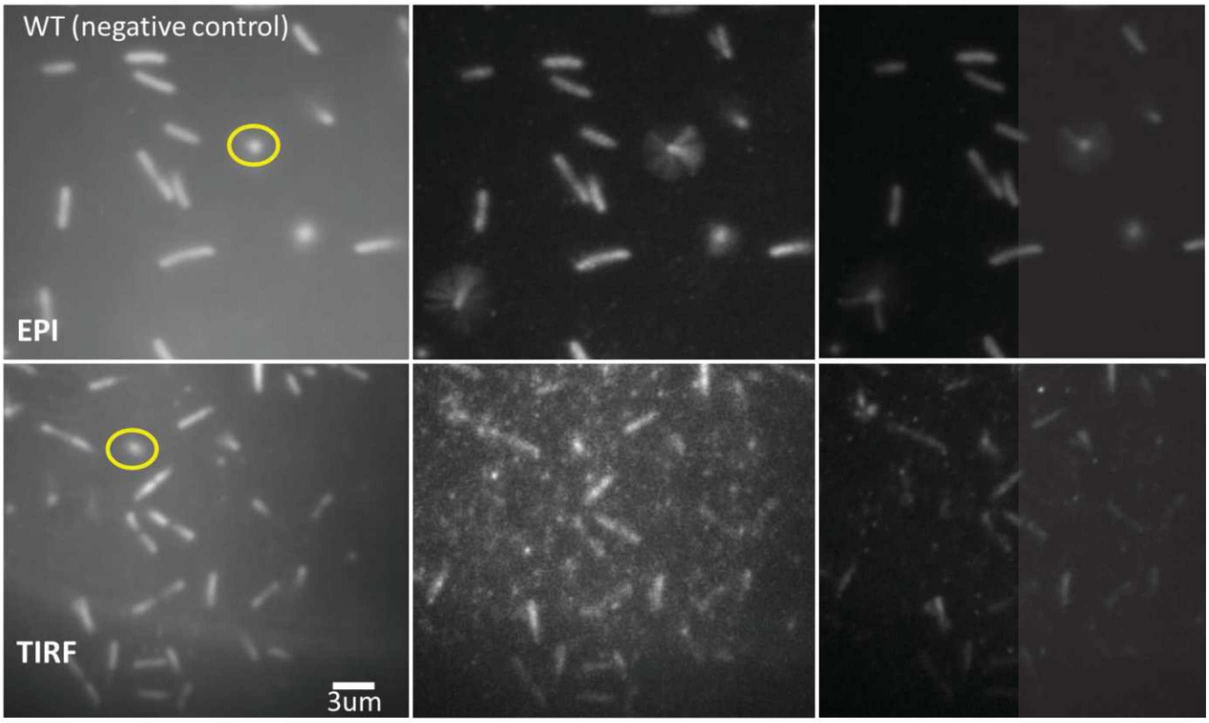
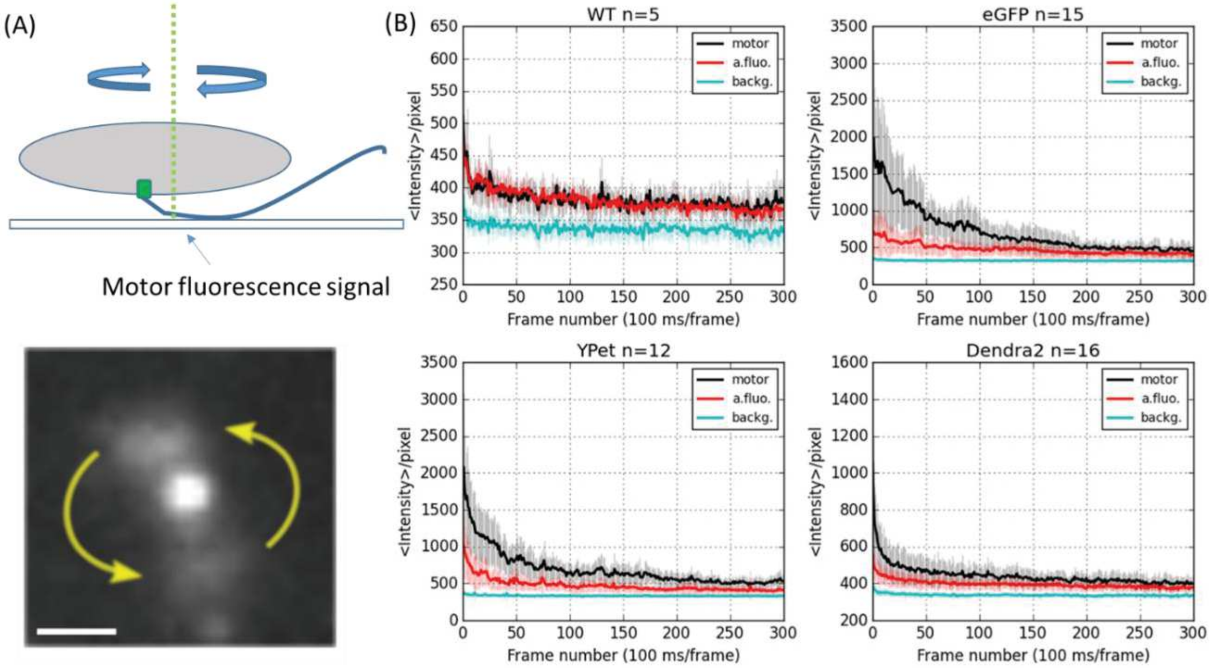


Figure 4.1 Example fluorescence images of the four strains in both Epi and TIRF modes. Brightness and contrast of the images were set by the auto-scale mode. The cells, expressing FP-MotB or MotB by low induction (0.002% arabinose), were either stuck or tethered on the glass slide. Some of the rotating tethered cells are indicated by yellow circles. The images were recorded at 100 ms/frame for 300 frames, and here the images are shown as SUM, MAX and STD (standard deviation) of the first 50 stacks of 300 frames. The rotating cells are depicted as dots with some cell body fluorescence signals around the dots. The stuck cells from the three FP-fusion strains show fluorescence spots in the cells, while no fluorescence spot formation is found from the WT strain.

4.2.2 Fluorescence signals at the center of rotation

The motor fluorescence intensities of the rotating cells (schematically shown in fig.4.2a) displayed in figure 4.1 were compared, and their photo-bleaching time trace are shown in figure 4.2b and 4.2c. Here, only the fluorescence spots from the center of rotating cells were taken as the motor fluorescence signal, which were calculated by mean intensity of the region of interest (ROI) of 3x3 pixel at the center of rotation. The cell body fluorescence background (auto-fluorescence) was calculated by 3x3 ROIs centered on several (typically 20-30) stuck cells in regions of the cell body not displaying fluorescent spots or clusters. The background signal was calculated by several 3x3 ROIs on the glass, and was rather constant throughout the whole frames. The cell body auto-fluorescence signal comparison of those four strains revealed that Ypet cells have the highest auto-fluorescence signal, followed by eGFP, Dendra2 and WT showing the lowest intensity. However, when the signal of interest divided by the noise level (SNR) was compared, eGFP showed the highest SNR (figure 4.2d).



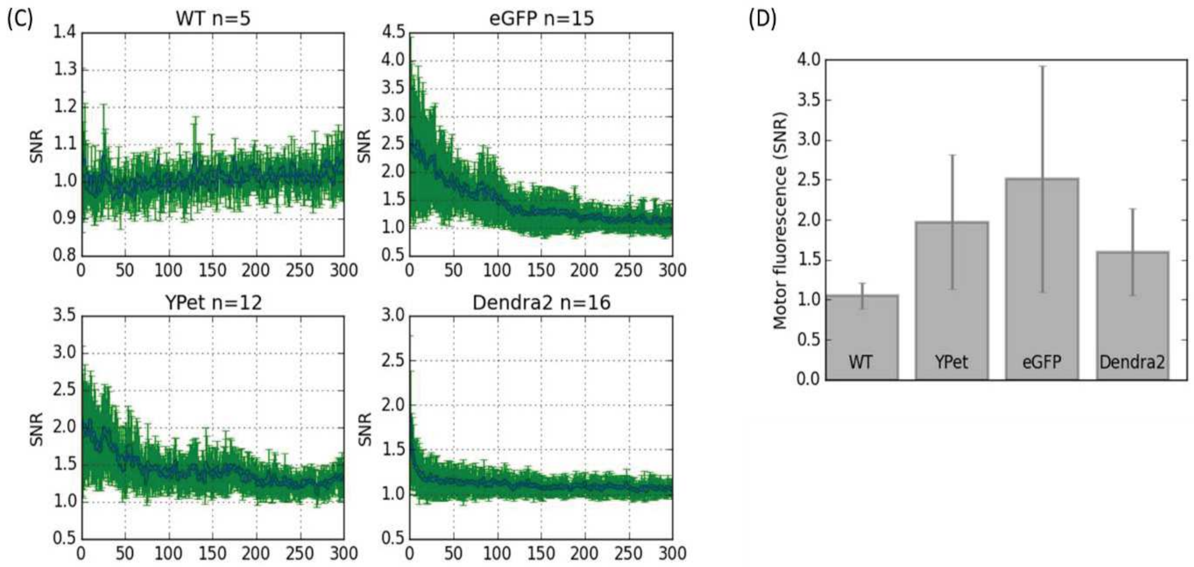


Figure 4.2 Motor fluorescence intensity of the four strains. (A) a schematic of a rotating tethered cells and an example image of a rotating cell with a fluorescence signal in the center of the rotation (eGFP strain). (B) the fluorescence intensity of the motor spot, the auto-fluorescence and the background is plotted over 30 seconds. The number of cells used for averaging are shown on the graphs. (C) the data on (B) is shown as a SNR. (D) the first frame SNR of the four strains.

Imaging the fluorescence signals right below the rotating beads on a stuck cell were also attempted, because detecting the fluorescence signal of the motor rotating a bead is a promising method to study the stator dynamics in relation to the torque generation at different loads. To image both the rotating bead and the eGFP fluorescence signal on the same cell, different illuminations and filters were required. Magnetic beads (NEB S1420S) that do not emit fluorescence signal by 488 nm laser were used to prevent any unwanted fluorescent noise background. To visualize the bead on the EMCCD, the emission filter in front of the camera was removed and very low 660 nm laser light (~ 2 mW) was used. Figure 4.3 shows an example cell with a rotating bead and the fluorescence image of the same cell. The fluorescence image was imaged with a 488nm laser by Epi-illumination that was focused on the upper side of the membrane. In this image, a bright fluorescent spot near the rotating bead was observed. However, there was a shift of about 8 pixels between the first movie of the rotating bead and the second acquisition of the motor fluorescence, (likely due to a tilt in the filter that had to be removed and replaced again depending on the illumination) which should be corrected to perfect the technique. Being able to detect the motor fluorescence intensity of the motor rotating a bead can bring many advantages over a tethered cell assay measurement. Therefore, fluorescence imaging of the motor spot of the cells rotating a bead should be optimized in the future.

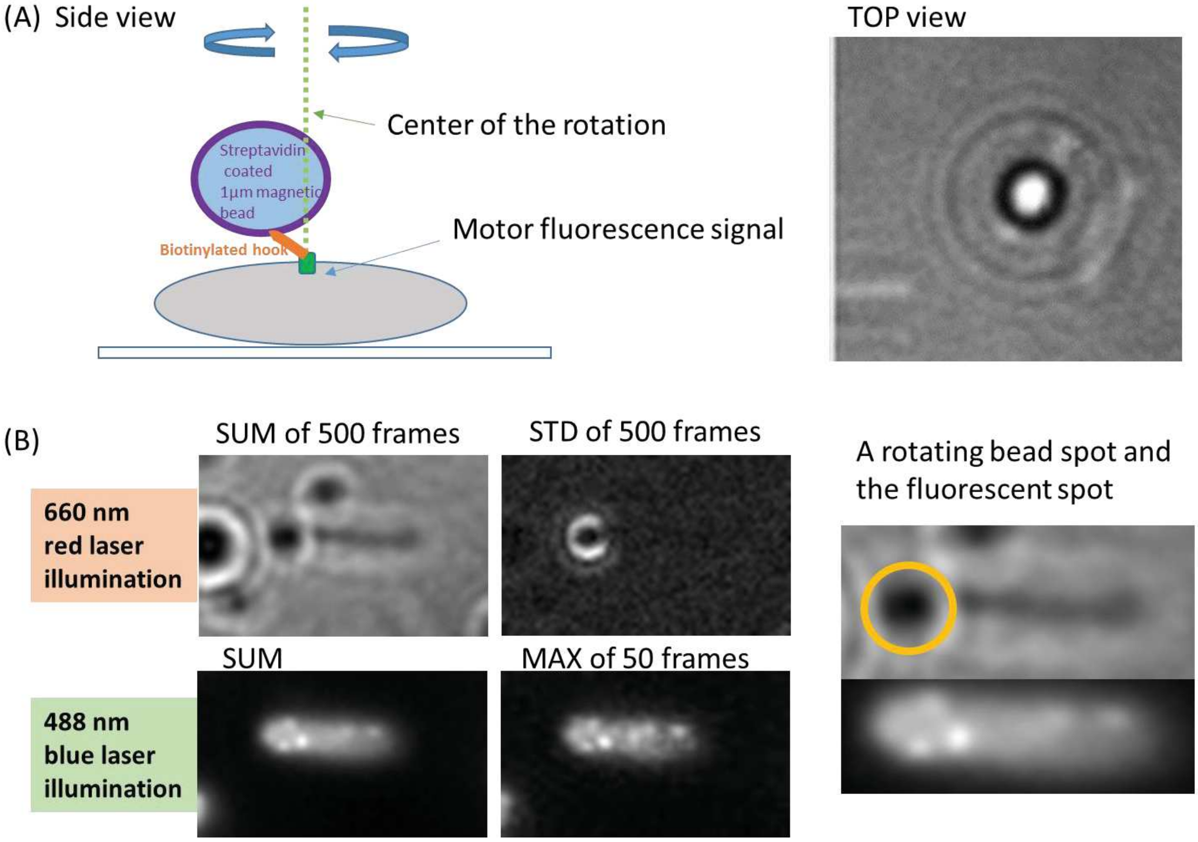
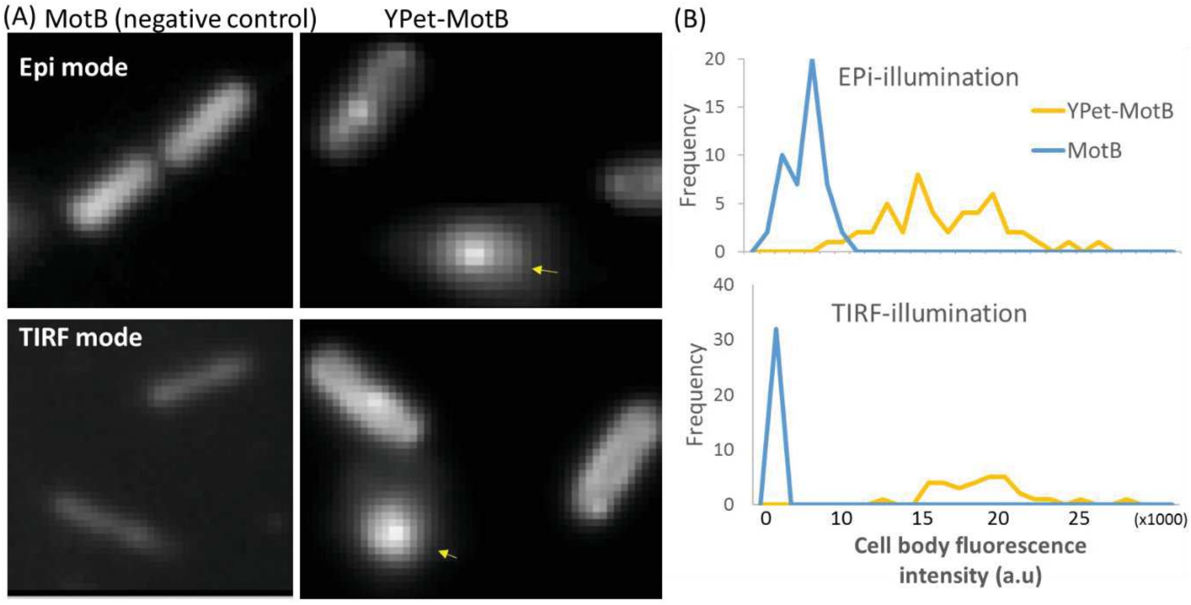


Figure 4.3 (A) A schematic figure of a rotating bead on a stuck cell from the side view, and an example image of a stuck cell with a bead attached from the top view. (B) An example cell with a rotating bead on the top. Unlike SUM image, the standard deviation (STD) image of a stack of 500 frames shows a bead only, which reveals that the bead was moving while everything else around was static. Sum and max images of the eGFP fluorescent signal image show a few of fluorescent spots on the cells near the bead.

4.2.3 YPet-MotB imaging in Epi illumination and in TIRF illumination

Fluorescence signals of the YPet-MotB fusion strains imaged by Epi- 488 nm laser excitation showed clusters of fluorescent spots (Fig 4.4a). However, unlike the eGFP-MotB strains, the fluorescent spots in these YPet cells were less apparent and their cell body fluorescence were higher than the WT negative control cells (Figure 4.4b). The distribution of the cell body fluorescent intensities of the YPet cells was wide compared to that of the negative control WT cells, suggesting high cell to cell variation and high number of diffusing YPet molecules in the cell. When the cells were imaged by TIRF illumination so that only the cell membrane in contact with the glass is illuminated, the cell body fluorescence intensity of the WT cells was reduced by half, while that of the YPet cells was not (figure 4.4b). This result may suggest that most of the cell body fluorescence intensities of the YPet cells are from the membrane rather than from the cytosolic part of the cell. The distributions of the signal to noise ratio (SNR) of the YPet cells (signal is the motor spot intensity) from Epi illumination and TIRF illumination are shown in figure 4.4c. TIRF mode slightly shifted the distribution to the right (a better SNR). The cells were grown in a high induction [0.1% arabinose] condition in this figure.



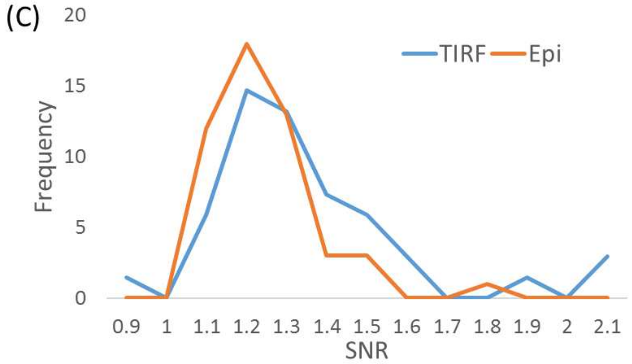


Figure 4.4 (A) Fluorescence imaging of the YPet-MotB cells and the MotB (negative control, WT) cells in Epi and TIRF illumination. The cells were *E. coli* strains JPA605 carrying either plasmid vector of pBAD33_MotA, MotB, and pBAD33_MotA, YPet-MotB, both expressing the stator genes by 0.1% arabinose induction condition. Rotating YPet cells showing the fluorescent spots in the rotational centers are indicated by yellow arrows. No clusters of the fluorescence spots formations were found in WT cells. (B) Histograms of the cell body fluorescence signals from Epi and TIRF modes. TIRF illumination reduced cell body autofluorescence for the WT cells but not for the YPet cells. The histograms are normalized and the number of cells measured were 48 and 32 for the YPet cells and 57 and 45 for the WT cells (Epi and TIRF, respectively). (C) SNR of the YPet motor spot fluorescence intensities in both Epi and TIRF mode are shown. The mean SNR for TIRF and EPi were 1.27 ± 0.17 and 1.20 ± 0.13, respectively.

4.3 Fluorescence signals of the cells grown in high and low level of gene induction conditions

Fluorescence signals of the eGFP-MotB cells grown in two different gene induction conditions (high and low) were compared. The cells were imaged by 488nm laser TIRF illumination for a period of about 20 s (200 ms x100 frames, frame transfer mode). The stack of images was summed into a single image and its background were subtracted using a subtract background function (rolling ball radius of 10) in Fiji_ImageJ. The “subtract background” commend on ImageJ removes smooth continuous backgrounds from the image based on the a “rolling ball” algorithm described in [13]. Each flied of view (FOV) had a similar number of cells (60~80 cells), and their background noise level were almost identical, thus the mean fluorescence intensity of the entire FOV of each sample were compared. Consistently, (about twice) higher total mean fluorescence signal was detected from [0.1% ara] FOVs than from [0.002% ara] FOV from all four different FOVs of each sample. Example images are shown in figure 4.5. This observation was confirmed in Epi-illumination as well, constantly showing nearly twice higher total fluorescence intensity in [0.1% ara] FOV than in [0.002% ara] FOV (data not shown).

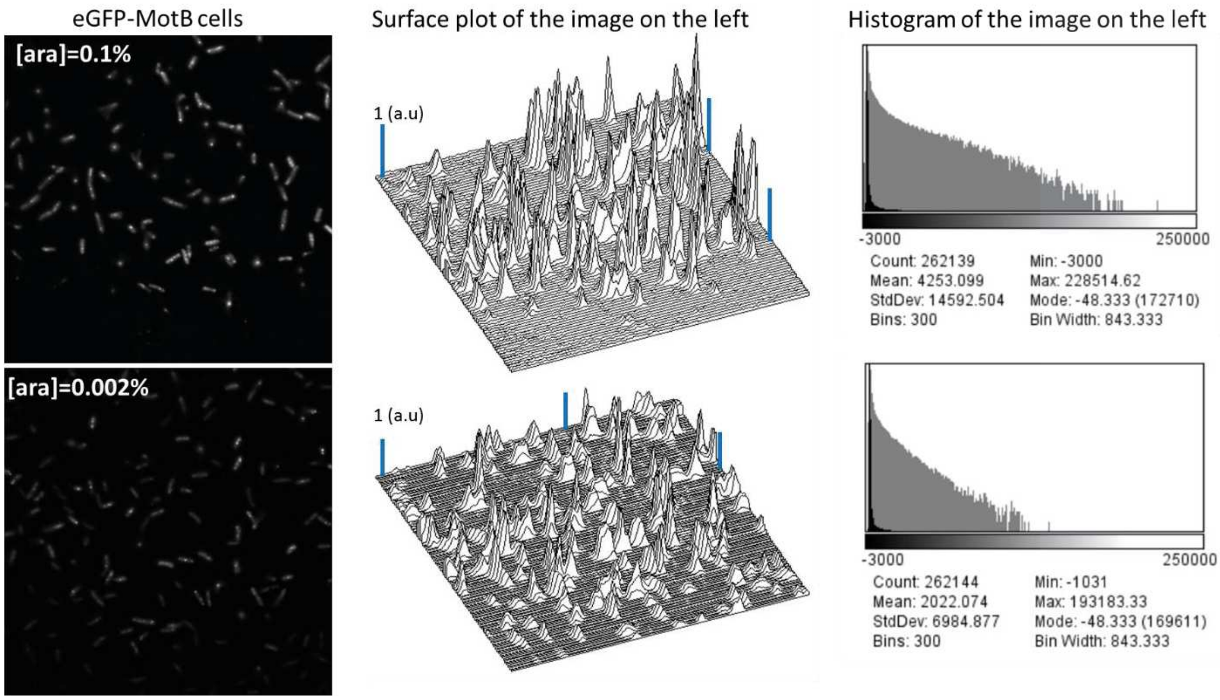


Figure 4.5 TIRF images of the eGFP-MotB cells grown in high [0.1% ara] and low [0.002% ara] induction conditions are shown. Their surface fluorescence intensity plots are shown with the same vertical scale bar, indicating the difference in emission from individual cells. Histograms of the entire

field of view pixel values are shown. The peak at low intensity represents the background, and the broad signals at higher intensity represents the fluorescence signals from the cells. The mean pixel intensities of [0.1% ara] sample is more than twice higher than that of [0.002% ara]. The x-axis range is the same in both histograms.

The difference in fluorescence signal intensity was also observed from the strains expressing the FP-MotB protein by native promoter and by the plasmid vector. The eGFP strain (JPA804, a gift of R.Berry’s lab) that expresses the 28AA-eGFP-MotB fusion protein by their native promoter (named eGFP_NP for simplicity) was imaged as a positive control since the fluorescence signal of this strain has already confirmed by a previous study [4]. The cells showed bright fluorescent spots mostly at the periphery of the cells by the 488 nm laser illumination (figure 4.6a). The eGFP strain (JPA605_pBAD33 MotA, eGFP-MotB) that expresses the eGFP-MotB fusion protein by pBAD33 expression vectors also showed eGFP fluorescence signal as concentrated clusters in the cell, mostly on the cell membrane as well, revealing the membrane localization of MotB proteins (figure 4.6a). This strain is named eGFP for simplicity.

The fluorescence signals by these two strains (GFP and eGFP_NP) were compared from the sum of the pixel intensities of 100 frames (with an exposure time of 0.2 s per frame in frame transfer mode). While both strains showed fluorescence clusters (fig 4.6a), eGFP_NP showed a better signal to noise ratio than eGFP: the mean cell body fluorescence of eGFP_NP was 16 % lower but the mean motor spot intensity was about 9% higher when compared to eGFP (figure 4.6b). This difference is observed by the distribution of eGFP motor signal to noise ratio (SNR: the spot fluorescence intensity divided by the corresponding stuck cell body fluorescence intensity) (figure 4.6c). The eGFP strain here was grown in a high induction condition (0.2% arabinose) overexpressing the stator genes, while eGFP_NP expresses the gene by their native promoter. The native expression level of the stator protein (MotB) is considerably lower than the gene induced by 0.2% arabinose by pBAD33 vector as shown by the immuno-blot experiment in [5]. The estimated number of the stator proteins in a native cell is ~190 ± 80 per cell [6]. This suggests that this cell body fluorescence intensity data may correspond to the difference in the expression levels of the two eGFP-MotB strains.

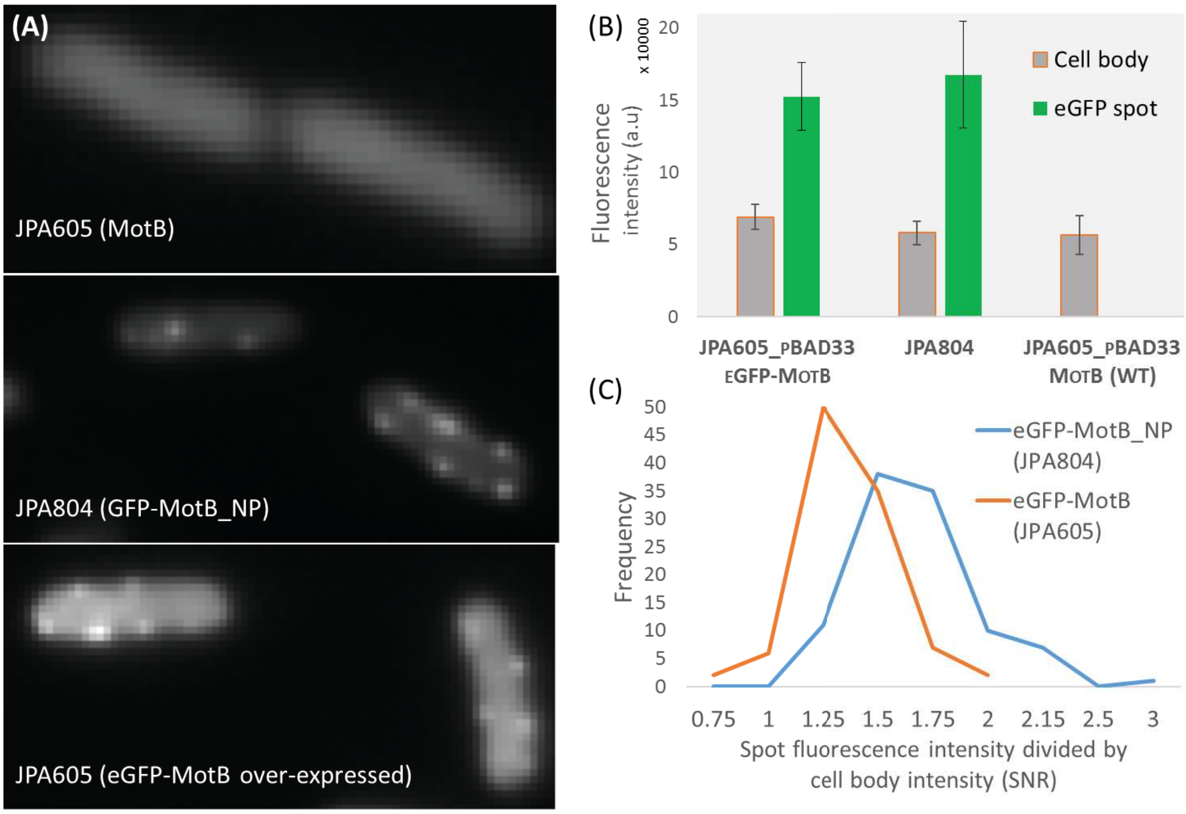


Figure 4.6 Comparison of the fluorescence intensity of cells expressing eGFP-MotB by native promoter and by a high induction condition vector expression system (pBAD33). (A) Representative images (sum of 100 stacks of images) of the cells from the negative control strain JPA605 (top), expressing MotB without eGFP tag by pBAD33 vectors, JPA804 (middle), expressing 28AA eGFP-MotB (eGFP_NP) and JPA605 (bottom), expressing eGFP-MotB by pBAD33 vectors. The grey scale of the images was automatically optimized brightness and contrast by ImageJ. No fluorescence spot formation was found in JPA605 negative control cells while the two eGFP expressing cells showed concentrated clusters of fluorescence signals mostly on the cell membrane. The cells were stuck on the poly L- lysine coated glass slide. The images were recorded in Epi-illumination by 488 nm laser, for 100 frames at 0.2 s exposure time with a -80 C cooled EMCCD camera (80 EM gain, 488 nm laser intensity: 20 mW 2/10). (B) Fluorescence intensities of the eGFP spots and that of the cell bodies of the three strains. (C) Normalized histogram showing the distributions of SNR (the spot fluorescence intensity divided by their cell body fluorescence intensity) of eGFP_NP and eGFP cells. The distribution was made by the 104 fluorescent spots in live 40 eGFP_NP cells and 82 spots in live 33 eGFP cells. eGFP_NP has a wider distribution than eGFP cells, and the mean SNR of eGFP_NP was 1.54±0.26 and the mean SNR of eGFP was 1.24±0.16.

4.4 Photo-converted Dendra2 imaging in live *E.coli*

Green-to-red photo-convertible fluorescent protein Dendra2 fused to MotB protein was imaged in order to test the possibility of resolving a super resolution image of the stators around the motor using a photo-activated localization microscopy (PALM). Dendra2 is characterized as monomeric, fast and efficient maturation, high contrast photo-conversion, bright fluorescence and high photostability fluorescent protein [7]. Dendra2 can be photoactivated (red form) by both UV-violet (e.g., 405 nm laser) or blue light (e.g., 488 nm laser), though activation with blue light is much less efficient than with UV-violet light. The activated Dendra2 possesses the maximum excitation at 553 nm and emission at 573 nm [7]. Dendra2 is a suitable fluorescent protein for the photoactivated localization microscopy for molecular counting, as demonstrated in [8] since Dendra2 photobleaches three times faster and blinks seven times less when compared to another photoconvertible fluorescent protein mEos2.

Dendra2 can be photo-activated by two methodologies [9]: The first method is to activate and excite them instantaneously by having both the photoactivation laser (405 nm) and the excitation laser (552 nm) on simultaneously. The laser power of the activating light should be considerably low (and slowly increased over time), while the laser power of the exciting laser should be high enough to subsequently photobleach the photoactivated Dendra2, so that fluorescent proteins are sequentially photoactivated, detected and photobleached one at a time and sparsely in the FOV (figure 4.7a). A second method is to have them photoactivated by a pulse of the activating laser followed by a pulse of the excitation laser (552nm) so that all the Dendra2s, photoactivated by each pulse, are photobleached before the next activation pulse (figure 4.9a). In each activation pulse (405 nm illumination time), one or few Dendra2 is photoactivated, and these photoactivated Dendra2 should be imaged until they photobleach (552 nm illumination time). This second method, despite intrinsic difficulties, is more reliable for the estimation of the stoichiometry of molecules in a closed proximity, as in principle it can discriminate blinking of a single fluorophore, and reduces the possibility to convert and count more than one molecule in a cluster.

Initially, the photoactivated red form of Dendra2s were visualized by the first method, having both lasers on simultaneously (images of the non-activated green form of Dendra2 are shown in figures 4.1). Dendra2 photoactivation, emission and photobleaching events were sequentially detected, revealing that the FP localized mostly at the cell membrane (figure 4.7). The reduced photoactivation efficiency was also observed as expected when the power of activation laser (405 nm) was reduced (figure 4.7e). WT cells was imaged in the same illumination condition as a negative control and showed no activation nor emission lights (figure 4.7e).

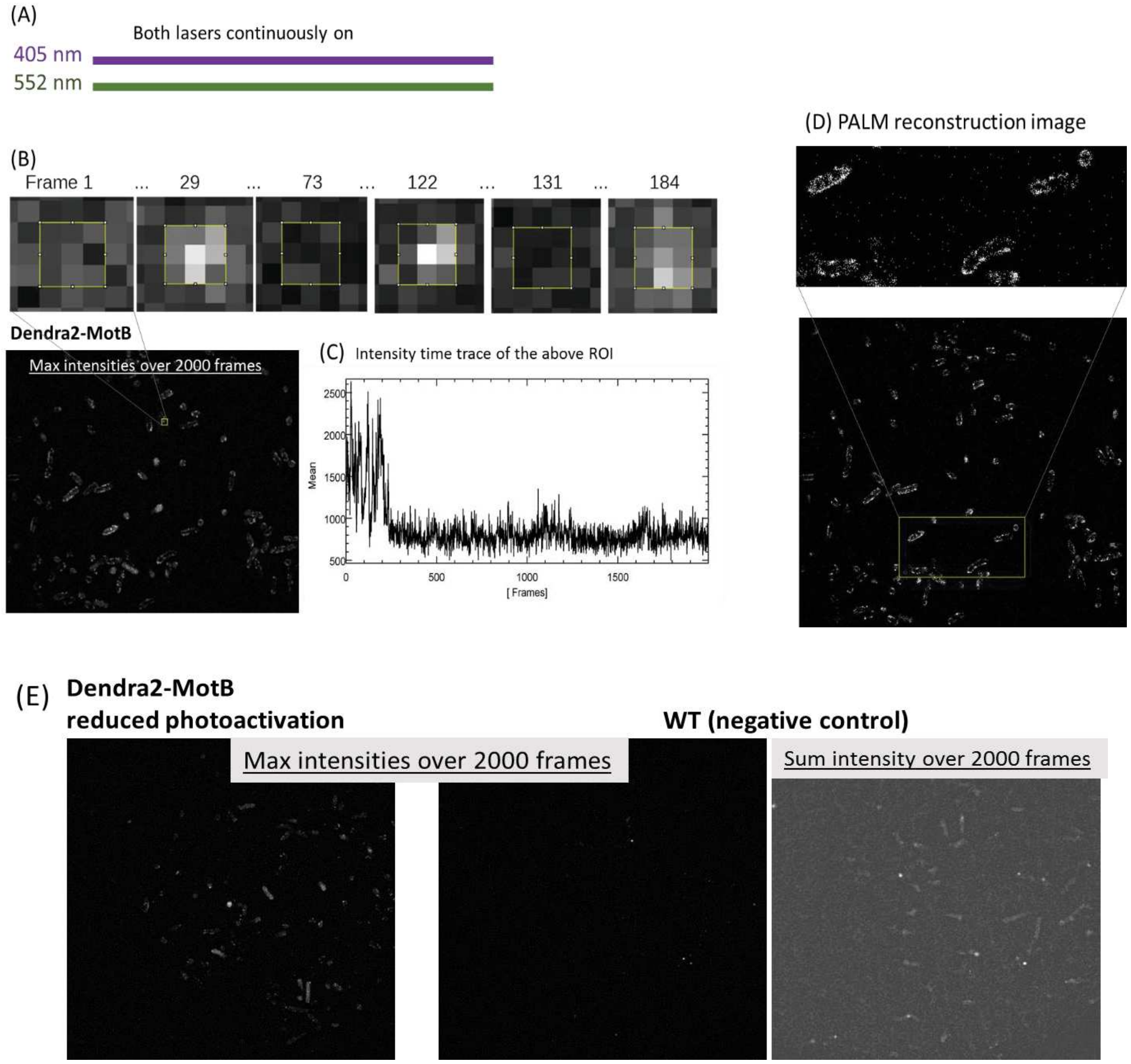


Figure 4.7 Simultaneous photoactivation and imaging of Dendra2-MotB in live *E.coli* cells. (A) Dendra2s are photoactivated, detected and photobleached by simultaneous illumination with the 405 nm activating laser and the 552 nm imaging laser. (B) Photoactivated Dendra2 fluorescence detected with the EMCCD camera shows fluorescence around the membrane and some spots on the membrane. EMCCD images of Dendra2 (yellow box ROI) in selected frames show fluorescence emission states and dark states (pixel size ~ 147 nm). (C) Time trace of the emission intensity of the fluorescence area shown above. (D) PALM reconstruction image (by QuickPALM analysis in imageJ) of the same FOV in (B). This image reveals that Dendra2-MotB are mostly localized at the cell membrane and some localized area at the membrane, representing possible location of the motors. (E) When 405nm activation laser light is reduced by more than half, less fluorescence signals were detected and WT negative control cells showed no fluorescence signal excited by both 405nm nor 552 nm laser illuminations. All images here were done in the same day (Scale bar, 1 µm.)

Figure 4.8 shows the emission time traces of some of the selected spots from the field of view in figure 4.7. The Dendra2 emission per peak just below 2000 (a.u) were most frequently observed, indicating potential single emission intensity of the activated Dendra2. The peaks above 2000 (a.u) are presumably multiple Dendra2 proteins (at the motor or nearby) emitting light at the same time. Multiple peaks with short off times followed by longer off times were often seen. The off time is non-activating/emitting period. The multiple peaks can be the multiple Dendra2 activated at the same time or Dendra2 blinking events [9].

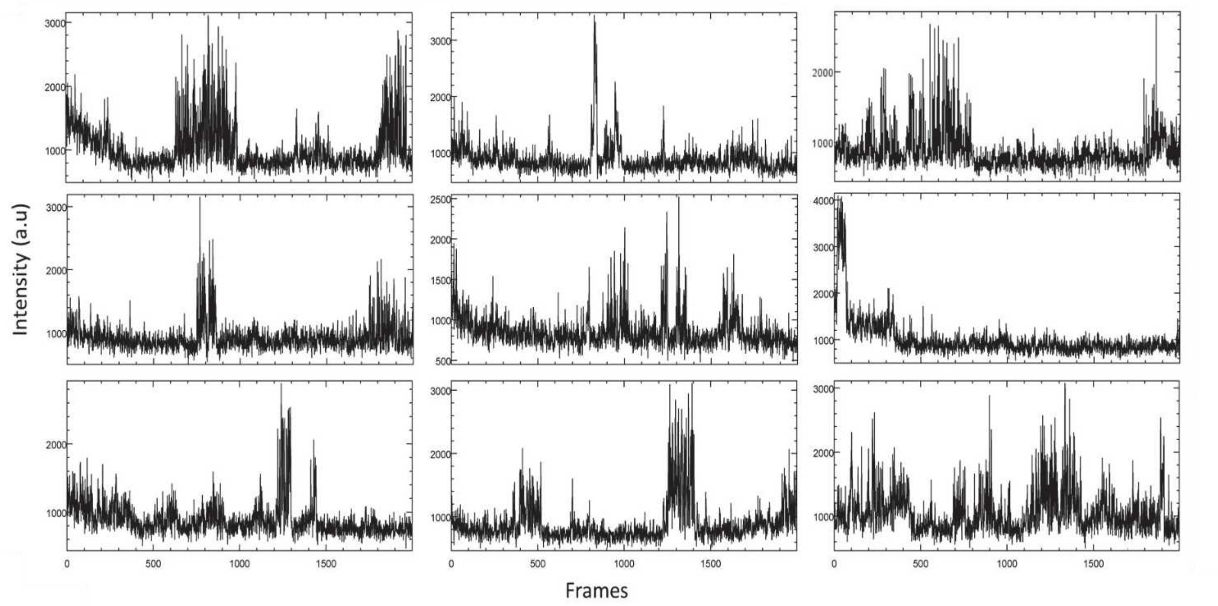
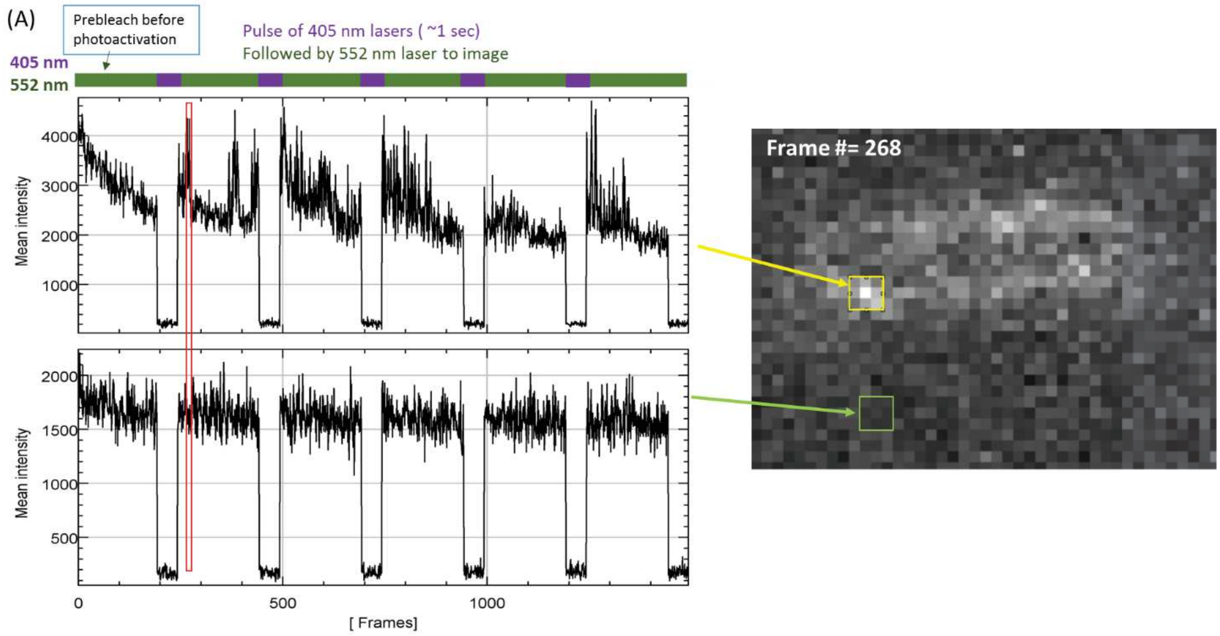


Figure 4.8. Emission intensity of 9 fluorescence clusters of Dendra2 (ROI of 3x3) from figure 4.7b. Each peak can represent emission events. The two lasers (405nm and 552nm) were simultaneously on and the frame rate was 20 ms.

Photoactivated Dendra2s were visualized by the second method as well. The cells were pre-bleached by illuminating the 552 nm lasers before activating Dendra2 to photobleach all the pre-activated (naturally red form) Dendra2 proteins (figure 4.9a). Then, upon a pulsed 405 nm laser illumination to photo-convert a fraction of the Dendra2 molecules, a pulse of the 552 nm illumination was followed to image the photo-converted Dendra2 molecules. The fluorescence emission was especially concentrated at the peripheral side of the cells, while less fluorescence was detected elsewhere. The sum and max images of the recorded frames (>1000 frames) revealed the (diffraction-limited) localization of the photoactivated Dendra2 molecules (figure 4.9c). Although these images are not the super-resolution images of Dendra2-MotB, they already provide different and improved/sharper fluorescence images without much cell body fluorescence background when compared to the fluorescence images by the conventional non-photoconvertible FPs (eGFP, YPet, and non-converted green form of Dendra2), confirming their membrane localizations.

Due to the current limit of the PALM spatial resolution, which is in the 20~40 nm range [1], however, the precise localization of more than 2 or 3 Dendra2-MotBs in the 45 nm diameter of motor is not possible. The number of stators at the motor should be limited to only two stators bound state, but we had no means to control and limit the motor to rotate at a fixed number of stators (only two stators in this case) at this stage. Therefore, resolving PALM super resolution image of only two stators (Dendra2-MotB) bound state in the motor could not be achieved in this study. However, it is possible to resolve super resolution images showing how stators are distributed in the entire cell membrane (not just focused on the localization of the bound stators in the motor), favorably from the Dendra2-MotB expressed by the native promoter. Super-resolved stator localization images may provide a detailed information on the stator expression level on the membrane as well as the patterns of the diffusing stators around the motor, which could potentially enlighten our current understanding of stator dynamics and stator turnovers.



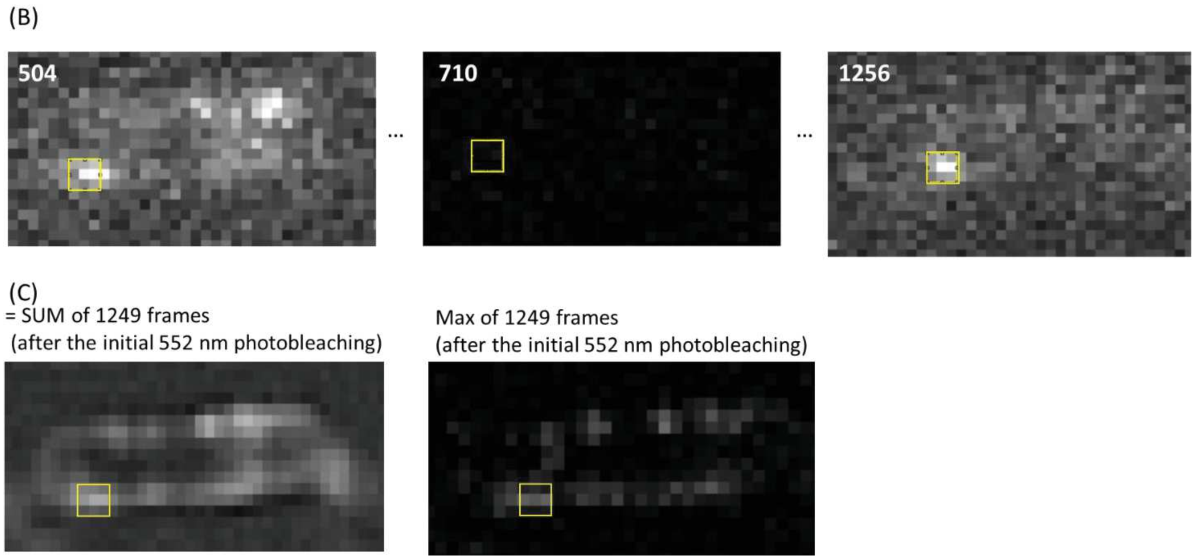


Figure 4.9 Dendra2-MotB imaging in live *E.coli* cell after sequential pulses of photo-activating 405 nm laser and exciting 552 nm laser. (A) naturally activated Dendra2s are initially photobleached by 502 nm laser (pre-bleach) prior to the photoactivation by 1 second of 405 nm laser pulse. The fluorescence intensity of the selected ROI at the cell membrane (top trace) shows intensity increase upon 405 nm laser illumination while another ROI outside of the cell area (bottom trace) shows no intensity change. A single frame is shown which is at a time point indicated as a red box in the intensity time trace. (B) three other images of the cell from (A) including the same Dendra2 ROI. The numbers on the images represent the frame number (frame rate=50ms, in frame transfer mode). (C) SUM and Max of the stack of 1249 frames of the same cell with the same selected ROI (a yellow box).

4.5 Dual recording set up and a proposed experiment

In order to record the function of the motor and the fluorescence signal of the stator bound to that same motor simultaneously, a dual recording set up was established (see figure 2.9 in chapter 2). The function of the motor can be monitored from the rotation of the cell body ("tethered cell" assay) or from the rotation of a bead ("tethered bead" assay) attached to a sheared flagellum or hook, recorded by a fast CMOS camera (typically at 500 – 1000 FPS, up to ~10000 FPS). At the same time, the fluorescence signal of the same motor can be recorded by an EMCCD camera at up to 20 ms per frame.

One experiment that can be done, in the future, in this dual recording set up is the measurement of stator binding and torque generation in a high temporal resolution (figure 4.10). The aim of such experiment would be to observe stator binding by both the fluorescence of the last incorporated stator and by a change in speed of the bead. In this way, it could be possible to observe if a time delay exits between the stator binding time and the speed increment of the motor (torque generation). This would indicate an activation time for the bound stator, which could shed light on the mechano-chemical cycle of the stator, still not well characterized [10-12]. Such a speed increment (delayed from the fluorescence step given by the incorporation of a stator) could be observed only if the delay is longer than 20 ms, corresponding to the highest time resolution of the EMCCD. Tracking a single eGFP molecule over an extended period is impossible due to the photobleaching. Thus, a TIRF illumination of tethered cells can be tried to photobleach only the eGFP molecules in the bottom membrane side of the cell where they are exposed to the laser. When the non-photobleached eGFP-MotB from other side of membrane (not exposed to the laser) incorporates to the motor that was rotating the cell body, a fluorescence signal could be observed from the center of the rotation while a speed increment could be detected by a CMOS camera. Ideally, small tethered cells with lowest drag values should be tried to better observe speed increment or decrements.

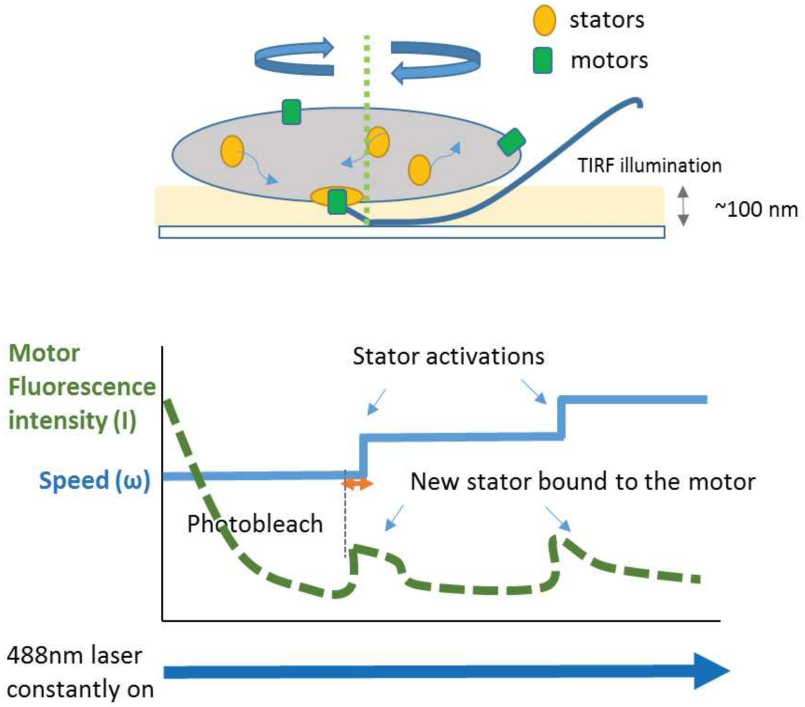


Figure 4.10 A schematic drawing of the proposed dual recording experiment in the TIRF illumination. The continuous 488 nm illumination photobleaches all the eGFP molecules where exposed to the TIRF illumination, and when a non-photobleached stator bound to the motor, an eGFP fluorescence signal followed by an increment of the speed can be detected. The potential time delay between stator incorporation and the torque generation is marked by an orange arrow.

4.6 Discussion

The fluorescence imaging of the live *E. coli* strains expressing the FP tagged stator proteins revealed that those fusion stators were localized at the cell membrane, as expected for the membrane proteins. 1~8 clusters of fluorescence spots per cell were also detected. These fluorescent clusters are most likely motor spots as confirmed by the fluorescence imaging of the rotating cells tethered on the glass surface, which showed a bright fluorescent spot at the center of the rotation. The tethered cells imaging can be optimized by TIRF illumination that focus on the membrane side. Thus, the cells grown at low induction (0.002 % ara) excited by TIRF illumination showed the most visible clusters of the fluorescent spots. In contrast to the cells expressing the FPs (eGFP, YPet and Dendra2)-MotB fusion proteins, the negative control cells (expressing MotB by the same vector) showed lower cell body fluorescence signals and no visible clusters of fluorescent spots at the cell membrane when illuminated by 488 nm laser.

Among the three FPs, YPet cells showed the highest cell body fluorescence compared to the other strains and tended to photobleach faster than eGFP. Dendra2 (unconverted green form) was less bright than eGFP and YPet. These relative fluorescence intensities of the three FPs are somewhat comparable to the known brightness of the three FPs (see table 4.1). However, a precise comparison of the three FPs can be made only after measuring the fluorescence intensity of each FP at the single molecule level. The fluorescence intensities of eGFP-MotB in two different induction conditions (low and high by the same plasmid vector system) were compared. This comparison revealed that the cells grown at the higher induction have a higher total fluorescence intensity, which likely represents the higher number of diffusing stators on the cell membrane in average. How the increased number of diffusing stators at the membrane effect the stator turnover or the switching frequency of the motor has yet to be measured. However, in this study, a reduced switching frequency (by half) was observed in the cells over expressing the stators (figure 3.16).

Photoactivated form of Dendra2 was imaged by 405 nm photo-activating laser and 552 nm exciting/imaging laser. The photo activated Dendra2 images clearly showed that Dendra2-MotB are localized at the cell membrane. This membrane localization was better visible by photo activated Dendra2 images than unconverted green form Dendra2 images. In summary, the fluorescence signals from all three FPs were detected, revealing that the three FP-MotB fusion proteins were well folded, transported, localized to the cell membrane and bound to the motors at the membrane. The novel dual recording set up can monitor the speed of the motor rotation and the stator fluorescence signals simultaneously. This dual recording may help to uncover the dynamic properties of stators and their activation mechanism at the highest spatial and temporal resolution.

References

1. Gahlmann, A., & Moerner, W. E. (2014). Exploring bacterial cell biology with single-molecule tracking and super-resolution imaging. *Nature Reviews. Microbiology*, *12*(1), 9–22.
2. Sahl, S. J., & Moerner, W. E. (2013). Super-resolution fluorescence imaging with single molecules. *Current Opinion in Structural Biology*, *23*(5), 778–87.
3. Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., … Cardona, A. (2012). Fiji: an open-source platform for biological-image analysis. *Nature Methods*, *9*(7), 676–682.
4. Tipping, M. J., Delalez, N. J., Lim, R., Berry, R. M., & Armitage, P. (2013). Load-Dependent Assembly of the Bacterial Flagellar Motor. Mbio.
5. Morimoto, Y. V, Nakamura, S., Kami-ike, N., Namba, K., & Minamino, T. (2010). Charged residues in the cytoplasmic loop of MotA are required for stator assembly into the bacterial flagellar motor. Molecular Microbiology, 78(5), 1117–29.
6. Leake, M. C., Chandler, J. H., Wadhams, G. H., Bai, F., Berry, R. M., & Armitage, J. P. (2006). Stoichiometry and turnover in single, functioning membrane protein complexes. Nature, 443(7109).
7. Chudakov, D. M., Lukyanov, S., & Lukyanov, K. a. (2007). Using photoactivatable fluorescent protein Dendra2 to track protein movement. *BioTechniques*, *42*(5), 553–565.
8. Lee, S.-H., Shin, J. Y., Lee, A., & Bustamante, C. (2012). Counting single photoactivatable fluorescent molecules by photoactivated localization microscopy (PALM). Proceedings of the National Academy of Sciences of the United States of America, 109(43), 17436–41.
9. Durisic, N., Laparra-Cuervo, L., Sandoval-Álvarez, A., Borbely, J. S., & Lakadamyali, M. (2014). Single-molecule evaluation of fluorescent protein photoactivation efficiency using an in vivo nanotemplate. Nature Methods, 11(2), 156–62.
10. Chowdhury, D. (2013). Stochastic mechano-chemical kinetics of molecular motors: A multidisciplinary enterprise from a physicist’s perspective. *Physics Reports*, *529*(1), 1–197.
11. Lo, C., Sowa, Y., Pilizota, T., & Berry, R. M. (2013). Mechanism and kinetics of a sodium-driven bacterial flagellar motor.
12. Mandadapu, K. K., Nirody, J. A., Berry, R. M., & Oster, G. (2015). Mechanics of torque generation in the bacterial flagellar motor. *Proceedings of the National Academy of Sciences*, *112*(32), E4381–E4389.
13. S.R. Sternberg, "Biomedical Image Processing", Computer, vol.16, no. 1, pp. 22-34, January 1983

Chapter 5. Evolutionary integration of foreign stators in the BFM: functional effects of compatibilizing mutations

5.1 Intro to the cBFM evolution experiment and aims

The macromolecular complexes of the cell, including protein complexes such as the BFM, are composed of multiple protein components that have a certain degree of structural and functional independence. This organization of the biological system is referred to as modularity, as discussed earlier in chapter 1.3. Comparative studies have provided overwhelming evidence indicating that the evolutionary origin, innovation and adaptation of protein complexes are involved in functional incorporation of pre-existing proteins. In order to gain an experimental insight into the potential of evolution to forge new functional interactions between proteins in a complex, we examined if evolution can cause functional integration of pre-existing proteins with a degree of fortuitous pre-adaptation. Specifically, we examined the incorporation of incompatible orthologous BFM components into the BFM and other cell systems of *E.coli*. Exploring the evolutionary dynamics of a macromolecular machines in a well-controlled system may lead to a progress in a better understanding of molecular details of the evolutionary transitions as well as the origins of biological complexity. In addition, a molecular machine modified by component exchange can provide a unique model system to explore their functional dynamics.

The laboratory of Bertus Beaumont designed and conducted an experimental evolution study on BFM (before I join the project) upon which the work described in this chapter builds further. The evolution experiment result is introduced in this section, and the follow up studies that I performed are discussed from the chapter 5.2. The experimental model system involves a series of *E. coli* strain in which the stator genes motA and motB had been exchanged with their orthologues (foreign stators) from different strains and species of bacteria. These strains are referred to chimeric BFM (cBFM) strains in this study. The evolution experiment was performed in the porous semi-solid agar (0.3% w/v), in which populations of motile cells can form a moving chemotaxic front, allowing rapid colonization of the petri dish from the point of inoculation. Thus, the chemotaxis abilities are measured by examining the migration distance of the cBFM strains from their center of origin in the semi-solid agar plates.

In principle, efficient motility promotes migration of the cells forming circles of chemotactic rings, whereas non-motile cells remain at the original spot as a single colony. Figure

5.1 (top two images) shows populations of non-motile and motile cells in semi-solid agar plates. Interestingly, another type of the chemotactic motility - colonies around the original spot of inoculation rather than forming chemotactic rings (bottom right in figure 5.1) - was also observed. We call this motility as atypical chemotaxis (ATC), and the underlying mechanism of such motility is unknown. When a cell evolves an improved ability to chemotaxis in semi-solid agar compared to their direct ancestor, they form a flare of motile population of cells. An example image of a flare is shown in figure 5.1 (bottom left image). Flares are wedge-shaped populations of cells capable of escaping from the chemotactic front of their ancestral populations in semi-solid agar. The chemotactic front speed focuses on the speed of the chemotactic front of the cells. Therefore, these flares are noticeable when the chemotactic front speeds of the two populations are distinctly different. The experiment looked for these flares, which can be referred to as adaptive evolutionary changes in chemotactic front speed of a particular genotype moved. These flares of motile bacteria cells have been observed by several studies [1-5], but the flares of cBFM are distinguishable than the flares from these studies for theirs (stator orthologs) chimeric nature of the motor.

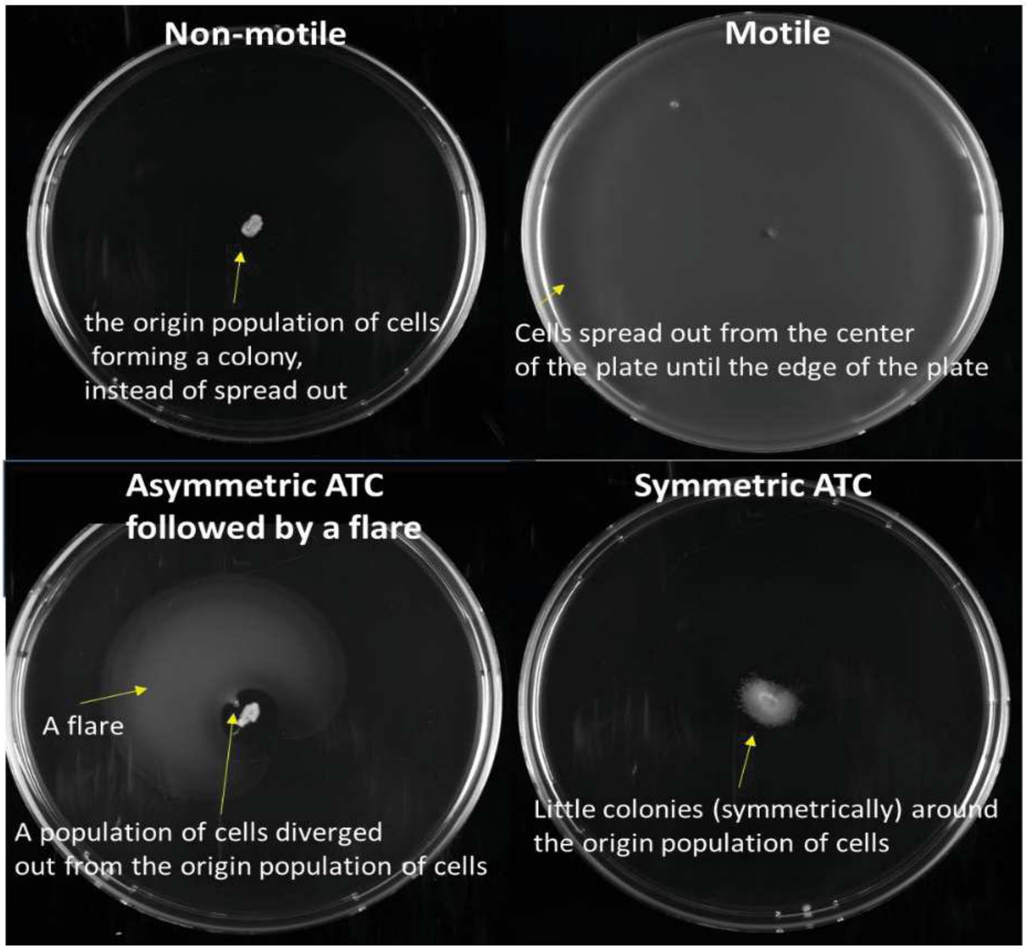
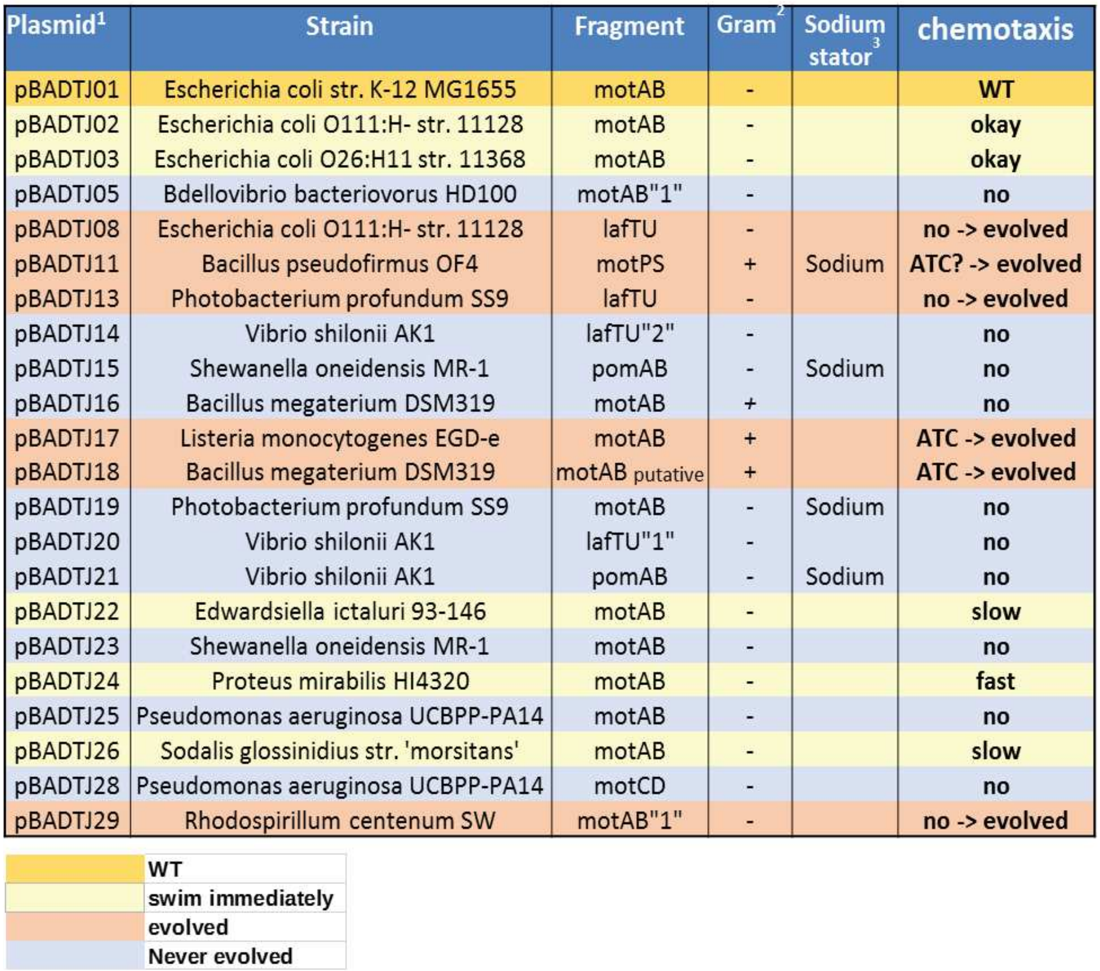


Figure 5.1 Examples of chemotaxis assay plates that show non-motile, motile and atypical chemotaxis (ATC) strains. Two types of ATC - asymmetric and symmetric - were observed. A flare diverging out from the ATC cells were often observed. The images were taken from the following strains 13-p, 8.1b-p, 11.1h-p and 29.1f-p (in a clockwise direction from non-motile cell).

The list of the cBFM and the outcome of the evolution experiment are shown in table 5.1. In brief, some of the strains were not capable of performing chemotaxis (population level of swimming); however, selection for motility caused some of them to be evolved to perform chemotaxis. Consequently, the cBFM foreign stators were categorized into the three compatibility groups: 'motile immediately', 'not motile but capable of evolving motility' and 'not motile and not able to evolve motility. Out of the total twenty-one cBFM, five cBFMs showed an immediately chemotaxis motility (immediately compatible), six showed evolved motility (potentially compatible) and the remaining ten cBFM showed no motility and never evolved to be motile (incompatible group) (Table 5.1). The potentially compatible cBFM strains were selected by allowing further growth of the non-immediately compatible cBFM strains on the semi-solid plates. When there is an evolution of increased fitness (chemotaxis ability in this case), a flare of a population of the cells can be observed. The evolution of increased fitness was checked in every 24 h. Most flares were observed within 6 days and the selection was continued for 35 days. For the detailed flares selection procedures, please refer to the PhD thesis of Regis Flohr. The strains that evolved to become motile represent functional integration of the pre-adapted yet incompatible stator orthologues, while the strains that never evolved to become motile represent incompatibility to achieve the functional integration. It is important to note that, however, some of the instances of these incompatible group strains might be because the orthologues stators were not expressed [6].

Table 5.1 A set of 22 MotAB homologs and their chemotaxis assay result



Plasmid[1]	Strain	Fragment	Gram[2]	Sodium stator[3]	chemotaxis
pBADTJ01	Escherichia coli str. K-12 MG1655	motAB	-		WT
pBADTJ02	Escherichia coli O111:H- str. 11128	motAB	-		okay
pBADTJ03	Escherichia coli O26:H11 str. 11368	motAB	-		okay
pBADTJ05	Bdellovibrio bacteriovorus HD100	motAB"1"	-		no
pBADTJ08	Escherichia coli O111:H- str. 11128	lafTU	-		no -> evolved
pBADTJ11	Bacillus pseudofirmus OF4	motPS	+	Sodium	ATC? -> evolved
pBADTJ13	Photobacterium profundum SS9	lafTU	-		no -> evolved
pBADTJ14	Vibrio shilonii AK1	lafTU"2"	-		no
pBADTJ15	Shewanella oneidensis MR-1	pomAB	-	Sodium	no
pBADTJ16	Bacillus megaterium DSM319	motAB	+		no
pBADTJ17	Listeria monocytogenes EGD-e	motAB	+		ATC -> evolved
pBADTJ18	Bacillus megaterium DSM319	motAB putative	+		ATC -> evolved
pBADTJ19	Photobacterium profundum SS9	motAB	-	Sodium	no
pBADTJ20	Vibrio shilonii AK1	lafTU"1"	-		no
pBADTJ21	Vibrio shilonii AK1	pomAB	-	Sodium	no
pBADTJ22	Edwardsiella ictaluri 93-146	motAB	-		slow
pBADTJ23	Shewanella oneidensis MR-1	motAB	-		no
pBADTJ24	Proteus mirabilis HI4320	motAB	-		fast
pBADTJ25	Pseudomonas aeruginosa UCBPP-PA14	motAB	-		no
pBADTJ26	Sodalis glossinidius str. 'morsitans'	motAB	-		slow
pBADTJ28	Pseudomonas aeruginosa UCBPP-PA14	motCD	-		no
pBADTJ29	Rhodospirillum centenum SW	motAB"1"	-		no -> evolved

🟨	WT
⬜	swim immediately
🟧	evolved
🟦	Never evolved

1.The numbers on the plasmids correspond to the group number of cBFM). 2. Gram-positive or gram negative bacteria. 3. All the other stators are proton translocating stators.

These phenotypic compatibility ranges were reflected on the genotypic similarity ranges of the cBFM stator orthologs (figure 5.2). The phylogenetic trees of the 22 motA and motB orthologues revealed that the immediately motile strains (yellow circled five cBFMs in figure 5.2) were indeed in close genetic distance to the wildtype motA and motB (*E.coli* K-12). The evolutionary modified compatible group (the six evolved cBFMs) is diverged into two groups in this phylogenetic tree. The first group with the cBFM 11, 17 and 18, which showed a ATC motility, and the second group with the cBFM 08, 13 and 29, which showed non-motile motility. Accordingly, these evolved cBFM strains can form two compatibility groups: the first group showed ATC motility from the beginning, then evolved to show normal motility in semi-solid agar, and the second group showed no motility from the beginning, then evolved to be motile in semi-solid agar.

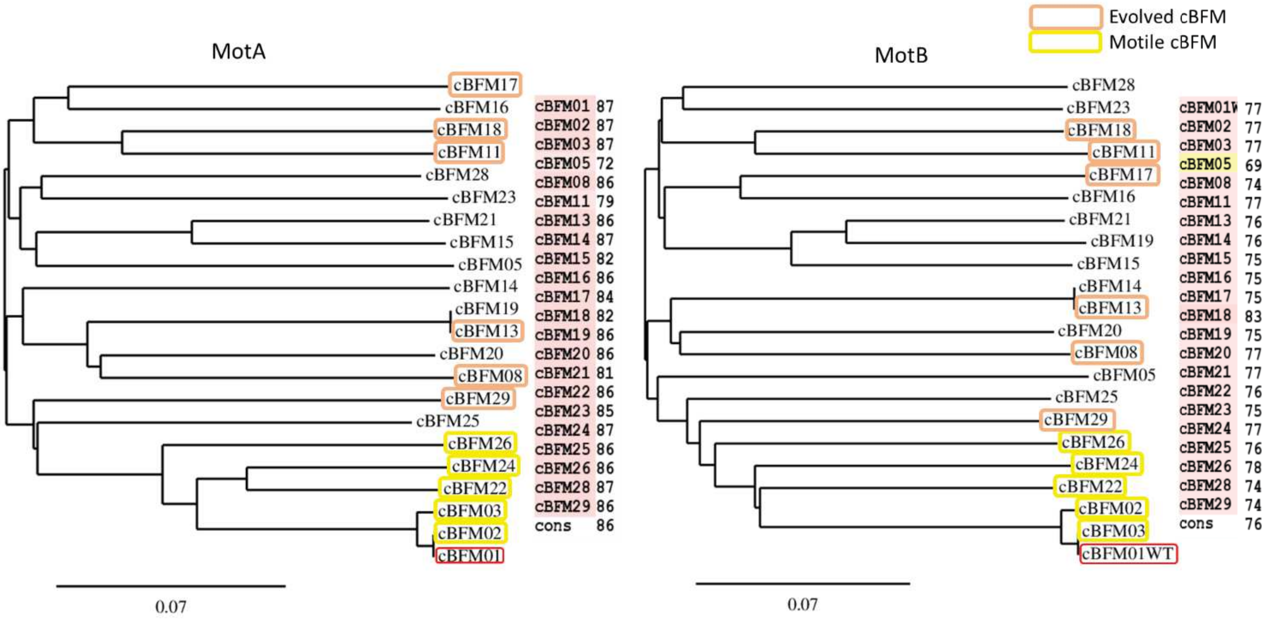


Figure 5.2 The phylogenetic trees of the 22 MotAB homologues. The trees were generated by the MotAB homologues amino acid sequences alignment by PSI/TM-Coffee (Align Proteins using Homology Extension against Reduced Databases) on website http://tcoffee.crg.cat/apps/tcoffee/do:tmcoffee and the online software TreeDyn 198.3. A scale bar is showing genetic distance. Sequence similarities (by PSI/TM-Coffee, transmembrane protein alignment) of the proteins are shown on the right side of each phylogenetic tree.

The cBFM that evolved to be motile showed a series of consecutive evolutionary integrations, which were observed as multiple generations of flares (up to sixth generations). The evolution trajectories from the six evolved strains, shown in figure 5.3, display certain difference within the cBFM group as well as between the cBFM groups. Certain lineages evolve further up to fifth or sixth flare/generations, while certain lineages stop evolving after the primary flare. In order to understand the pattern of genotypic changes occurred in those flares of strains and to see how the mutational trajectories fit in the fitness landscape, the entire pool of the evolved cBFM strains were sequenced and began to determine the order of mutations in some lineages. As an initial attempt to achieve the overall goal of understanding of the molecular details of the evolutionary transitions, here, we focused on the identification and analysis of the mutations of the early stage of evolutionary paths (primary flares), with an aim to resolve the functional effects of the mutations at the three different motility levels (population swimming, single cell swimming, single motor rotation). In particular, functional analysis of the cBFM at the single motor level was a primary interest in this chapter, because the rotation of the BFM is the fundamental element of bacterial cell swimming motility. The aims of this chapter are listed below.

i. identify cases in which the first beneficial mutation had occurred in the stators.

ii. look at the mutated amino acids on the sequence alignment of the 22 motAB orthologues and the literature to see what is known about changes at these sites.

iii. observe the motility of the non-evolved and evolved (cBFM-P) strains at the single cell level.

iv. characterize the motility of the non-evolved and evolved (cBFM-P) strains at the single BFM level.

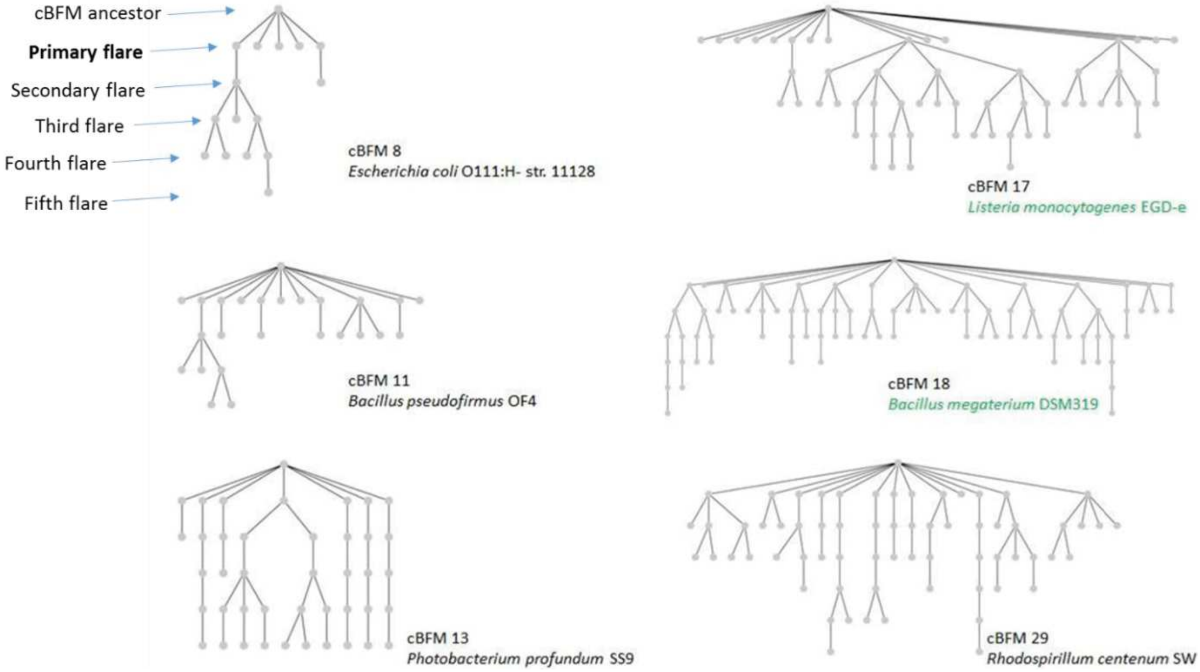


Figure 5.3 Evolutionary flare trajectories. cBFM8 had five primary flares, of which the first and the last flares evolved further (secondary, third, etc…). The other cBFM strains also exhibit multiple evolution trajectories in a similar manner. The last flares of each lineage are called end points. We focused on the primary flare strains in this chapter (figure made by Regis Flohr, a PhD student at Beaumont lab).

5.2 Primary flare strain (cBFM-P) preparations and motile strains screening

In order to understand the nature of evolutionary modifications that occurred in the early stages of the evolution, we first asked if any mutations had occurred in motA and motB. If so, are those mutations sufficient to improve the motility? To answer such questions, chemotaxis assay screening was performed with cBFM-P strains. The **cBFM-P** strains are the original motAB knockout *E.coli* K-12 parental cells that carry the plasmid vectors, encoding the foreign stator genes, isolated from the evolved primary flare strains. Figure 5.4 illustrates how the cBFM-P strains were prepared. The motility of the cBFM-P strains was tested in semi-solid agar (chemotaxis assay) in the same way as it was performed for the cBFM evolution experiment. The cBFM-P strains were prepared and their chemotaxis assays were performed together with Regis Flohr, a PhD student at Beaumont's lab.

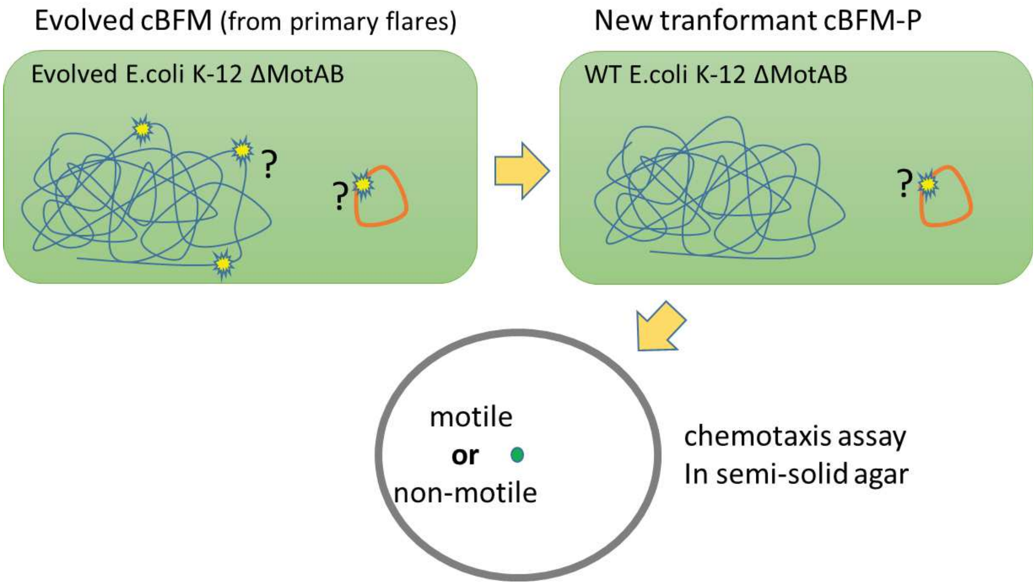
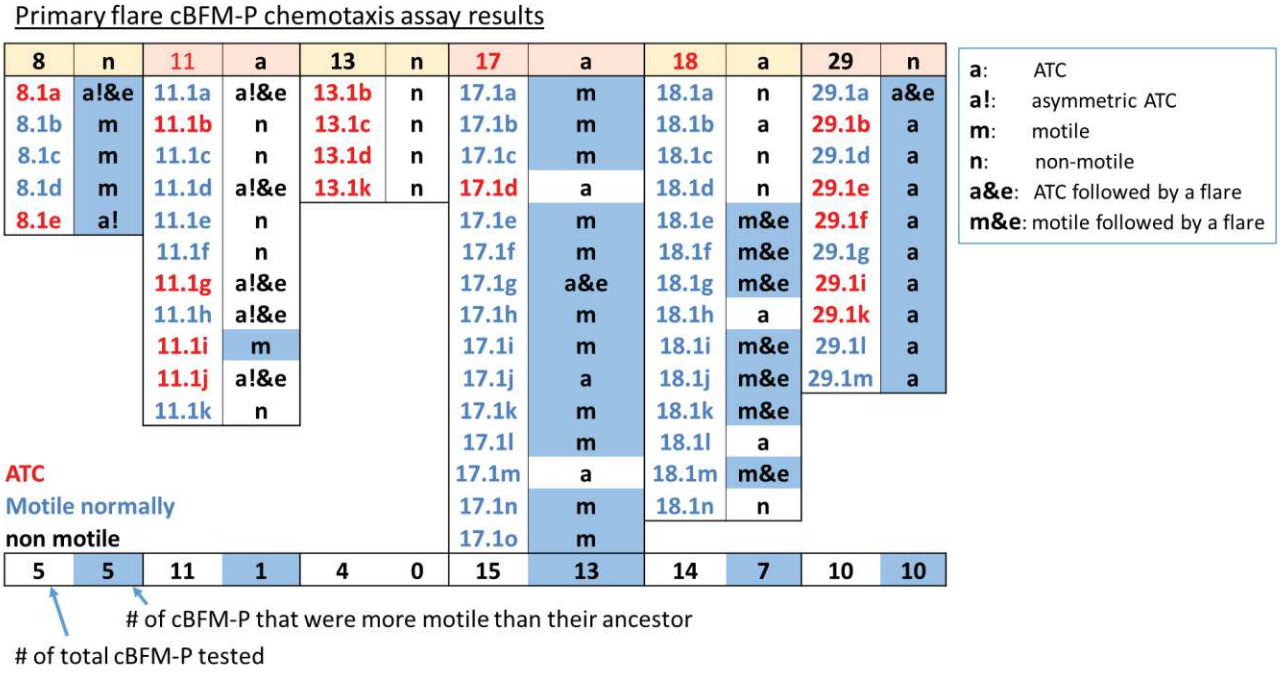


Figure 5.4 Primary flare strains cBFM-P preparation: the plasmid vectors encoding motAB homologue genes from the evolved cBFM strains were isolated and transformed back to the original motAB knockout *E.coli* K-12 strains, where there is no additional mutation. The potential mutations are shown as yellow stars on the genome (tangled blue line) and on the plasmid vector (red circular line). The motility of these cBFM-P strains was tested in semi-solid agar and compared that to their ancestor's motility.

Total 59 cBFM-P strains from six evolved foreign stators (cBFM 8,11,13,17 and 18) were prepared. The list of the cBFM-P strains tested and their chemotaxis assay screening results are shown in figure 5.5. The chemotaxis assays revealed that 36 strains out of 59 strains (~61%) improved chemotaxis than their ancestors. All the strains tested from the cBFM8-P and 29-P were more motile than their ancestors, while 13 out of the 15 strains from the cBFM 17-P and half of the cBFM18-P strains showed an improved motility compared to their ancestors. In contrast, cBFM11-P and 13-P had one or none of such improved motile strains. This chemotaxis screening result implies that certain changes had occurred in the plasmid vectors, which led to a better fitness in the given constraint (chemotaxis adaptation). In order to determine whether there were the genetic modifications in the stator genes, which were sufficient to improve the motility, we sequenced the foreign stators genes of the cBFM-P.



Primary flare cBFM-P chemotaxis assay results

8	n	11	a	13	n	17	a	18	a	29	n
8.1a	a!&e	11.1a	a!&e	13.1b	n	17.1a	m	18.1a	n	29.1a	a&e
8.1b	m	11.1b	n	13.1c	n	17.1b	m	18.1b	a	29.1b	a
8.1c	m	11.1c	n	13.1d	n	17.1c	m	18.1c	n	29.1d	a
8.1d	m	11.1d	a!&e	13.1k	n	17.1d	a	18.1d	n	29.1e	a
8.1e	a!	11.1e	n			17.1e	m	18.1e	m&e	29.1f	a
		11.1f	n			17.1f	m	18.1f	m&e	29.1g	a
		11.1g	a!&e			17.1g	a&e	18.1g	m&e	29.1i	a
		11.1h	a!&e			17.1h	m	18.1h	a	29.1k	a
		11.1i	m			17.1i	m	18.1i	m&e	29.1l	a
		11.1j	a!&e			17.1j	a	18.1j	m&e	29.1m	a
		11.1k	n			17.1k	m	18.1k	m&e		
						17.1l	m	18.1l	a		
						17.1m	a	18.1m	m&e		
						17.1n	m	18.1n	n		
						17.1o	m				
5	5	11	1	4	0	15	13	14	7	10	10

ATC
Motile normally
non motile

# of cBFM-P that were more motile than their ancestor
# of total cBFM-P tested

a: ATC
a!: asymmetric ATC
m: motile
n: non-motile
a&e: ATC followed by a flare
m&e: motile followed by a flare

Figure 5.5 The chemotaxis motility screening results of the cBFM-P strains. The ancestors (cBFM) chemotaxis motilities are indicated in the first column by three colors: red colored strains indicate ATC motility, blue indicates normal motility, black indicates no motility in semi-soft agar. Six different types of motilities were observed by cBFM-P strains. a: ATC; a!: asymmetric ATC; m: motile; n: non-motile; a&e: ATC followed by a flare; m&e: motile followed by a flare. The motility test result varies greatly from one group to another.

3.3 Beneficial mutations of the MotA/B homologues

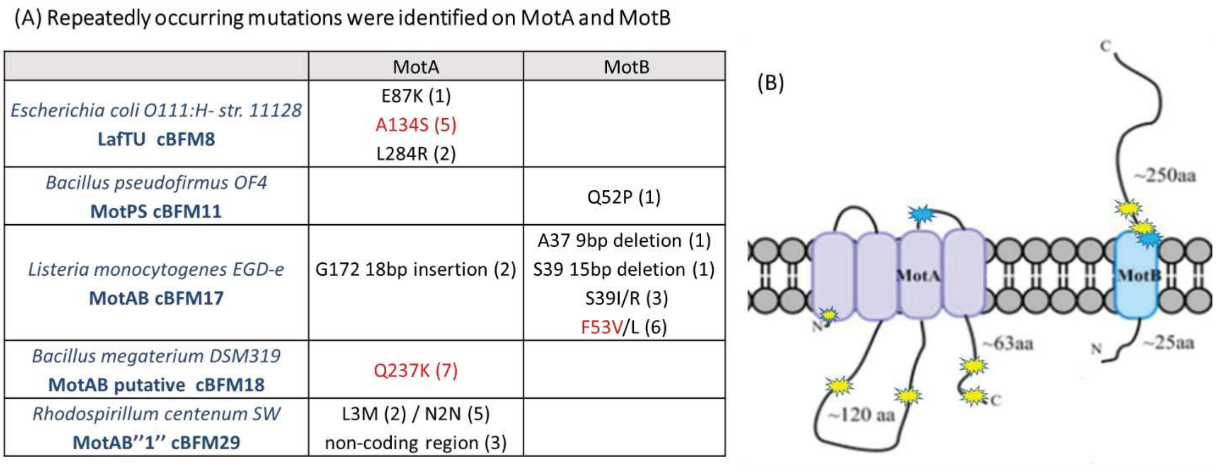
To identify the genetic modification occurred in those foreign stator genes, those that showed an improved motility than their direct ancestors (the 36 plasmid vectors) were sent out for Sanger sequencing. The sequencing reads were primarily focused on the sequences encoding motA and motB genes and the promoter sequences (the upstream of motA). In addition to this sanger sequencing, whole genome sequencing of the entire cBFM populations was performed by next generation sequencing (NGS). The sequenced cBFM populations (by NGS) are pooled samples of the end-point genotypes of the replicate evolutionary lineages of each cBFM strain. The NGS data confirmed that there are no additional mutations outside the coding region on the plasmids, except for cBFM29. Further analysis of this NGS data will not be discussed in this study.

The cBFM-P motAB mutations are listed in Table 5.2, together with their chemotaxis speeds measured by Regis Flohr. The number of mutations identified per each cBFM group varied: three different genotypes from the cBFM8, one genotype from cBFM11, seven different genotypes from cBFM17, one genotype from cBFM18 and three different genotypes from cBFM29. These mutations are beneficial mutations since they improved population motilities in semi-solid agar. The beneficial mutations are mutations that cause positive and beneficial effects, thereby they are rare [7]. The seven independent primary flares of the cBFM18 carried an identical mutation Q237K on motA. The identical mutations were also observed by other cBFM groups, which indicates the repeatability of evolution [8]. In contrast to the cBFM18, multiple types of causal mutations in both motA and motB homologues were found in the cBFM17, which indicates that there were several possibilities to functionally integrate the cBFM 17 foreign stator, while a single possibility was available for the cBFM18 foreign stator. Most mutations were single nucleotide polymorphisms (SNPs), though four insertion and deletion mutations were found in cBFM17. Mutations were also found on the non-coding region and at the beginning of motA from cBFM29. These mutations were single nucleotide point mutations including a silent mutation.

Table 5.2 The list of beneficial mutations found on the primary flare motAB homologues and the corresponding cBFM-P strains' chemotaxis speeds.

Strain	Mutations on plasmid		nucleotide seqence	Chemotaxis speed	Stdev
1 (WT)	MotA	MotB		0.9 ± 0.05	0.05
Escherichia coli O111:H- str. 11128					
8.1a-P	A134S LafT		GCG->TCG	ATC	
8.1b-P	A134S LafT / L284R LafT		GCG->TCG / CTG->CGG	0.610	0.04
8.1c-P	A134S LafT / L284R LafT		GCG->TCG / CTG->CGG	0.650	0.065
8.1d-P	A134S LafT / E87K LafT		GCG->TCG / GAA->AAA	0.900	0.04
8.1e-P	A134S LafT		GCG->TCG	ATC	
Bacillus pseudofirmus OF4					
11.1i-P		Q52P MotS	CAG->CCG	0.494	0.028
Listeria monocytogenes EGD-e					
17.1a-P		S39I MotB	AGT->ATT	0.444	0.041
17.1b-P		F53V MotB	TTT->GTT	0.407	0.025
17.1c-P		F53V MotB	TTT->GTT	0.425	0.029
17.1e-P		F53V MotB	TTT->GTT	0.434	0.046
17.1f-P		S39R MotB	AGT->CGT	0.451	0.046
17.1g-P	G172 18bp insertion MotA		ACTCATTCATGCGATGGG	Slow motile	
17.1h-P		F53L MotB	TTT->TTA	0.382	0.048
17.1i-P		F53V MotB	TTT->GTT	0.429	0.049
17.1j-P	G172 18bp insertion MotA		ACTCATTCATGCGATGGG	Slow motile	
17.1k-P		A37 9bp deletion MotB	CCTCCAGTT	0.554	0.029
17.1l-P		F53L MotB	TTT->TTG	0.440	0.043
17.1n-P		S39I MotB	AGT->ATT	0.564	0.041
17.1o-P		S39 15bp deletion MotB	AGTTCAGTTGATGCA	0.635	0.044
Bacillus megaterium DSM319					
18.1e-P	Q237K MotA		CAA->AAA	Slow motile	
18.1f-P	Q237K MotA		CAA->AAA	Slow motile	
18.1g-P	Q237K MotA		CAA->AAA	Slow motile	
18.1i-P	Q237K MotA		CAA->AAA	Slow motile	
18.1j-P	Q237K MotA		CAA->AAA	Slow motile	
18.1k-P	Q237K MotA		CAA->AAA	Slow motile	
18.1m-P	Q237K MotA		CAA->AAA	Slow motile	
Rhodospirillum centenum SW					
29.1a-P	L3M MotA		CTG->ATG	ATC	
29.1b-P	N2N MotA		AAC->AAT	ATC	
29.1d-P	N2N MotA		AAC->AAT	ATC	
29.1e-P	L3M MotA		CTG->ATG	ATC	
29.1f-P	N2N MotA		AAC->AAT	ATC	
29.1g-P	N2N MotA		AAC->AAT	ATC	
29.1i-P	N2N MotA		AAC->AAT	ATC	
29.1k-P	non-coding region		GAT->TAT	3/9 minimal ATC	
29.1l-P	non-coding region		GAT->TAT	7/9 minimal ATC	
29.1m-P	non-coding region		GAT->TAT	5/9 minimal ATC	

The motility of the cBFM-P strains at the population level (chemotaxis assay) was checked to investigate the phenotypic effects of the mutation, prior to the single cell and single motor levels of the functional investigations (section 5.4 and 5.5). The population swimming speeds of the cBFM-P strains were determined by measuring the distance the front of a population of cells had travelled per hour in semi-solid agar 8h after inoculation. The range of chemotaxis motility was from ATC to slow motile and normal motile. The phenotype data was in an agreement with the genotype data (Table 5.2). Single point mutations on either motA or motB were sufficient to make cBFM 11, 17, 18 (the ATC motile compatibility group B) motile, suggesting the possibility of a higher structural similarity between the stators of cBFM 11,17,18 and the stators of *E.coli* BFM. Figure 5.6 shows how many times each mutation occurred and the location of the mutations. Most mutations occurred more than once, especially, a preliminary replicate experiment (performed by Thierry Janssens, postdoc in Beaumont lab, unpublished observations) also found the same mutations A134S, F53V and Q237K. Such observations - same mutations evolving in replicate experiments - were observed by other studies as well [8-10].



(A) Repeatedly occurring mutations were identified on MotA and MotB

	MotA	MotB
Escherichia coli O111:H- str. 11128 **LafTU cBFM8**	E87K (1) A134S (5) L284R (2)	
Bacillus pseudofirmus OF4 **MotPS cBFM11**		Q52P (1)
Listeria monocytogenes EGD-e **MotAB cBFM17**	G172 18bp insertion (2)	A37 9bp deletion (1) S39 15bp deletion (1) S39I/R (3) F53V/L (6)
Bacillus megaterium DSM319 **MotAB putative cBFM18**	Q237K (7)	
Rhodospirillum centenum SW **MotAB"1" cBFM29**	L3M (2) / N2N (5) non-coding region (3)	

Figure 5.6 (A) Single nucleotide mutations and insertion/deletion mutations were found in motA and motB. The numbers in parenthesis represent how many times the mutation occurred. The mutations that were also found in the preliminary replicate evolution experiment are in red. (B) the locations of the mutations are shown on the putative MotAB topology.

Sequence alignments of the 22 motA and motB homologues were performed to see functional, structural, or evolutionary relationships between the sequences (figure 5.7). A sequence alignment arranges multiple DNA or protein sequences, and identifies the regions of similarity between the sequences. The alignment revealed that most mutations remain in a proximity to the previously known functionally important residues, such as torque generating conserved charged residues (R90E, E98K, E150K) on the cytoplasmic loop of MotA [11-14, 17], two proline residues (P173, P222) at the interface between the cytoplasmic domain and the membrane domain of MotA [15], and the putative proton binding residue A32 on MotB TM1 [16]. Previous studies have shown that many mutations on the cytoplasmic loop of MotA either abolished or impair torque generation. In addition, many of the mutations involve in loss or gain of a proline residue, suggesting that these mutations disrupt function by altering the protein conformation rather than by directly affecting the residues that interacts with the FliG rotor components [17]. The sequence alignment revealed that all the mutations identified were not strictly in the conserved residues, and many of them were involved in either a change of charged residue (E87K, Q237K) and or a gain of proline residue (Q52P) (figure 5.7).

Evolutionary modification in cBFM8

In cBFM8, a single mutation of A134S LafT (equivalent to MotA), which showed atypical (ATC) chemotaxis, occurred in all five cBFM8 strains including the evolved cBFM8 strains from the previous replicate evolution experiment. Additional mutations at either L284R or E87K in LafT generated normal chemotaxis, especially the E87K mutation (cBFM8.1d-P) made the strain chemotaxis as fast as the WT strain. The L284R mutation is located at the end of C-terminus site, and the mutation E87K (a reversal charge shift from the negatively charged glutamic acid to the positively charged lysine) is in a proximity to the torque generating (charged) residues. Interestingly, these three residues (E87K, A134S and L284R) substituted their original amino acids to the corresponding *E.coli* MotA amino acids (figure 5.7). For example, the glutamic acid (E) at residue 87 is substituted to the lysine (K), which is the next (one residue after) corresponding amino acid of *E.coli* MotA at this site. This result may provide an insight into how non-compatible components overcome the non-compatibility issue, such as the foreign stators acquired better structural fits to the rest of *E.coli* flagellar motor components.

Evolutionary modification in cBFM11

Unlike the other evolved foreign stators, cBFM11 MotPS of *Bacillus pseudofirmus* is a sodium stator [18]. The mutation Q52P on MotS is in a domain called MotB plug domain (residues 53 to 66 in motB). It is shown that in frame deletion of this domain causes proton flow leakage, influencing cell growth [19, 20]. It is also shown that the stators of *Bacillus* flagellar motors can alter their ion specificity by the mutations (i.e., Q43S, A46Q and M33L) [21, 22], which are near the plug domain. This single point mutation improved the motility from ATC to normal chemotaxis.

Evolutionary modification in cBFM17

One insertion modification on MotA and six different mutations (including two deletions) at three different sites on MotB (at sites A37, S39 and F53) were found in cBFM17. Deletion or insertion mutation is supposed to cause a greater structural change. The 5 AA insertion, repeated 5AA upstream at the G172 of MotA, was found in the periplasmic region between TM3 and TM4, which is the pore forming domain along with the TM of MotB [23]. This insertion mutation strains (17.1g and j) showed the slowest chemotaxis speed (slow motile), though generated second/third (or more) flare trajectories. Two deletion mutations were found near the pore forming domain of MotB (at A37 and S39). The 5AA deletion at S39 strain (cBFM17.1o-P) chemotaxis faster (0.635 ± 0.044) than the 3AA deletion at A37 strain (cBFM17.1k-P) (0,554 ± 0.029). The single nucleotide modifications S39I/R and F53V/L improved chemotaxis motility from the ATC motility of cBFM17 to normal chemotaxis. *E.coli* codon usage bias may took an advantage in the cBFM17 mutations [24]. For example, AGT (S) codon is less frequently used in *E.coli* than ATT (I) or CGT (R), which may facilitate a better expression of the protein.

Evolutionary modification in cBFM18

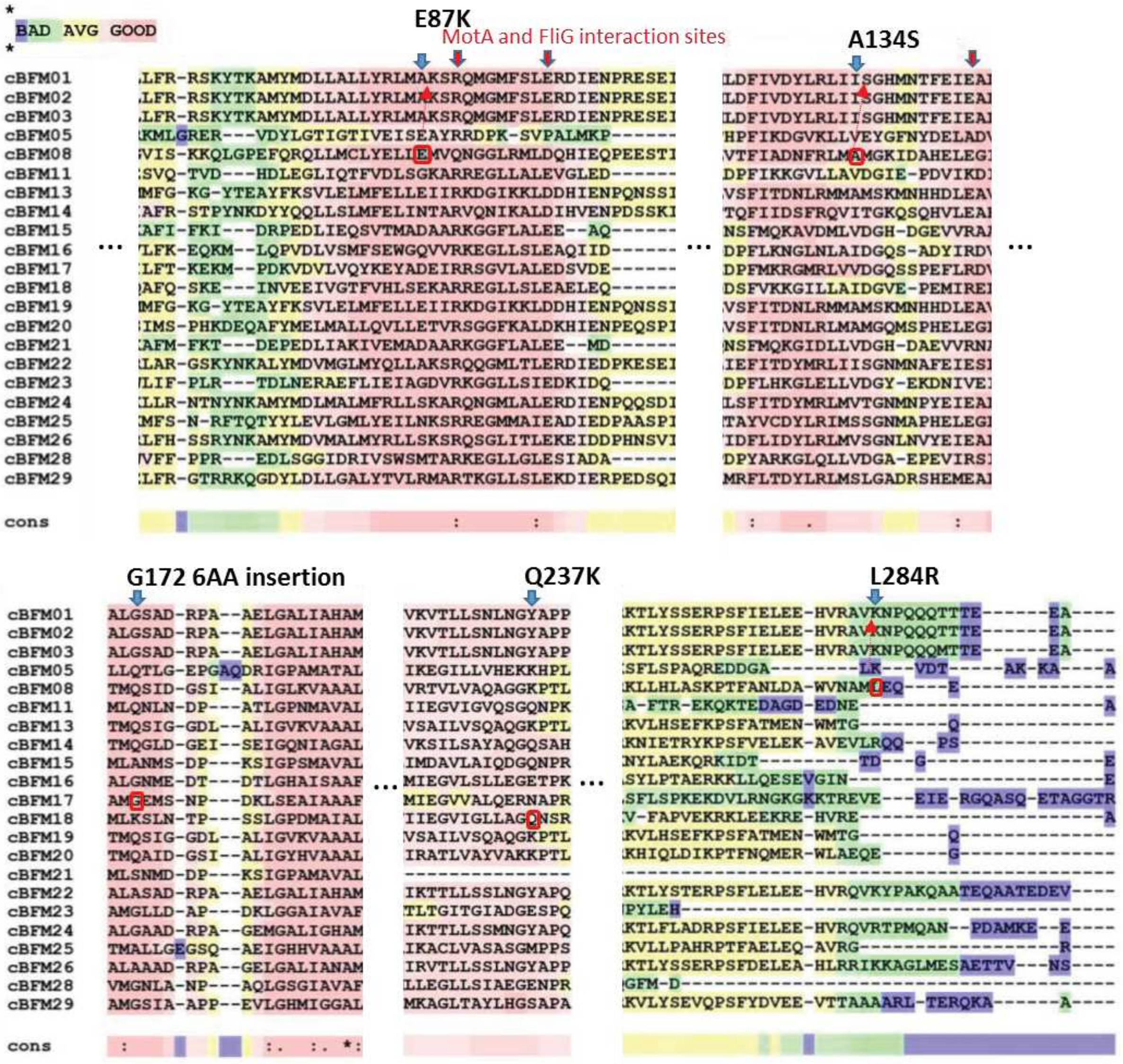
Independently evolved all eight strains of the cBFM18 showed only one single point mutation (Q235K) on MotA. Previous evolution experiment also showed the identical Q235K mutation (mutational recurrence). This mutation improved the ATC motility of cBFM18 to normal yet very slow chemotaxis motility. *E.coli* codon usage bias may influenced such an improvement, since the CAA (Q glutamine) codon is three times less frequently used than the AAA (K lysine) codon.

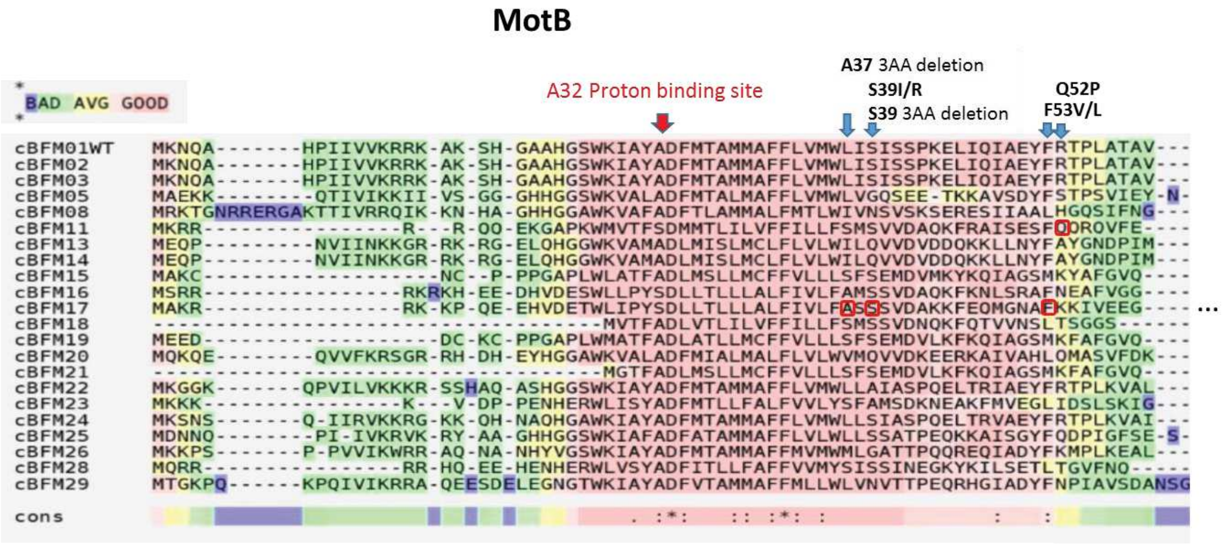
Evolutionary modification in cBFM29

Three mutations (L3M, N2N and in non-coding region) at the first and second AA of MotA were found from cBFM29. The N2N is a silent mutation changed nucleotide codon from AAC to AAT, and L3M mutation changed CTG codon to ATG, which are not favored by *E.coli* codon usage. Considering the location of these mutations, they do not affect the structure of the stator complex neither. The motilities of all cBFM29-P strains were ATC without an exception. Further analysis of their plasmid vectors (by the NGS data from a pool of the entire cBFM29 flares/strains) revealed that there were some mutations in the P15A origin of replication site in their plasmid vectors. Consequently, the copy number of these plasmids went up about 20x than other plasmids, suggesting that the improved ATC motility from non-motile strains was due to an altered expression levels.

In summary, all mutations (except for the cBFM29 mutations) were found near the functional domains: 1) the cytoplasmic loop of MotA where they interact with the rotor component to generate torque and 2) the periplasmic region near the pore forming transmembrane segment of MotB. Thus, most mutations were identified around the well-conserved and functionally essential residues. Perhaps, the underlying reason behind this is due to the fact that the foreign stators that we introduced were essentially wildtype stators, which embraces all the operative functional domains, but not compatible with the rest of the *E.coli* rotor components. Consequently, the mutations that could enhance the structural fitness/compatibility with the rest of the *E.coli* motor components could have improved the functionality. In this regard, the mutations identified here may provide an insight into the structural difference between the *E.coli* (WT) stator and the foreign stator, in a condition that the mutation introduced functional/structural modification instead of expression level modification.

MotA


BAD AVG GOOD
E87K
MotA and FliG interaction sites
A134S
G172 6AA insertion
Q237K
L284R
cBFM01
cBFM02
cBFM03
cBFM05
cBFM08
cBFM11
cBFM13
cBFM14
cBFM15
cBFM16
cBFM17
cBFM18
cBFM19
cBFM20
cBFM21
cBFM22
cBFM23
cBFM24
cBFM25
cBFM26
cBFM28
cBFM29
cons



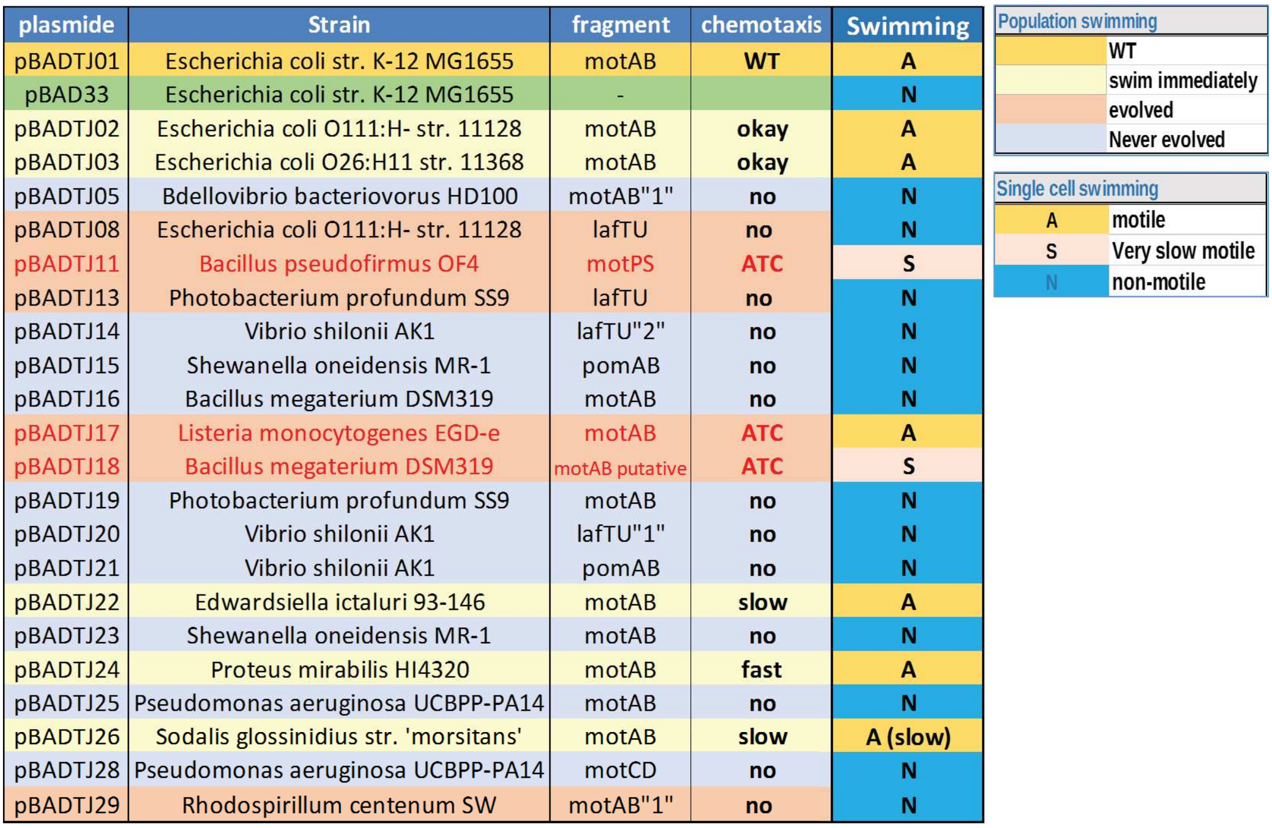
```
cBFM01WT  MKNQA-------HPIIVVKRRK-AK-SH-GAAHGSWKIAYADFMTAMMAFFLVMWLISISSPKELIQIAEYFRTPLATAV---
cBFM02    MKNQA-------HPIIVVKRRK-AK-SH-GAAHGSWKIAYADFMTAMMAFFLVMWLISISSPKELIQIAEYFRTPLATAV---
cBFM03    MKNQA-------HPIIVVKRRK-AK-SH-GAAHGSWKIAYADFMTAMMAFFLVMWLISISSPKELIQIAEYFRTPLATAV---
cBFM05    MAEKK--------QTIVIKKII-VS-GG-GHHGGSWKVALADFMTALMAFFLVMWLVGQSEE-TKKAVSDYFSTPSVIEY-N-
cBFM08    MRKTGNRRERGAKTTIVRRQIK-KN-HA-GHHGGAWKVAFADFTLAMMALFMTLWIVNSVSKSERESIIAALHGQSIFNG---
cBFM11    MKRR----------------R---R-QQ-EKGAPKWMVTFSDMMTLILVFFILLFSMSVVDAQKFRAISESFQDRQVFE----
cBFM13    MEQP---------NVIINKKGR-RK-RG-ELQHGGWKVAMADLMISLMCLFLVLWILQVVDVDDQKKLLNYFAYGNDPIM---
cBFM14    MEQP---------NVIINKKGR-RK-RG-ELQHGGWKVAMADLMISLMCLFLVLWILQVVDVDDQKKLLNYFAYGNDPIM---
cBFM15    MAKC-------------------NC--P-PPGAPLWLATFADLMSLLMCFFVLLLSFSEMDVMKYKQIAGSMKYAFGVQ----
cBFM16    MSRR----------------RKRKH-EE-DHVDESWLLPYSDLLTLLLALFIVLFAMSSVDAQKFKNLSRAFNEAFVGG----
cBFM17    MAKR----------------RK-KP-QE-EHVDETWLIPYSDLLTLLLALFIVLFASSSVDAKKFEQMGNAEKKIVEEG----
cBFM18    -----------------------------------MVTFADLVTLILVFFILLFSMSSVDNQKFQTVVNSLTSGGS------
cBFM19    MEED-------------------DC-KC-PPGAPLWMATFADLATLLMCFFVLLLSFSEMDVLKFKQIAGSMKFAFGVQ----
cBFM20    MQKQE--------QVVFKRSGR-RH-DH-EYHGGAWKVALADFMIALMALFLVLWVMQVVDKEERKAIVAHLQMASVFDK---
cBFM21    -----------------------------------MGTFADLMSLLMCFFVLLLSFSEMDVLKFKQIAGSMKFAFGVQ----
cBFM22    MKGGK-------QPVILVKKKR-SSHAQ-ASHGGSWKIAYADFMTAMMAFFLVMWLLAIASPQELTRIAEYFRTPLKVAL---
cBFM23    MKKK----------------K---V-DP-PENHERWLISYADFMTLLFALFVVLYSFAMSDKNEAKFMVEGLIDSLSKIG---
cBFM24    MKSNS-------Q-IIRVKKRG-KK-QH-NAQHGAWKIAYADFMTAMMAFFLVMWLLSIASPQELTRVAEYFRTPLKVAI---
cBFM25    MDNNQ--------PI-IVKRVK-RY-AA-GHHGGSWKIAFADFATAMMAFFLVLWLLSSATPEQKKAISGYFQDPIGFSE-S-
cBFM26    MKKPS-------P-PVVIKWRR-AQ-NA-NHYVGSWKIAYADFMTAMMAFFMVMWMLGATTPQQREQIADYFKMPLKEAL---
cBFM28    MQRR----------------RR-HQ-EE-HENHERWLVSYADFITLLFAFFVVMYSISSINEGKYKILSETLTGVFNQ-----
cBFM29    MTGKPQ------KPQIVIKRRA-QEESDELEGNGTWKIAYADFVTAMMAFFMLLWLVNVTTPEQRHGIADYFNPIAVSDANSG

cons                                          . :*:   :: :*: :            :    :
```

Figure 5.7 The amino acid sequence alignment of the 22 MotAB homologues. Only the segment of sequences where the mutations were found are shown. The mutation sites are indicated by red squares and the functionally important residues such as MotA-FliG interacting sites (A90, E98, E150) and the proton binding sites (D32) are indicated by red arrows. Note: 6AA insertion mutation means 6AA upstream of G177 is repeated twice.

5.4 Single cell swimming motility of cBFM-P

In the previous sections, the motility of the cBFM was described at the population level in semi solid agar (chemotaxis). In order to appreciate their motility at the level of basic causal element, and to see how the population level chemotaxis motility can be comparable to the single cell or the single motor level of motilities, their swimming ability at the single cell level in liquid media as well as the single motor rotation (section 5.5) were observed. The strains were grown in ½ LB media and their swimming was also observed in ½ LB media to have the same motility condition as in the chemotaxis experiment. Three independent replicates per strains, together with the positive and negative control in each time, were observed and recorded in AVI video files (see chapter 2.4.2 for the detailed experiment procedure). Qualitative measurement of the single cell swimming abilities was recorded manually. Three different swimming states were monitored: (A) most cells (>80%) swim actively, even in a cripple way, (S) a few cells (<5 % of the population) swim slowly or tumble, and (N) no cells swim (see Table 5.3). When cells do not swim, they are stationary with Brownian motions, but when cells tumble or swim slowly, they exhibit directional movements and marked (S, semi-motile). We were interested in the capability of single cell swimming in liquid from the following four groups: 1) the original cBFM strains that showed immediate chemotaxis motility (the compatibility group A), 2) the original cBFM strains that showed ATC motility and the strains evolved to be motile (the compatibility group B and C), 3) the original cBFM strains that never chemotaxis (the compatibility group D) and 4) all the unique primary flare mutants (total of 18 cBFM-P strains) from the compatibility group B and C. Firstly, the ancestor 22 cBFM strains swimming ability in liquid media were examined (Table 5.3).

Table 5.3 List of the cBFM strains with their swimming abilities in the population level (chemotaxis column) and in the single cell level (swimming column).



plasmide	Strain	fragment	chemotaxis	Swimming
pBADTJ01	Escherichia coli str. K-12 MG1655	motAB	WT	A
pBAD33	Escherichia coli str. K-12 MG1655	-		N
pBADTJ02	Escherichia coli O111:H- str. 11128	motAB	okay	A
pBADTJ03	Escherichia coli O26:H11 str. 11368	motAB	okay	A
pBADTJ05	Bdellovibrio bacteriovorus HD100	motAB"1"	no	N
pBADTJ08	Escherichia coli O111:H- str. 11128	lafTU	no	N
pBADTJ11	Bacillus pseudofirmus OF4	motPS	ATC	S
pBADTJ13	Photobacterium profundum SS9	lafTU	no	N
pBADTJ14	Vibrio shilonii AK1	lafTU"2"	no	N
pBADTJ15	Shewanella oneidensis MR-1	pomAB	no	N
pBADTJ16	Bacillus megaterium DSM319	motAB	no	N
pBADTJ17	Listeria monocytogenes EGD-e	motAB	ATC	A
pBADTJ18	Bacillus megaterium DSM319	motAB putative	ATC	S
pBADTJ19	Photobacterium profundum SS9	motAB	no	N
pBADTJ20	Vibrio shilonii AK1	lafTU"1"	no	N
pBADTJ21	Vibrio shilonii AK1	pomAB	no	N
pBADTJ22	Edwardsiella ictaluri 93-146	motAB	slow	A
pBADTJ23	Shewanella oneidensis MR-1	motAB	no	N
pBADTJ24	Proteus mirabilis HI4320	motAB	fast	A
pBADTJ25	Pseudomonas aeruginosa UCBPP-PA14	motAB	no	N
pBADTJ26	Sodalis glossinidius str. 'morsitans'	motAB	slow	A (slow)
pBADTJ28	Pseudomonas aeruginosa UCBPP-PA14	motCD	no	N
pBADTJ29	Rhodospirillum centenum SW	motAB"1"	no	N

Population swimming	
🟨	WT
⬜	swim immediately
🟧	evolved
🟦	Never evolved

Single cell swimming	
A	motile
S	Very slow motile
N	non-motile

The five original cBFMs that were immediately functional in semi-agar (cBFM #2,3,22,24,26, light yellow colored in Table 5.3) were actively swimming in liquid media as well, though they were swimming slower than WT. However, the evolved original cBFM strains (cBFM #8,11,13,17,18,29) showed surprisingly diverse single cell swimming abilities, ranging from no swimmers at all (N) to a few swimmers (S) and active swimmers (A). Especially, unlike our earlier expectation, cBFM17 cells swam in liquid, raising a question on the functional/cellular mechanism of the ATC motility in semi-solid agar. A study [25] also observed this type of functional mismatch between the chemotaxis ability and the swimming ability from several FliG and motA/B double mutations strains. They reported that the mutations permitted a better chemotaxis than swimming motility were in or near the transmembrane regions of both MotA and MotB, suggesting a possible correlation between proton translocation and motor rotation at high load. The other two ATC strains (cBFM 11 and 18) showed a cripple way of slow swimming in liquid, with only few cells exhibiting slow swimming activities or tumbling. The other three evolved strains (cBFM 8, 13, 29) that showed no chemotaxis ability exhibited no swimming activity in liquid, along with the ten non-evolved and non-chemotaxis strains (cBFM 5, 14, 15, 16, 19, 20, 21, 23, 25, 28), as expected.

The results of the single cell swimming assay with the strains that carries the unique mutations of the foreign stators (cBFM-P) are shown in Table 5.4. The motilities of their ancestor strains' (denoted as strain 8, 11, 17, 18 and 29) are also shown. Table 5.4 provides a comprehensive summary of the data to see how the mutations (change of genotype) affect the two types of phenotypic characteristics (chemotaxis and single cell swimming). In brief, all ATC motile stains also showed few "slow and tumbling" swimmers in liquid. All normal chemotaxis strains showed active swimmers in liquid, except that 17.1g strain (slow motile chemotaxis) did not show any sign of motility in liquid.

Table 5.4 List of cBFM-P strains with their motilities in semi solid agar (chemotaxis) and in liquid (swimming)

Strain	Mutations on plasmid		Chemotaxis speed	Swimming
1 (WT)	MotA	MotB	0.9 ± 0.05	A
Escherichia coli O111:H- str. 11128				
8			no	N
8.1a-P	A134S LafT		ATC	S
8.1b-P	A134S LafT / L284R LafT		0.610	A
8.1d-P	A134S LafT / E87K LafT		0.900	A
Bacillus pseudofirmus OF4				
11			ATC	S
11.1i-P		Q52P MotS	0.494	A
Listeria monocytogenes EGD-e				
17			ATC	A
17.1a-P		S39I MotB	0.444	A
17.1b-P		F53V MotB	0.407	A
17.1f-P		S39R MotB	0.451	A
17.1g-P	G172 18bp insertion MotA		Slow motile	N
17.1h-P		F53L MotB	0.382	A
17.1k-P		A37 9bp deletion MotB	0.554	A
17.1l-P		F53L MotB	0.440	A
17.1n-P		S39I MotB	0.564	A
17.1o-P		S39 15bp deletion MotB	0.635	A
Bacillus megaterium DSM319				
18			ATC	S
18.1e-P	Q237K MotA		Slow motile	A (slow)
Rhodospirillum centenum SW				
29			no	N
29.1a-P	L3M MotA		ATC	S
29.1d-P	N2N MotA		ATC	S
29.1k-P	non-coding region		3/9 minimal ATC	S

Single cell swimming	
A	**motile**
S	**Very slow motile**
N	**non-motile**

5.5 Single motor rotation motility of cBFM-P

The functional dynamics of the series of cBFM-P motors at the single motor level were characterized. Single motor rotation was measured by monitoring the rotation of a tethered cell, instead of monitoring the rotation of a bead attached to the flagellar, because the tethered cell assay provides a higher yield to find the rotating cell bodies than the bead assay, and there were total 17 strains to test. Rotation of the cell body also provides a higher load on the motor than the rotation of a bead, which makes a suitable method to measure and compare the functional dynamics of the cBFM motors that underwent to swim in a viscous semi-solid agar environment.

Tethered cell assays were performed by attaching a sheared flagellar on a glass surface, and the rotating cell bodies were monitored. Three independent replicate experiments were performed per sample. A schematic illustration of a tethered cell is shown in figure 5.8a. The evolved foreign stators and non-evolved ancestor foreign stators were expressed in *E.coli* JPA605 strain, which is the same strain used for the fusion stator bead assays (a RP437 derivative MotAB knockout sticky fliC). The cells were grown in 0.1% arabinose induction condition. Their growth time to reach certain OD600 varied; similar or extended growth time was required for the cBFM-P strains to reach OD600 of 0.55 with respect to the WT strain. The extended growth time were no more than 1.5 hours (30% more than for the WT strain), except for the cBFM8.1c. This strain, exceptionally, required much longer time to grow: about three time longer than the WT, which possibly influenced their motor functions. The sticky FliC strain- the strain with more hydrophobic flagellar - was used since it allows an efficient tethering process. See chapter 2.4 and 2.5 for the detailed experiment procedure and analysis. Total 7 different measurements were recorded for the detailed characterizations of each cBFM-P motor: the percentage of observing spinners per field of view (% spinners), mean CCW torque values, the degree of rotation speed variation (torque fluctuation), switching frequency (/s), mean CW time, mean CCW time and CCW bias (figure 5.8).

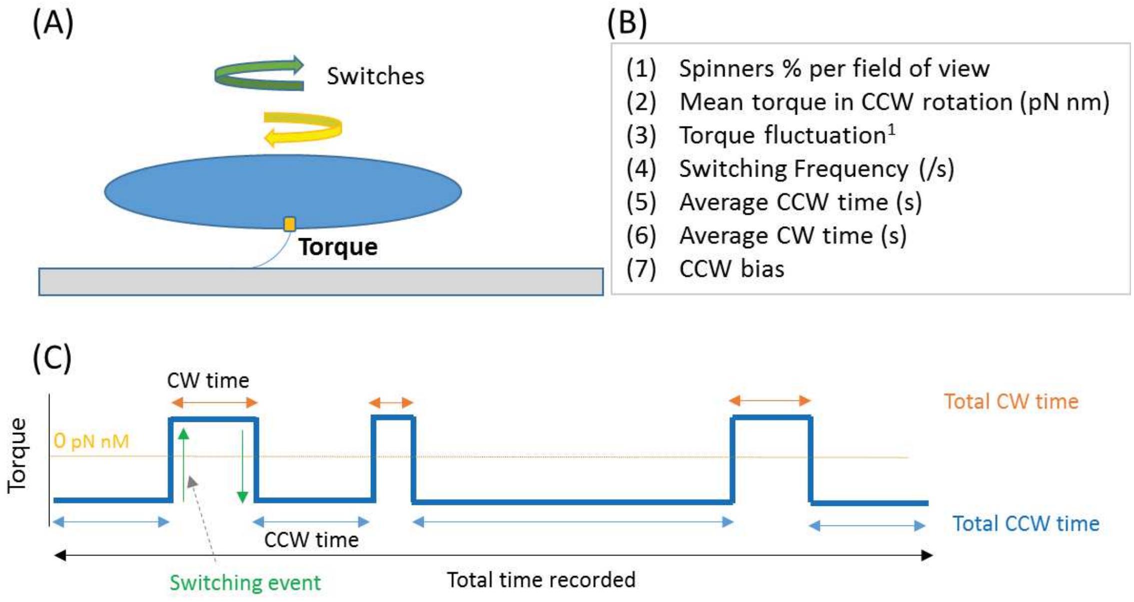


Figure 5.8 (A) a schematic drawing of the tethered cell assays. Sticky flagellar filaments of the motors were stuck on the glass slide. (B) seven different quantitative measurements were made for each cBFM strain. The assay generates a torque vs time trace shown in C. Torque fluctuation[1] was calculated by dividing the standard deviation of CCW torque by the mean torque value. (C) From the torque vs time trace, the six measurements (2) ~ (7) shown in (B) can be obtained. Switching is counted when the direction of rotation switched from either CCW (counter-clockwise) to CW (clockwise) or from CW to CCW crossing the 0-torque value.

The number of tethered cells rotating varied by strains. The percentages of rotating cells in a field of view (FOV) were measured by dividing the number of rotating cells per FOV by the total number of cell in the same FOV. This measurement is important since it can tell about how homogenously the foreign stators are functionally compatible in the large population of cells. The total number of the cells per FOV was in a range of 100 ~ 400 cells, and three different microscopic FOVs from three independent replicate experiments were considered. The likelihoods of finding the spinning cells were high (> 20%) for most of the strains. But, the cBFM 8.1a, 17, 17.1o -P showed ~10%, and the cBFM11, 8.1c-P showed less than 1% likelihood of finding spinning cells. No spinners were found from the cBFM8 strain. The percentage of finding spinners from the cBFM18 was unexpectedly high, despite their previously seen low motility; the population and single cell level of their motility was ATC and few cells crippling swimming motility, respectively.

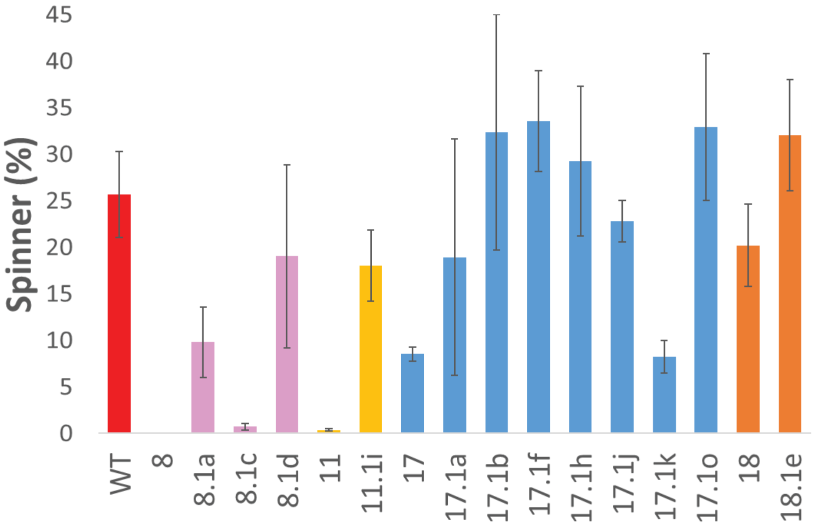


Figure 5.9 The mean percentages of spinners out of total number of cells per field of view. Typically, total number of the cells on a glass surface was in a range of 100-400. The mean percentages (%) were obtained from total of 9 different microscopic fields during the three replicate experiments. The spinners % can be categorized into four: high (>20%), medium (~10%), low (~1%) and none (0%). The error bars indicate standard deviations.

The cell body rotation speeds were converted into torque values, after taking into account the drag coefficients from each cell (see ch. 2.5.1). Unlike the motors from the bead assay, all of them were CCW biased motors, thus the CCW torque values were considered as their torque values. The mean torque values of the cBFM(-P) strains are shown in a bar graph in figure 5.10. The mean torque generated by the WT motors was 470 ± 148 pN nm (the mean rotation speed of WT motor was about 4 ± 1.6 Hz with a maximum rotation speed of ~8 Hz). This is (~66%) lower than the known torque value of WT BFM (~1300 pN nm) [26]. Since we are looking at the relative torque changes of cBFM in comparison to the WT motor, this WT torque difference may not be an issue in this scope of comparison, though this may remain disputable. The four cBFM ancestor strains (8, 11, 17 and 18 cBFM) showed mean torque values of 0 pN nm (cBFM8), 48 ± 21 pN nm (cBFM11), 204 ± 123 pN nm (cBFM17) and 252 ± 131 pN nm (cBFM18), and their evolved motors (cBFM-P) showed either similar or improved torques.

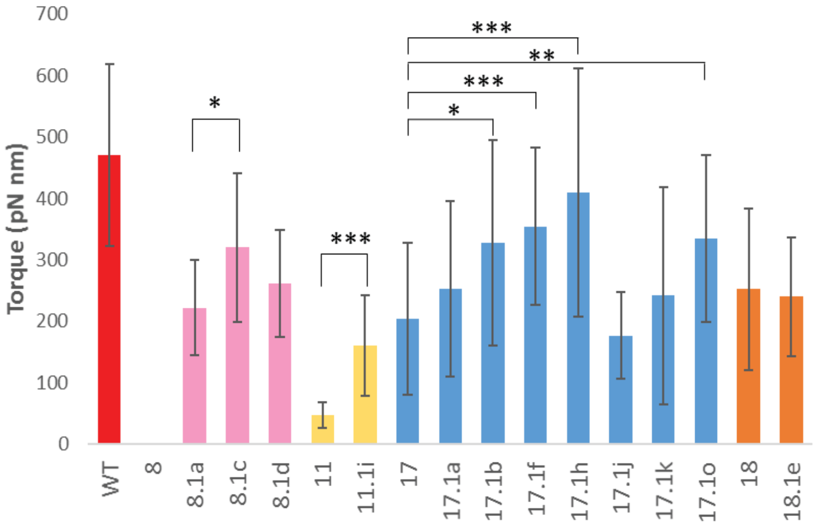


Figure 5.10 Mean torque values of the cBFM motors (both ancestor cBFM 8, 11, 17, 18 and their evolved cBFM-P) and the WT motor. The number of cells measured were: 31, 0, 29, 20, 30, 11, 44, 27, 28, 39, 29, 34, 28, 27, 41, 39 and 35 in the same order as in the graph (left to right). The error bars indicate standard deviations. Statistical t-Test (Two-Sample Assuming Unequal Variances) was performed between the ancestor strains and the evolved strains, and between the evolved strains for cBFM8: *P < 0.01, **P < 0.001, ***P < 0.0001.

The speed and torque distributions of the WT tethered cell is shown in figure 5.11. The speed distributions of other 16 cBFM(-P) strains can be found in appendix C. When compared to the WT motors rotated a 1.1 µm bead, the WT tethered cells (rotated its own body) showed two noticeably different functional characteristics. 1) No CW biased cells were observed from the tethered cells. This is reflected on the speed histogram: only red color CCW biased cells (figure 5.11a) were observed, unlike the bead assay speed histogram generated *both* red color CCW biased cells and blue color CW biased cells (figure 3.6 a). 2) The tethered cells had a higher CW resident time on average (1.7 s for tethered cell and 0.23 s for the bead assay), including the switching events with the extended CW resident time (of more than 1.5 sec). A wider CW resident time distribution of the WT tethered cells can be seen in figure 5.11c. Assuming that rotating a cell body impose a higher load on the motor than rotating a 1.1 µm bead, this result is consistent to the motor switching under high load observed by [33]. Their finding was clockwise (CW) intervals lengthened appreciably at high loading whereas counterclockwise (CCW) intervals remained about the same for all loads. The WT strain used, the growth condition, the induction condition [arabinose]=0.1% and the recording time (3 to 5 min) were the same in both tethered cell assay and bead assay measurements.

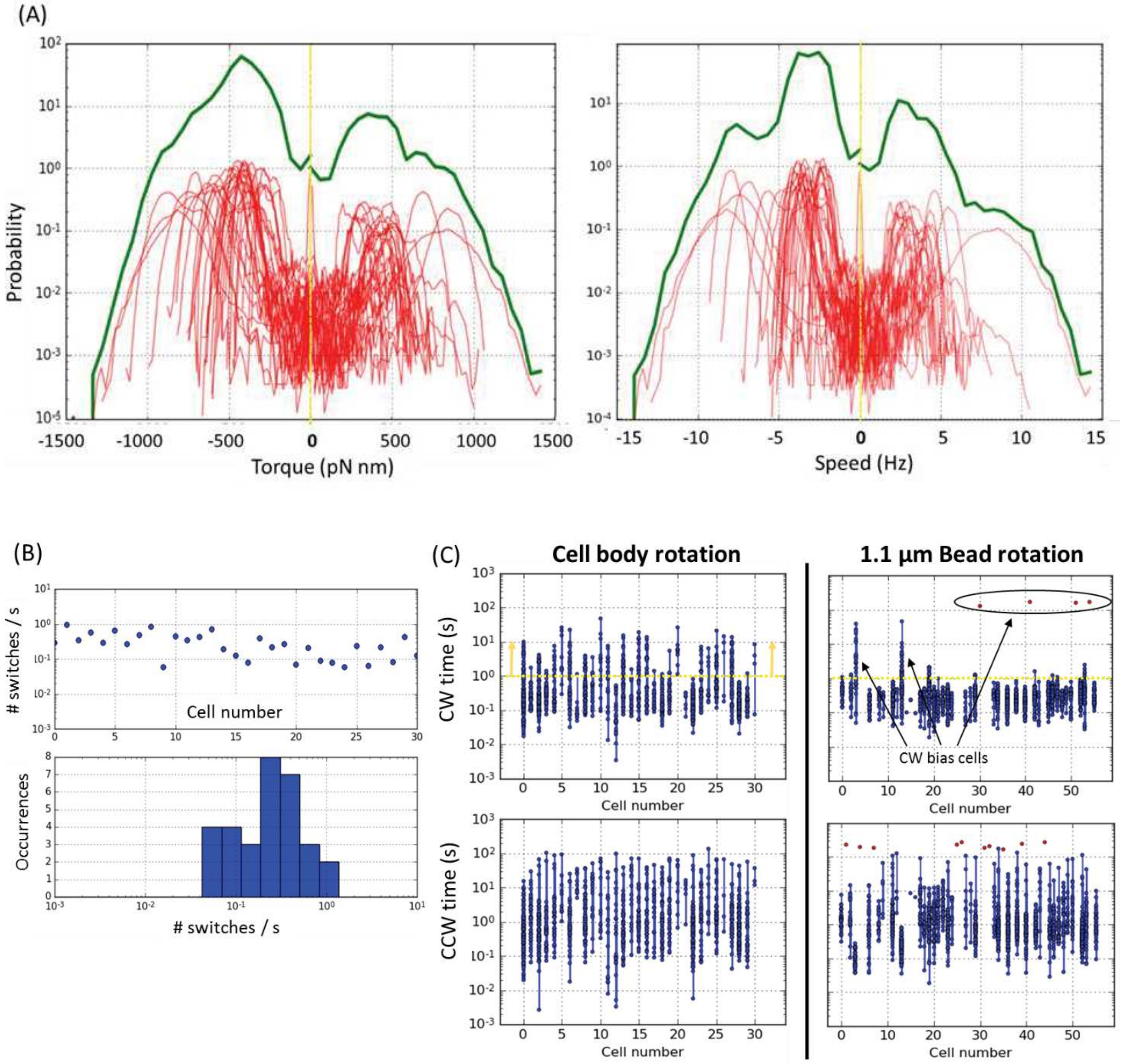
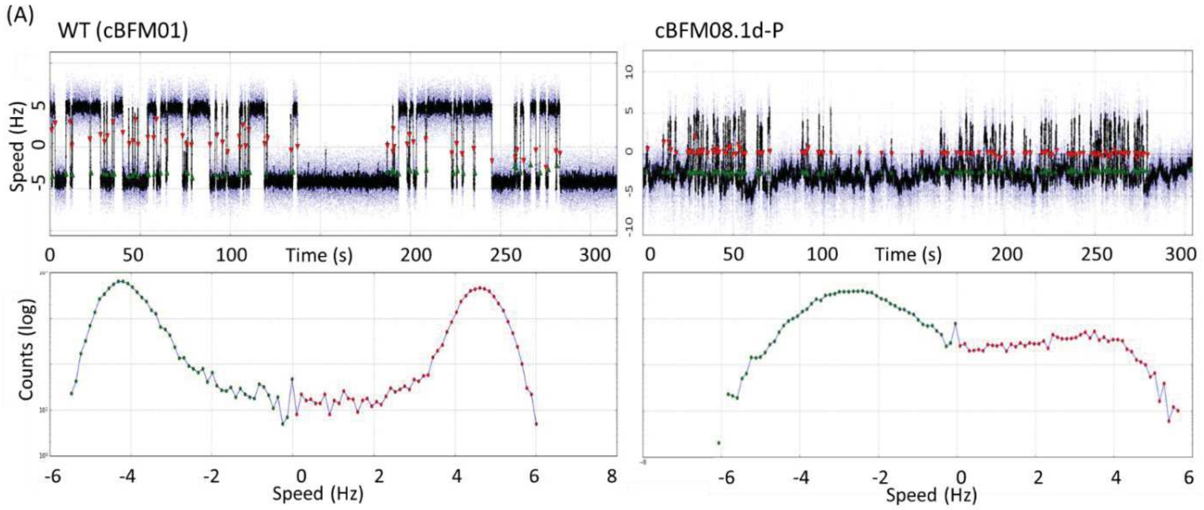


Figure 5.11 Summary of the WT motor tethered cell assay results. (A) torque and speed histograms of the WT tethered cells. The green line represents cumulative histograms of 31 motors. Only CCW biased cells (red color) were observed, whereas CW bias cells (blue color) were also observed during the bead assay (see figure 3.6a). (B) the distribution of switching frequencies are shown by individual cells (upper) and by total switching events (lower). All the tethered cells measured switched during the time course of recording (average time of 262s). Average of switching frequency was 0.31 ± 0.23 (/s), which is close to that of the bead assay 0.32 ± 0.21 (/s). (C) the CW and CCW time distribution of the WT motors are represented by individual cells. Each dot represents switching time (intervals), and the switching events above the yellow line (1s) in CW time graphs represent the CW resident time of more than 1 s.

Interestingly, while recording the rotation speeds of the tethered cells, it was observed that the rotation speeds of the cBFM motors fluctuate noticeably, while the WT motors tend to have a rather constant rotation speed with low or no degree of fluctuation (figure 5.12). The degree of the torque fluctuation (equivalent to the rotation speed variance) was measured by the standard deviation of the mean CCW torque of the individual cell divided by its mean CCW torque. This value indicates how much the rotation speed varies (%) from the mean rotation speed. The mean values for each motor are shown in figure 5.12b. All cBFM motors showed about roughly two-times (in a range of 1.3~2.8 times) higher torque fluctuation values than that of the WT motors. These increased torque fluctuations may reflect the unstable nature of the foreign stator - rotor interactions. According to [27], at a fixed rotation rate (constant number of bound stators), a motor rotates more smoothly (decreased fluctuation) with a higher step size (a step refers to the angular degree moved by a single power stroke mechanism of the stator-rotor interaction force), suggesting that the higher torque fluctuations were due to the lowered step sizes per revolution of the cBFMs. Since the torque of the cBFM was generated by the foreign stators, causing altered stator and rotor interactions in the *E.coli* K-12 motor, it can be speculated that their power stroke mechanism with the *E.coli* FliG rotor units was different than that of the WT motors.



(B)

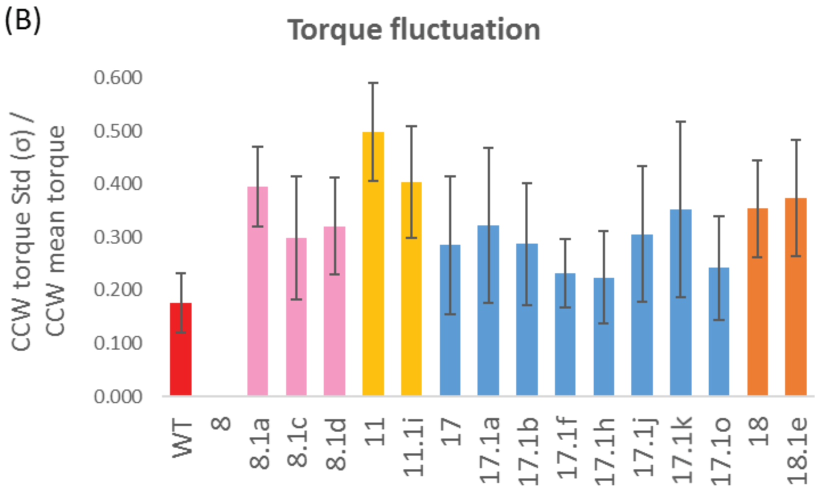


Figure 5.12 (A) example rotation speed vs time traces of the WT motor and the cBFM8.1d motor showing the higher degree of rotation speed variation by the cBFM8.1d motor than that observed by the WT motor. This difference in speed variation is reflected on the speed histogram: a peak around -4 Hz for the WT motor is observed whereas a wider peak around -2 ~ -4 Hz is seen for the cBFM8.1d-P motor (note: negative speed is CCW rotation in the tethered cell assay). (B) Mean torque fluctuations of the cBFM and WT motors. The mean torque fluctuation value (X): the rotation speed varies/fluctuates in X % from the mean speed. The error bars indicate standard deviations.

The mean switching frequencies of the WT and the cBFM motors are shown in figure 5.13, and the distribution of the WT motors switching frequencies was shown in figure 5.11b. The switching frequency was measured by dividing the total number of switching events by the total recording time. All strains showed a wide range of switching frequencies. The mean switching frequency of the WT motors was not very different than that of some of the cBFM motors. However, there were certain percentages (5~20 %) of the cBFM motors that did not switch during the time course of recording (~5 minutes), while all WT motors switched at least once. Especially, about 20% of the cBFM8.1c and 8.1d -P and cBFM18 cells did not show switching events (data not shown). The main difference of the WT motors and the cBFM motors in switching dynamics was more apparent in their CW switching time, as shown in figure 5.15.

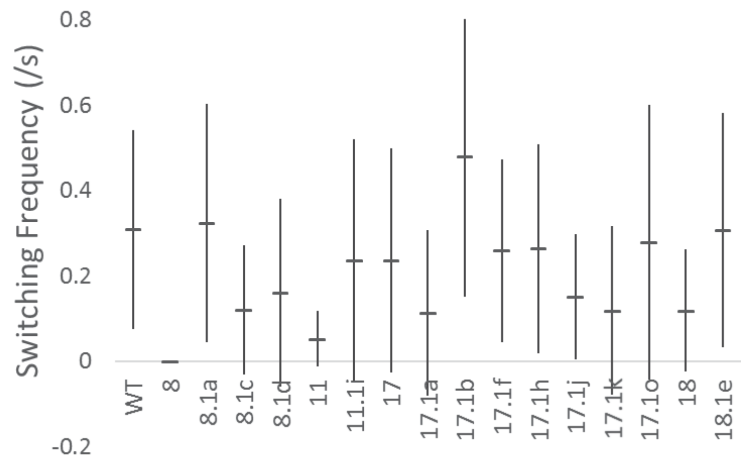


Figure 5.13 Switching frequencies of the WT motors and all the evolved (cBFM-P) and their ancestor cBFM motors. The error bars indicate standard deviations.

Unlike the other three ancestor strains (cBFM8,11 and 18), the ancestor strain cBFM17 already showed a moderate torque generation and switching frequency. According to the switching frequencies and mean torque values relationship plot in figure 5.14, the cBFM strains performed the best (or closest to the WT motor performance) are the four cBFM17 strains, 17.1b, o, f and h. cBFM17.1b has a F53V mutation, 17.1o has a 5AA deletion at S39, 17.1f has a S39R mutation and 17.1h has a F53L mutation (Table 5.2). Thus, the 5AA deletion at S39 motB (BFM17.1o-P) and the single nucleotide point mutations S39R and F53L (cBFM17.1f and h -P, respectively) improved torque generations, and the F53V (cBFM17.1b) improved both torque generation and switching frequency. In contrast, the 3AA deletion at A37 motB (cBFM17.1k-P), the S39I motB mutation (cBFM17.1a-P) and the 6AA insertion between the TM3-TM4 G173 of motA (cBFM17.1j-P) lowered the switching frequency and did not improved the torque generation. Strikingly, even though all mutations were found in the three motB residues (37, 39, 53) consistently, except for the cBFM17.1j, distinctly different functionalities at the single motor level were observed, suggesting the complexity and importance of the residues between the 37 to 53 in motB. For example, the S39R (17.1f-P) mutation, which changed serine to a positively charged arginine, improved torque generation while the S39I (17.1a-P) mutation did not. The F53L (17.1h-P) mutant only improved torque generation while F53V (17.1b-P), which is a proline substitution mutation, improved both torque and switching frequencies.

The strains cBFM8 did not rotate at all, but their evolved foreign stator motors improved both torque and switching frequencies. The single nucleotide change of A134S in LafT (cBFM8.1a) produced a functional motor from the non-functional motor (cBFM8), and the additional mutations (L284R and E87K) further improved torque generations (cBFM8.1b and c). cBFM11 performed the worst, but a mutation Q52P in motS (cBFM11.1i) improved both the torque generation and the switching frequencies. The evolved motor of the cBFM18 (cBFM18.1e) has a Q237K mutation in the cytoplasmic region of motA, where MotA possibly interacts with FliG proteins, and this evolved motor showed an improved switching frequency.

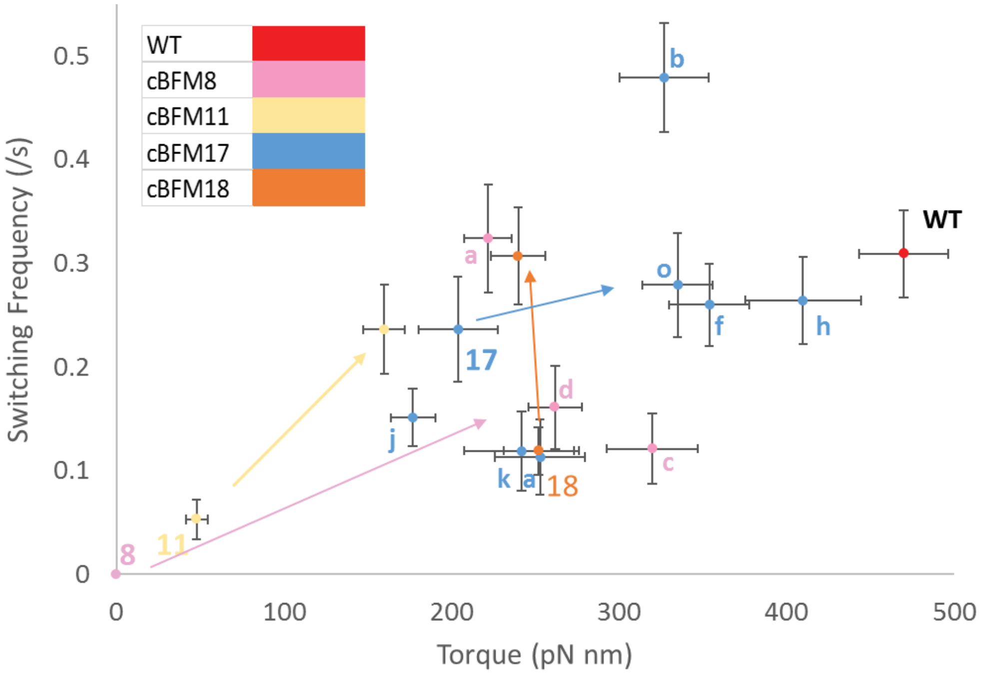


Figure 5.14 Relationship between switching frequencies and torque values of each cBFM strains. The x axis gives the average torque values shown in 5.10, and the y axis gives the average switching frequencies of each cBFM strains shown in 5.13. The arrows indicate how the evolved strains migrated from the ancestor strains (cBFM 8, 11, 17 and 18) toward the WT motor in this plot. The error bars indicate standard errors of the means.

The mean switching frequencies were plotted against the mean CW time and against the CCW bias (figure 5.15). When CW resident time of the WT motors and of the cBFM motors were compared, the cBFM motors showed shorter CW resident time on average than that of the WT motors (x-axis in figure 5.15a). The 17.1k-P and cBFM11 showed longer averaged CW time but with wide error bars. This difference is reflected on their higher CCW bias values (the percentage of time spent in CCW rotation) (x-axis in figure 5.15b). 13 motors out of the 31 WT motors (~42%) had CW time of longer than 1 s, but most of the cBFM motors did not have the CW time of longer than 1 s. In this plot, the three evolved cBFM17 motors (17.1a, 17.1k and 17.1j), which showed a poor performance in both switching and torque generation, were located relatively closely each other as well as the other four evolved cBFM17 motors (17.1b, o, f, h), forming two groups of cBFM17 family. This average CW time is also reflected on the CCW bias values, generating almost an identical mirror relationship when compared to the switching frequencies (except for the cBFM11 and cBFM17.1J, k -P motors).

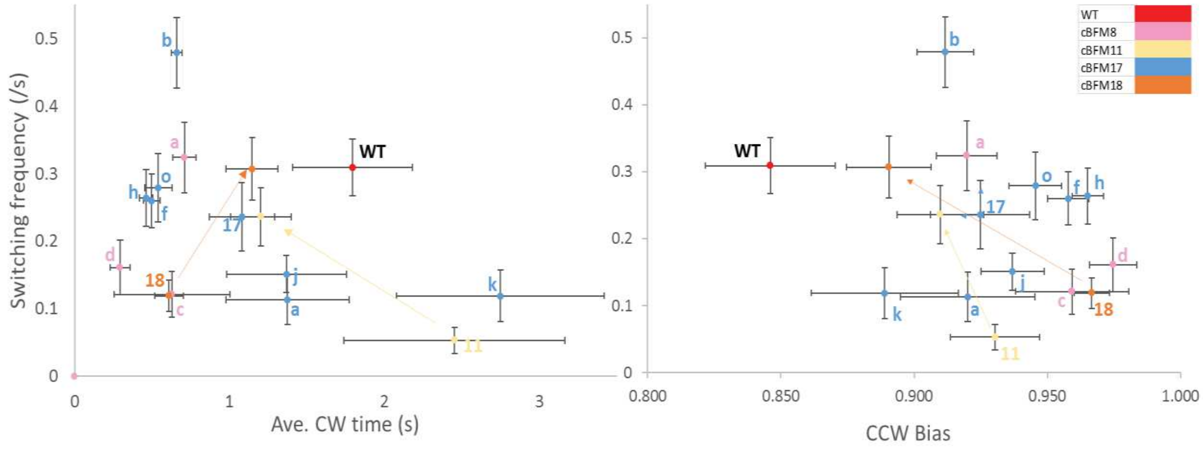
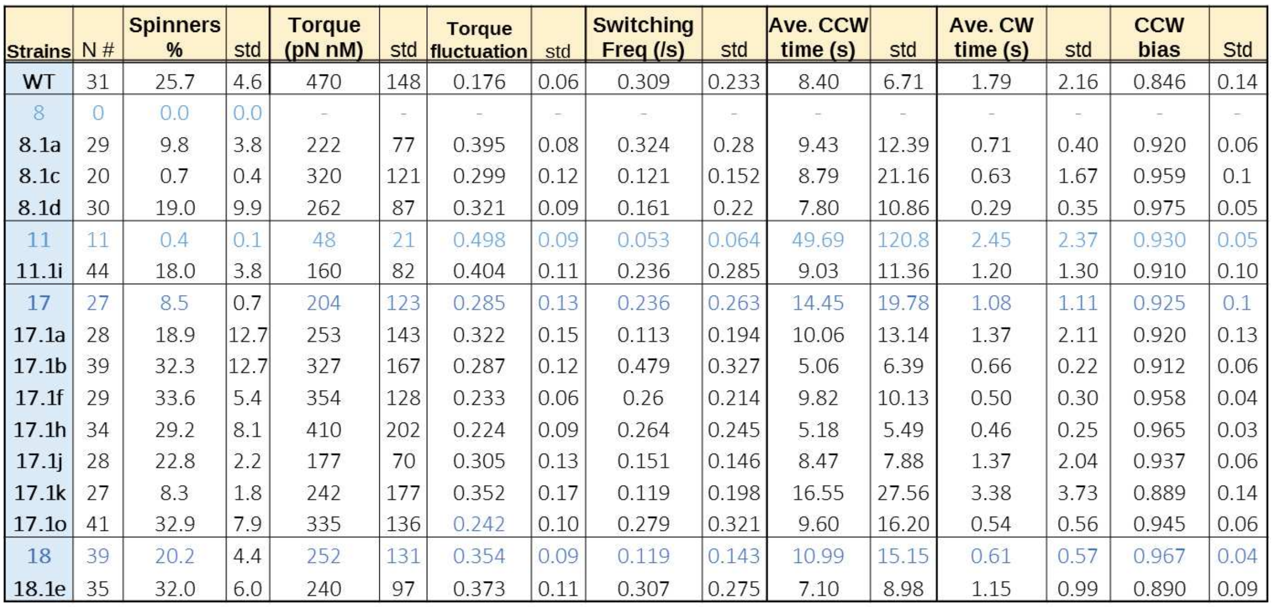


Figure 5.15 Switching frequencies of the WT motor and the evolved (and non-evolved) foreign stator motors are plotted against the average CW resident time (left) and the CCW bias (right). The average CW time and the mean CCW bias of each cBFM strain are shown on the abscissa, and the mean switching frequencies are plotted on the ordinate. The arrows indicate how the ancestor cBFM migrated towards the WT motor. All cells were CCW bias motors. The error bars indicate standard errors of the means.

Table 5.5 Comprehensive data table of the cBFM-P strains (N# is the number of cells measured).



Strains	N #	Spinners %	std	Torque (pN nM)	std	Torque fluctuation	std	Switching Freq (/s)	std	Ave. CCW time (s)	std	Ave. CW time (s)	std	CCW bias	Std
WT	31	25.7	4.6	470	148	0.176	0.06	0.309	0.233	8.40	6.71	1.79	2.16	0.846	0.14
8	0	0.0	0.0	-	-	-	-	-	-	-	-	-	-	-	-
8.1a	29	9.8	3.8	222	77	0.395	0.08	0.324	0.28	9.43	12.39	0.71	0.40	0.920	0.06
8.1c	20	0.7	0.4	320	121	0.299	0.12	0.121	0.152	8.79	21.16	0.63	1.67	0.959	0.1
8.1d	30	19.0	9.9	262	87	0.321	0.09	0.161	0.22	7.80	10.86	0.29	0.35	0.975	0.05
11	11	0.4	0.1	48	21	0.498	0.09	0.053	0.064	49.69	120.8	2.45	2.37	0.930	0.05
11.1i	44	18.0	3.8	160	82	0.404	0.11	0.236	0.285	9.03	11.36	1.20	1.30	0.910	0.10
17	27	8.5	0.7	204	123	0.285	0.13	0.236	0.263	14.45	19.78	1.08	1.11	0.925	0.1
17.1a	28	18.9	12.7	253	143	0.322	0.15	0.113	0.194	10.06	13.14	1.37	2.11	0.920	0.13
17.1b	39	32.3	12.7	327	167	0.287	0.12	0.479	0.327	5.06	6.39	0.66	0.22	0.912	0.06
17.1f	29	33.6	5.4	354	128	0.233	0.06	0.26	0.214	9.82	10.13	0.50	0.30	0.958	0.04
17.1h	34	29.2	8.1	410	202	0.224	0.09	0.264	0.245	5.18	5.49	0.46	0.25	0.965	0.03
17.1j	28	22.8	2.2	177	70	0.305	0.13	0.151	0.146	8.47	7.88	1.37	2.04	0.937	0.06
17.1k	27	8.3	1.8	242	177	0.352	0.17	0.119	0.198	16.55	27.56	3.38	3.73	0.889	0.14
17.1o	41	32.9	7.9	335	136	0.242	0.10	0.279	0.321	9.60	16.20	0.54	0.56	0.945	0.06
18	39	20.2	4.4	252	131	0.354	0.09	0.119	0.143	10.99	15.15	0.61	0.57	0.967	0.04
18.1e	35	32.0	6.0	240	97	0.373	0.11	0.307	0.275	7.10	8.98	1.15	0.99	0.890	0.09

5.6 Discussion

How macromolecular complexes such as molecular machines have evolved is a challenging yet very festinating question. Compositional evolution suggests that they might have evolved through the stepwise addition of pre-existing protein components [28]. The BFM is an excellent model system that can provide an experimental insight to test this compositional evolution theory [29-31]. Prior to the present study, the modularity and evolvablability of BFM is tested by exchanging the orthologues of the stators in *E.coli* BFM, which we called cBFM. Some of the non-compatible foreign stators in cBFM were evolved to be functionally compatible, and the result of this evolution experiment was shown in table 5.1. Followed up from this evolution experiment, 1) the mutations occurred on motA and motB during the first stage of the evolution and 2) their functionalities (in three levels of motilities) were examined to understand the molecular details of the evolutionary transitions.

When the evolved foreign stators were put back to the *E.coli* K-12 ΔMotAB strain (resulting strains were called cBFM-P), more than half (~62%) of the cBFM-P strains showed improved chemotaxis ability than their ancestors, suggesting that the modification on the stator genes played a vital role during evolution. Total 15 different beneficial mutations on motA and motB foreign genes were identified. Strong mutational variations between the different cBFM groups (strain-specific constraints) were also observed (Table 5.2 and figure 5.6). Most mutations were single nucleotide polymorphisms (SNPs), though four insertion and deletion mutations were found on motA and motB from cBFM17. The mutations were found near the functional sites of the stators: stator-rotor interacting cytoplasmic loops and the proton channel forming domains. Some of them were repeatedly discovered by independently evolved strains as well. This mutational recurrence suggests that evolution is repeatable by favoring certain genetic modifications better than others [9].

The three-different level of motilities (population, single cell and single motor) of the non-evolved cBFM (ancestors) and their evolved cBFM-P strains were examined. When their population swimming motilities and their single cell swimming motilities were compared, these two motilities were mostly in agreements, except for the cBFM17 strain (Table 5.3 and 5.4). Unexpectedly, cBFM17 showed atypical motility (ATC) in semi-solid agar but normal active swimming in liquid. Other ATC strains showed a few fractions of cells tumbling/crippling way of swimming in liquid.

Their motilities at the single motor level revealed more detailed functional differences. In comparison to the WT motor, all cBFM showed reduced torque generation (~53 %) and switching frequencies (~70%) on average (figure 5.14). Two noticeably different functions in cBFM and WT motors was observed. First is the higher degree of the cBFM rotation speed variations, which measured as torque fluctuations. WT motors usually rotated at a fixed rotation speed without a much speed fluctuation, while most of the cBFM rotated in wider range of rotation speeds, exhibiting higher speed fluctuations (figure 5.12). This fluctuation analysis of the rotational speeds generated by the foreign stators may suggest the different stepping mechanisms of the WT motors and of the motors driven by foreign stators [27]. The dwell time at CW rotation of the cBFM-P was also relatively shorter than WT; cBFM-P did not rotate in CW for extended period time of more than 1 s, while certain percentage of the WT motors rotated in CW for a period of more than 1 s (note that all tethered cells were CCW biased motors).

The evolved foreign stator motors improved motor performances compared to their ancestor motors. The cBFM8.1a-P with A134S mutation improved the motor from non-functional to functional motor. The cBFM8.1d-P (double mutations E87K and A134S on motA) showed the most improved functionality in consideration of the three parameters: percentage of spinners in total number of cells (spinners percentage), torque and switching frequency. These two mutations E87K and A134S are located in near the two well-known stator-rotor interaction sites (R90 and E150). The cBFM8.1a-P (A134S) showed improved torque generation and switching frequency (about twice higher than cBFM8.1d-P), but a lower spinner percentage (~10%) than cBFM8.1d-P. Thus, an additional mutation E87K on cBFM8.1d-P improved the spinner percentage (~20%), but lowered its switching frequency. cBFM11 showed a poor motor performance, including the very few spinners (~0.04 %) per field of view. Its evolved motor – the cBFM11.1i-P with a beneficial mutation of Q52P MotB- improved motor functions in all three parameters (18% of finding spinners, 34% of torque generation and 76% of switching frequency of the WT motor). The mutation (Q52P) is near the pore forming domain, and a specific implication on the ion specificity of this domain was discussed in [21]. The cBFM18 motors, unexpectedly, functioned well (except their switching ability) despite their poor performance in the previous two motilities assays. The cBFM18.1e-P, with a mutation Q237K on cBFM18, improved switching frequencies. The evolved stators of the cBFM17 seem to have divided into two groups: the cBFM 17.1 b, f, h, o, which improved motor functions and the cBFM 17.1 a, j, k, which did not. As a final point, the figure 5.14 revealed that the cBFM phenotypes (torque generation and switching frequency) can be varied in all combinations by mutations, i.e., one up while another is stable or both up. This suggests that these phenotypes can be modulated by mutations in an orthogonal manner. In other words, different BFM properties can by independently tuned by mutations. This lack of pleiotropic effects suggests that the mutations are modularly encoded within the stator, thereby increase the evolvability of a system.

The cBFM 29 had mutations at the second and third residues of the N-terminus and in the non-gene coding region of the plasmid vectors, as well as the mutations at the origin of replication on the vector. Thus, it is highly speculated that the improved motilities observed by the evolved cBFM29 strains are due to the altered gene expression level [6] or due to the N-terminal codon bias [32]. The gene expression level can be altered by protein folding rates, codon bias, other modifications on the vectors such as DNA methylation (epigenetics). Therefore, it is possible that an altered foreign stator expression level, which is not direct the structural and functional element, influenced their improved chemotaxis motility during the evolution experiment. To check this possibility, the foreign stator protein expression level analysis will be carried out as a follow up study.

Reference

1. Paulick, A., Koerdt, A., Lassak, J., Huntley, S., Wilms, I., Narberhaus, F., & Thormann, K. M. (2009). Two different stator systems drive a single polar flagellum in Shewanella oneidensis MR-1. *Molecular Microbiology*, *71*(4), 836–850.
2. Fujinami, S., Terahara, N., Lee, S., & Ito, M. (2007). Na(+) and flagella-dependent swimming of alkaliphilic Bacillus pseudofirmus OF4: a basis for poor motility at low pH and enhancement in viscous media in an "up-motile" variant. *Archives of Microbiology*, *187*(3), 239–47.
3. Ito, M., Hicks, D. B., Henkin, T. M., Guffanti, A. A., Powers, B. D., Zvi, L., … Krulwich, T. A. (2004). MotPS is the stator-force generator for motility of alkaliphilic Bacillus, and its homologue is a second functional Mot in Bacillus subtilis. *Molecular Microbiology*, *53*(4)
4. Togashi, F., Yamaguchi, S., Kihara, M., Aizawa, S. I., & 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. *Journal of Bacteriology*, *179*(9), 2994–3003.
5. Mohari, B., Licata, N. a, Kysela, D. T., Merritt, P. M., Mukhopadhay, S., Brun, Y. V, & Setayeshgar, S. (2015). Novel Pseudotaxis Mechanisms Improve Migration of Straight- Swimming Bacterial Mutants Through a Porous Environment, *6*(2), 1–12.
6. Lind, P. A., Tobin, C., Berg, O. G., Kurland, C. G., & Andersson, D. I. (2010). Compensatory gene amplification restores fitness after inter-species gene replacements. *Molecular Microbiology*, *75*(5), 1078–1089.
7. Desai, M. M., & Fisher, D. S. (2007). Beneficial mutation-selection balance and the effect of linkage on positive selection. *Genetics*, *176*(3), 1759–1798.
8. Bauer, B., & Gokhale, C. S. (2015). Repeatability of evolution on epistatic landscapes. Scientific Reports, 5, 9607.
9. Gompel, N., & Prud'homme, B. (2009). The causes of repeated genetic evolution. *Developmental Biology*, *332*(1), 36–47.
10. Justin R. Meyer1,2, Devin T. Dobias3, Joshua S. Weitz4, Jeffrey E. Barrick2,5, Ryan T. (2012) Repeatability and Contingency in the Evolution of a Key Innovation in Phage Lambda. Science. 335
11. Lair, D. A. F. B. (1998). Electrostatic interactions between rotor and stator in the bacterial flagellar motor, *95*(May), 6436–6441.
12. Morimoto, Y. V, Nakamura, S., Kami-ike, N., Namba, K., & Minamino, T. (2010). Charged residues in the cytoplasmic loop of MotA are required for stator assembly into the bacterial flagellar motor. *Molecular Microbiology*, *78*(5), 1117–29.
13. Morimoto, Y. V, Nakamura, S., Hiraoka, K. D., Namba, K., & Minamino, T. (2013). Distinct roles of highly conserved charged residues at the MotA-FliG interface in bacterial flagellar motor rotation. Journal of Bacteriology, 195(3), 474–81.
14. Takekawa, N., Kojima, S., & 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. Journal of Bacteriology, 196(7), 1377–85.
15. Braun, T. F., Poulson, S., Gully, J. B., Empey, J. C., Van Way, S., Putnam, a, & Blair, D. F. (1999). Function of proline residues of MotA in torque generation by the flagellar motor of Escherichia coli. *Journal of Bacteriology*, *181*(11), 3542–51.
16. Asai, Y., Yakushi, T., Kawagishi, I., & Homma, M. (2003). Ion-coupling Determinants of Na+-driven and H+-driven Flagellar Motors. *Journal of Molecular Biology*, *327*(2), 453–463.
17. Zhou, J., & Blair, D. F. (1997). Residues of the cytoplasmic domain of MotA essential for torque generation in the bacterial flagellar motor. *Journal of Molecular Biology*, *273*(2), 428–39.

18. Fujinami, S., Terahara, N., Lee, S., & Ito, M. (2007). Na(+) and flagella-dependent swimming of alkaliphilic Bacillus pseudofirmus OF4: a basis for poor motility at low pH and enhancement in viscous media in an "up-motile" variant. *Archives of Microbiology*, *187*(3), 239–47.

19. Hosking, E. R., Vogt, C., Bakker, E. P., & Manson, M. D. (2006). The Escherichia coli MotAB Proton Channel Unplugged. Journal of Molecular Biology, 364(5), 921–937.
20. Kojima, S., Imada, K., Sakuma, M., Sudo, Y., Kojima, C., Minamino, T., Namba, K. (2009). Stator assembly and activation mechanism of the flagellar motor by the periplasmic region of MotB. Molecular Microbiology, 73(4), 710–8.
21. Terahara, N., Krulwich, T. a, & Ito, M. (2008). Mutations alter the sodium versus proton use of a Bacillus clausii flagellar motor and confer dual ion use on Bacillus subtilis motors. *Proceedings of the National Academy of Sciences of the United States of America*, *105*(38).
22. Terahara, N., Sano, M., & Ito, M. (2012). A Bacillus flagellar motor that can use both Na+ and K+ as a coupling ion is converted by a single mutation to use only Na+. PloS One, 7(9), e46248.
23. Kim, E. a, Price-Carter, M., Carlquist, W. C., & 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*(43), 11332–9.
24. Codon usage in *E.coli* genes data table. The arabidopsis Research Companion. http://www.sci.sdsu.edu/~smaloy/MicrobialGenetics/topics/in-vitro-genetics/codon-usage.html
25. Togashi, F., Yamaguchi, S., Kihara, M., Aizawa, S. I., & 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. *Journal of Bacteriology*, *179*(9), 2994–3003.
26. Berg, H. C. (2008). Bacterial flagellar motor. *Current Biology*, *18*(16), R689–91.
27. Samuel, a D., & Berg, H. C. (1995). Fluctuation analysis of rotational speeds of the bacterial flagellar motor. *Proceedings of the National Academy of Sciences of the United States of America*, *92*(8), 3502–3506.
28. Marsh, J. a., & Teichmann, S. a. (2015). Structure, Dynamics, Assembly, and Evolution of Protein Complexes. Annual Review of Biochemistry, 84(1), 551–575.
29. Snyder, L. a S., Loman, N. J., Fütterer, K., & Pallen, M. J. (2009). Bacterial flagellar diversity and evolution: seek simplicity and distrust it? *Trends in Microbiology*, *17*(1), 1–5.
30. Liu, R., & Ochman, H. (2007). Stepwise formation of the bacterial flagellar system, *104*(27).
31. Gophna, U., Ron, E. Z., & Graur, D. (2003). Bacterial type III secretion systems are ancient and evolved by multiple horizontal-transfer events. *Gene*, *312*(1-2), 151–163.
32. Kennedy, M. C., Stout, C. D., Thomson, A. J., Holm, R. H., Hagen, K. S., Babcock, G. T., Hatchikian, E. C. (2013). Causes and Effects of N-Terminal Codon Bias in Bacterial Genes, 475–480.
33. Bacterial flagellar switching under load. (2003). *Nature*, *423*(June), 2003.

Chapter 6. Concluding remarks and future researches

In chapter 3, we quantitatively characterized switching dynamics of the BFM driven by the stators tagged by three different FPs (fusion stators). Owing to the unexpected switching behaviors observed by these motors, the symmetric nature of the bi-directional rotations of the BFM was re-visited. The asymmetric switching (ASW) of the fusion stator motors and the incomplete switching described in the conformational spread model (described in figure 1.8a) share a similar feature in a sense that no symmetry of the CW and CCW rotation speeds is reached during a single switch. However, the ASW is distinguishable from the incomplete switches for the following reasons. 1) the ASW was observed consistently without an exception, whereas the incomplete switches occur occasionally due to the incomplete conformational change of the entire C-ring. 2) the ASW was accompanied by three other switching modifications (prolonged CW or CCW resident time, the reduced switching frequencies and the extended switching time), revealing that various aspects of the switching dynamics of the motor are altered for the ASW of the fusion stator. 3) the ASW has a tag-specificity, meaning that the level of asymmetry is different by which FP is fused to and by the presence or absence of linkers. Furthermore, a tendency of the FP tagged motors (as well as for the WT motors) to pause or slow down around 0 Hz during switching events were often observed (appendix D), suggesting the existence of a transition state (at 0 Hz) between the CW and the CCW states.

The switching dynamics of the fusion stator motors can be in line with the mechanics of the torque generation model that described in figure 1.5. This model takes the four subunits of MotA as a bundle of four gears, where two of each are responsible for either CCW rotation or CW rotation via contact with the FliG proteins. The conformational change of the C-terminus torque generating domain of the FliG (~180 °C) facilitates this transition, so that it interacts with each two pairs of MotA. In a scenario that the FP-tag hinders the transition of the FliG conformational change, which described in figure 3.25, the ASW of the fusion stator motors is caused by the FliG stuck in somewhere between these two pairs of MotA. Although an additional experiment will be required to validate this model, the fusion stator study may provide an important insight in our current understanding of how the symmetric bi-directional rotations can be achieved.

An experiment that can be done to test the mechanics of torque generation model is by examining the symmetricity of torque generated by the product of two MotA genes fused together, one of which is mutated to be non-functional (MotA-MotA*). If this model is valid, the torque generated by each CW and CCW rotation state can be dramatically different, because one side of

MotA responsible for a certain direction is designed to be non-functional. Control experiments can be the same assays with multiple other combinations of the MotA fusions, such as two wildtype MotA fused together (MotA-MotA) and five others (WT, MotA*, MotA*-MotA, MotA-MotA*, motA*-motA*). Such experiment, together with the fusion stator study, may provide an important insight to support the model by [2].

Another remark from this thesis study is that we observed the motors rotate exclusively either CCW or CW in strains lacking CheY gene (the chemotaxis regulator) [MT03 in Appendix B]. This result can be a useful addition to the previous report that the motors rotate exclusively CCW in such a motor. The existence of CW-biased motors was also prominent throughout the whole fusion stator bead assay, raising a question on the nature of bi-directionality once more. Furthermore, the dwell times in CW rotation (CW resident time) of the motors rotating a cell body were longer than those observed in motors rotating a 1.1 μm bead. This observation is consistent with the previous observation that the CW resident time lengthened appreciably at high load (above 1.44 μm beads) [1]. The motors rotating a cell body showed CCW bias exclusively as well, suggesting that there are certain correlations between the bias, the time spent in CW rotation state and the load. Since the rotational direction is determined by the conformational states of the C-ring of the motor, this suggests that the C-ring can be in either CW or CCW state only when the imposing load is relatively not too high (< a 1.1 μm bead), but the C-ring can form CCW state only when the imposing load is relatively high (like a cell body). The dwell times in CW rotation were also altered greatly by the FP-tagged stators and by the foreign stators (cBFM). The motors with the FP-tags showed lengthened CW resident time (as if they are experiencing a higher load than the actual load of a 1.1 μm bead), on the other hand, the cBFM did not show the lengthened CW resident time like the WT motors did in the same condition. Identifying the correlation between the load on the motor and the dwell time of each CW and CCW state may provide an important step forward in both mechano-sensing and the switching dynamics of the motor.

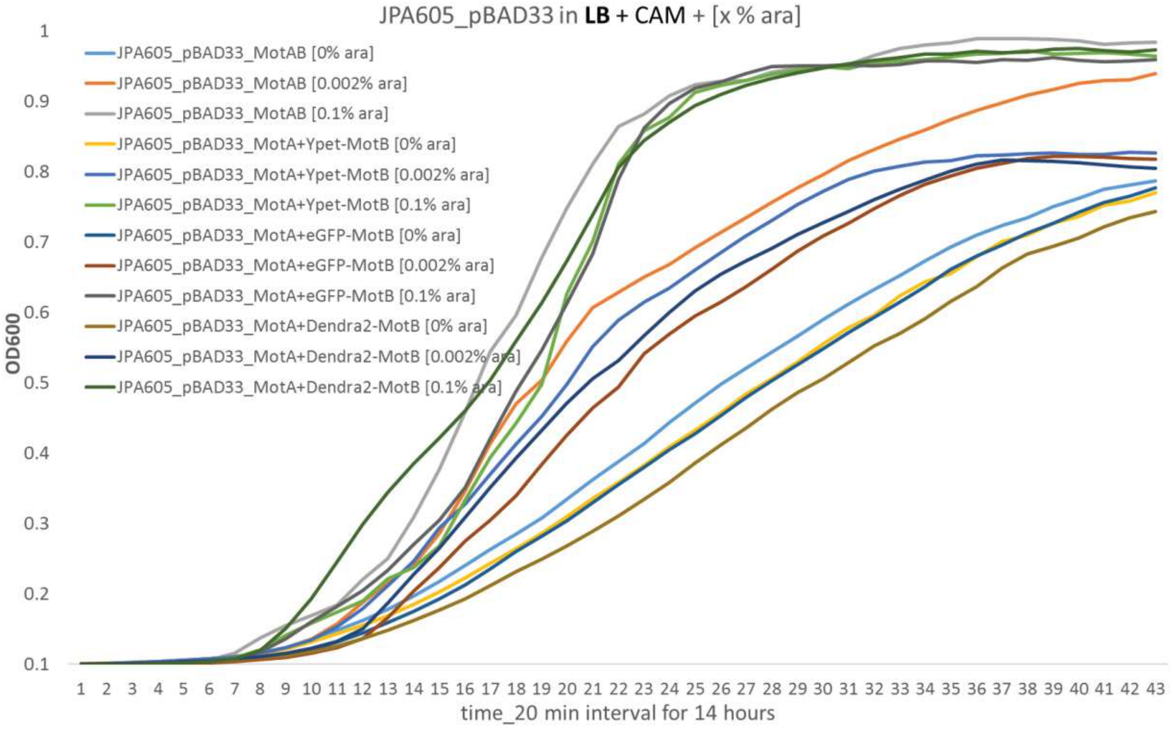
The discovery that an anchored component (MotB) of a multi-protein complex diffuses in the membrane and exchanges rapidly with the motor expanded the conventional ‘static’ view of molecular complexes [2]. The underlying mechanisms of the stator turnover are not clear. However, one may speculate that the mechanics involved in torque generation (energy transduction as a result of the conformational change and the interactions with the FliG proteins) eventually damage or alter the functionality of each stator. Typically, the BFM rotates at a constant speed for an extended period during bead assays, without making any clear signs of stators leaving or engaging. Therefore, stator turnover seems remarkably fast and efficient, likewise, prompt stator activation and inactivation mechanism around the motor seems inevitable as well. Stator turnover opened numerous intriguing questions: is each stator only used for a certain period and degraded (finite functional life time), or are they recycled? Are there functional and physiological benefits of the stator turnover? Do components of other macromolecular complexes also undergo dynamic exchange? How the stator turnover rates change in relation to the stators expression level, in relation to the mechanical load, and in relation to the ion motive force? The last question can be answered by observing the rates of the successive incorporation of the stators in the different conditions that can alter the following three parameters: the stator expression level can be altered by different inducer concentrations, the mechanical load can be altered by applying magnetic force to trap the bead rotated by the motor, the ion motive force can be altered by applying different sodium concentrations buffer. PomAPotB (a sodium stator that can be utilized in *E.coli*) should be used here instead of MotAMotB (proton stator).

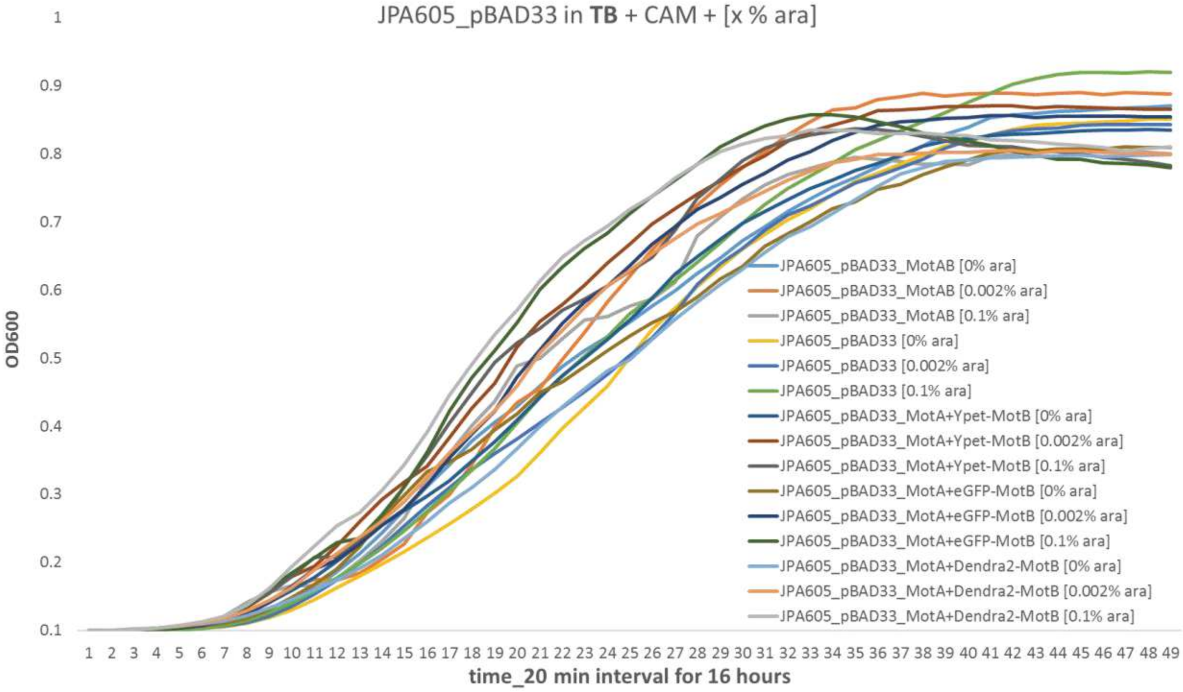
References:

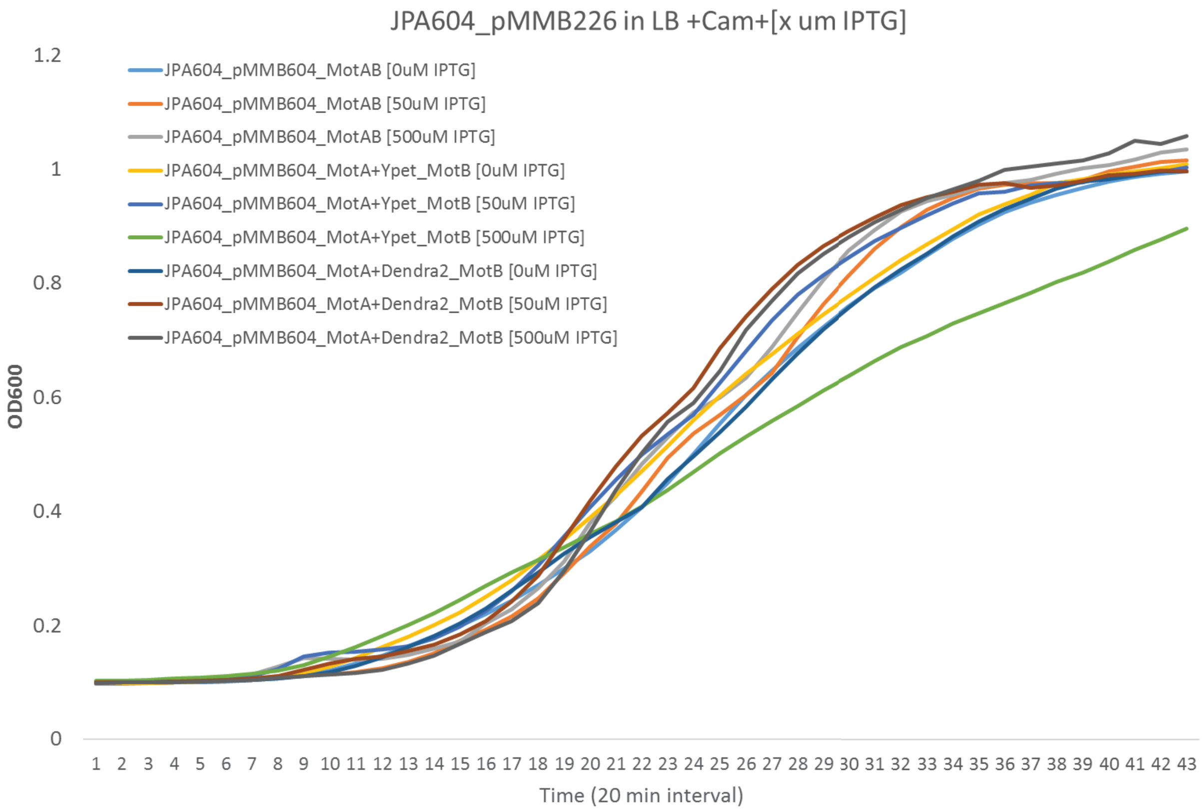
1. Karen A. Fahrner, Ryu, W. S., & Berg, H. C. (2003). Bacterial flagellar switching under load. Nature, 423(June), 2003.
2. Mandadapu, K. K., Nirody, J. a, Berry, R. M., & Oster, G. (2014). Mechanics of torque generation in the bacterial flagellar motor, 1–18.

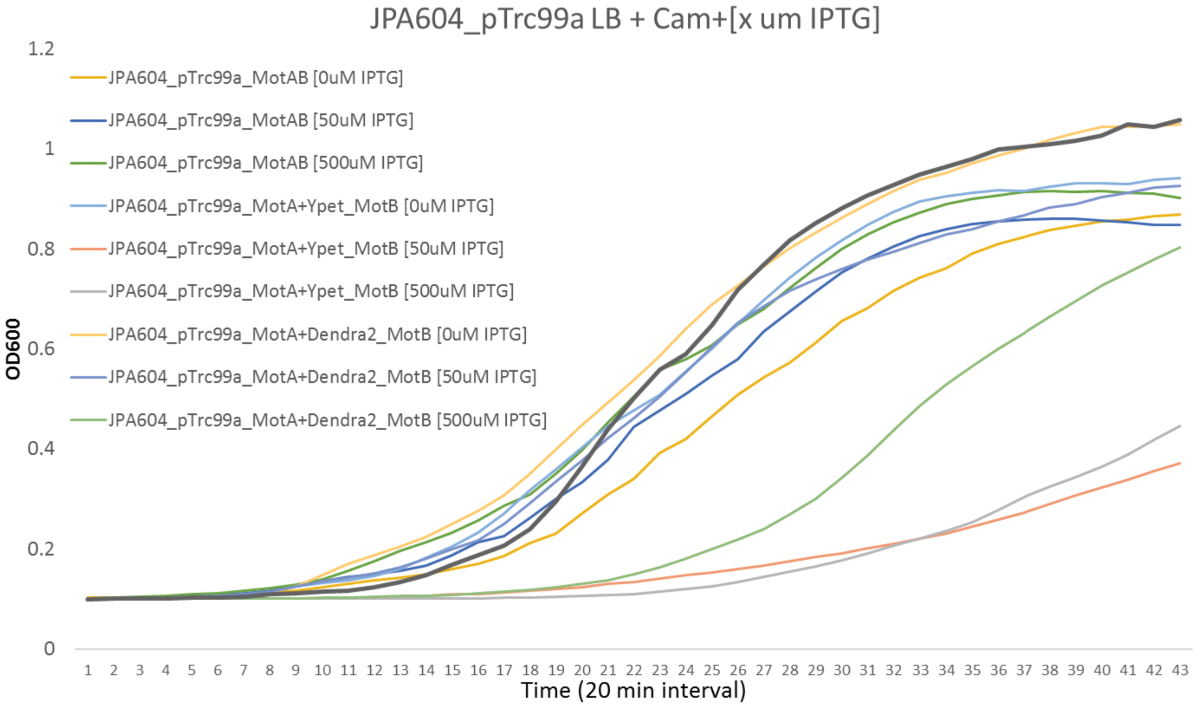
Appendix

A. OD600 growth curves

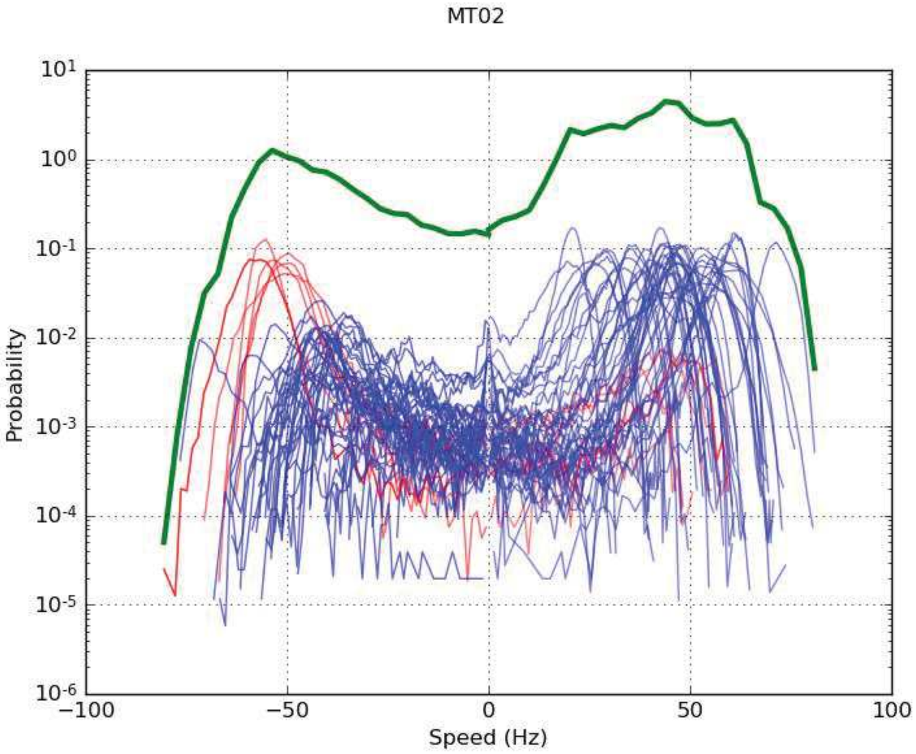


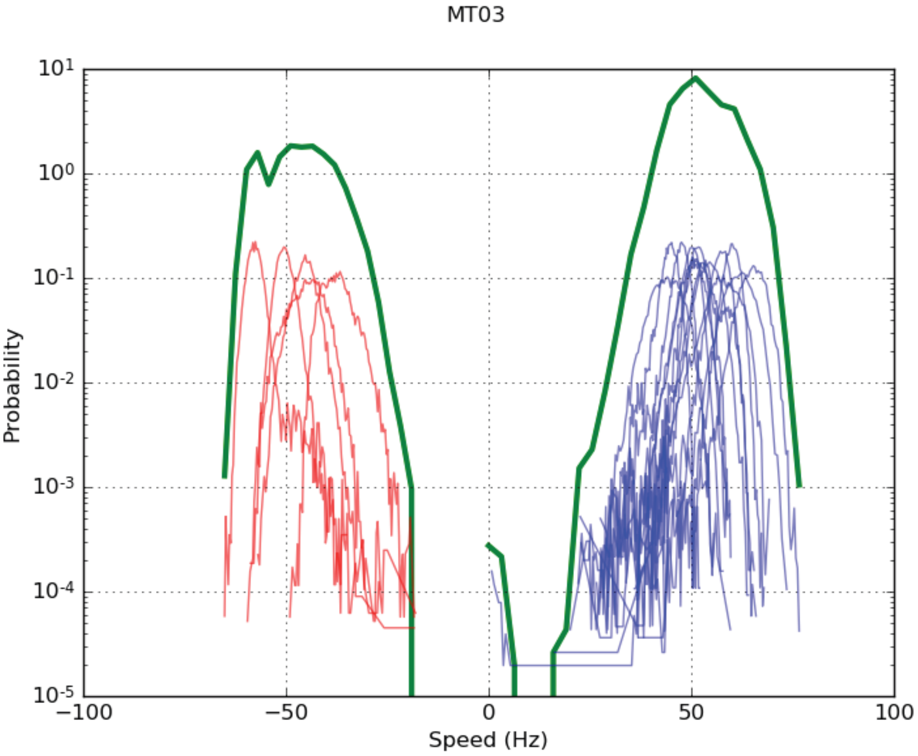


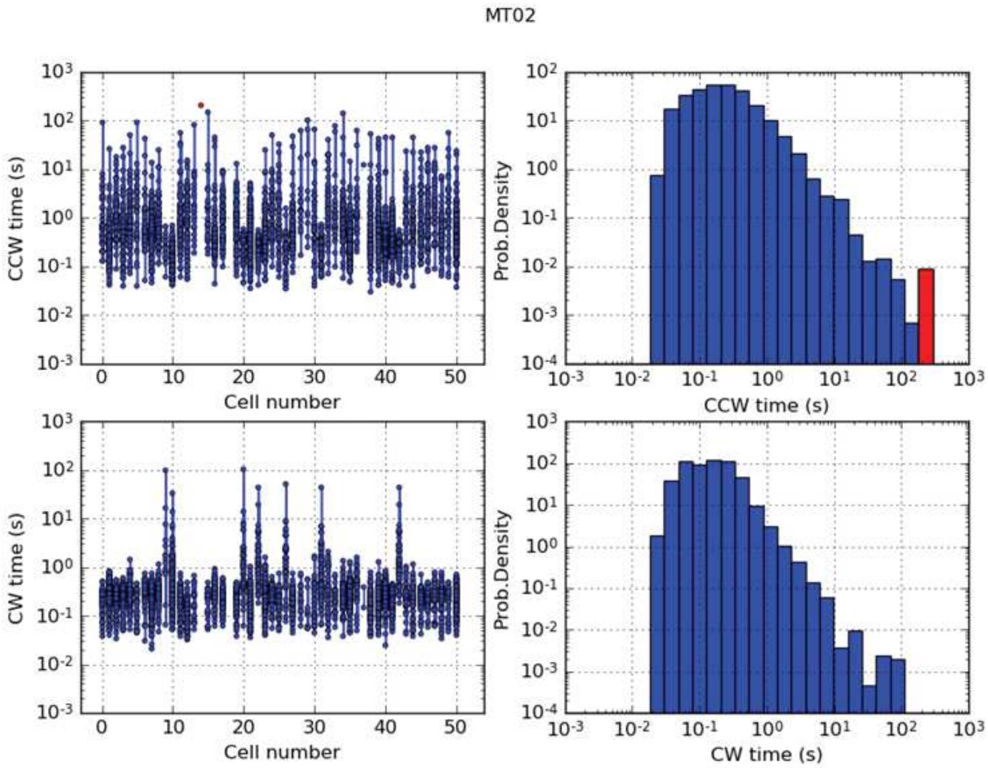

JPA604_pMMB226 in LB +Cam+[x um IPTG]
JPA604_pMMB604_MotAB [0uM IPTG]
JPA604_pMMB604_MotAB [50uM IPTG]
JPA604_pMMB604_MotAB [500uM IPTG]
JPA604_pMMB604_MotA+Ypet_MotB [0uM IPTG]
JPA604_pMMB604_MotA+Ypet_MotB [50uM IPTG]
JPA604_pMMB604_MotA+Ypet_MotB [500uM IPTG]
JPA604_pMMB604_MotA+Dendra2_MotB [0uM IPTG]
JPA604_pMMB604_MotA+Dendra2_MotB [50uM IPTG]
JPA604_pMMB604_MotA+Dendra2_MotB [500uM IPTG]
OD600
1.2
1
0.8
0.6
0.4
0.2
0
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43
Time (20 min interval)

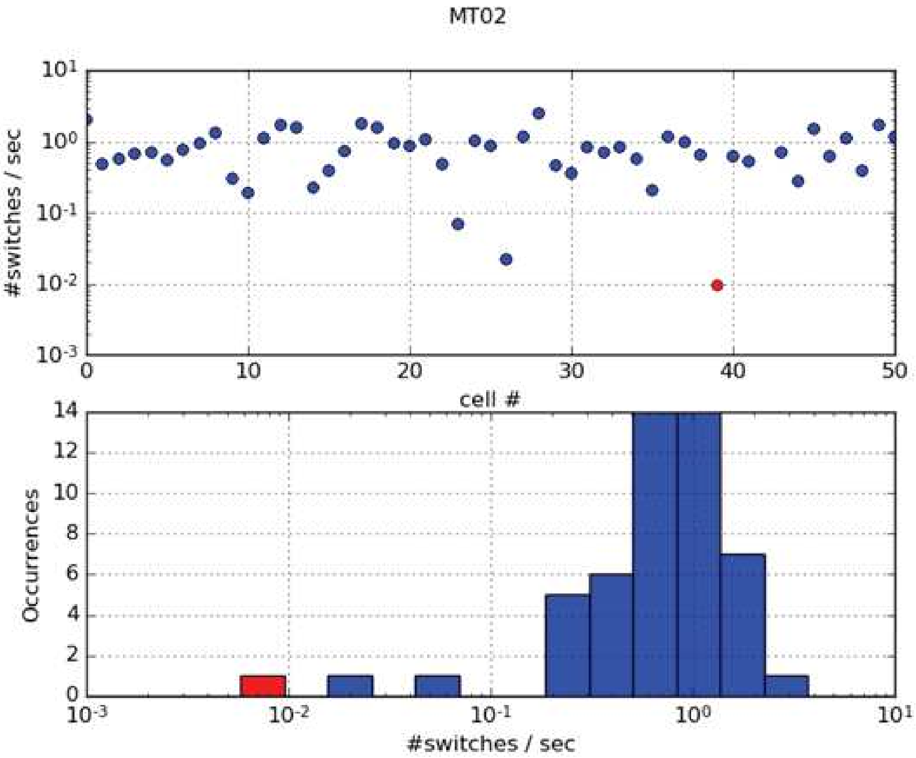

JPA604_pTrc99a LB + Cam+[x um IPTG]
JPA604_pTrc99a_MotAB [0uM IPTG]
JPA604_pTrc99a_MotAB [50uM IPTG]
JPA604_pTrc99a_MotAB [500uM IPTG]
JPA604_pTrc99a_MotA+Ypet_MotB [0uM IPTG]
JPA604_pTrc99a_MotA+Ypet_MotB [50uM IPTG]
JPA604_pTrc99a_MotA+Ypet_MotB [500uM IPTG]
JPA604_pTrc99a_MotA+Dendra2_MotB [0uM IPTG]
JPA604_pTrc99a_MotA+Dendra2_MotB [50uM IPTG]
JPA604_pTrc99a_MotA+Dendra2_MotB [500uM IPTG]
OD600
1.2
1
0.8
0.6
0.4
0.2
0
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43
Time (20 min interval)

B. WT BFM rotation (MT02 and MT03) with a 1.1 µm bead, see Table 2.1 for the strains information.

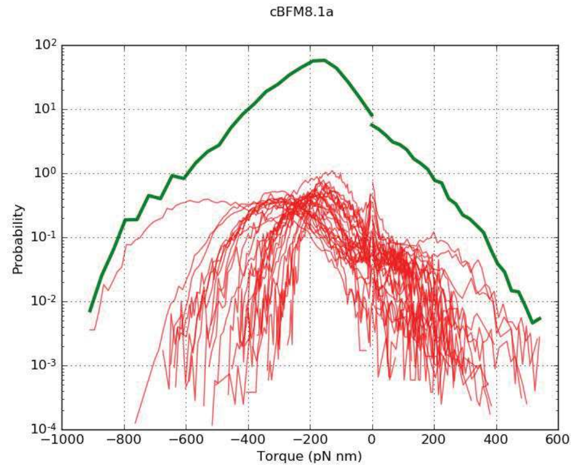


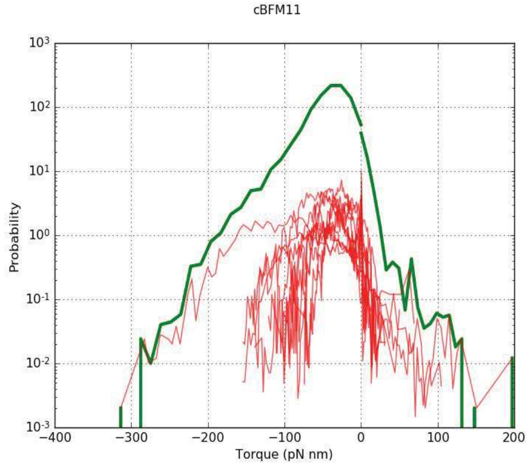


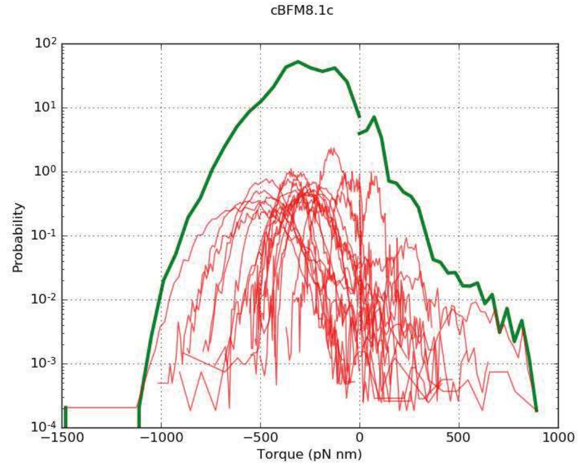

MT02
CCW time (s)
Cell number
Prob.Density
CCW time (s)
CW time (s)
Cell number
Prob.Density
CW time (s)
10^{3}
10^{2}
10^{1}
10^{0}
10^{-1}
10^{-2}
10^{-3}
10^{-4}
0
10
20
30
40
50

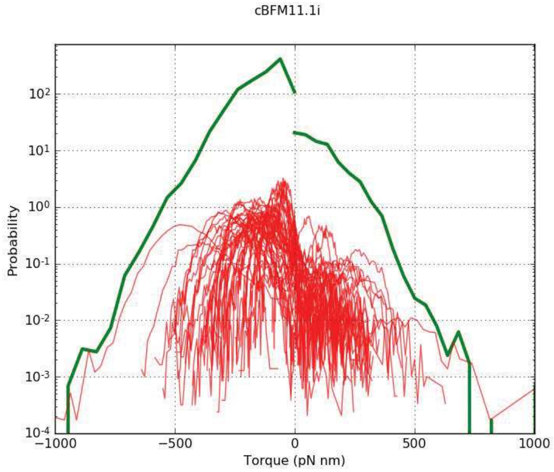

MT02
#switches / sec
cell #
Occurrences
#switches / sec
10^{1}
10^{0}
10^{-1}
10^{-2}
10^{-3}
0
10
20
30
40
50
14
12
10
8
6
4
2
0

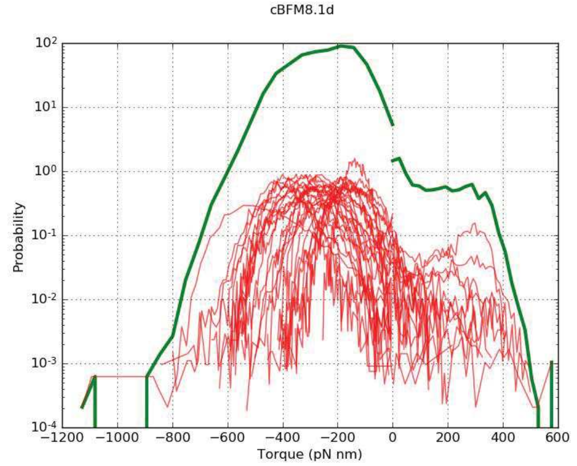
C. cBFM-P torque and speed histograms

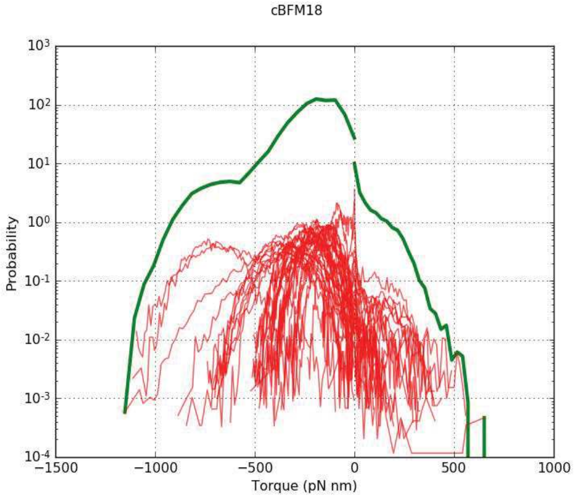


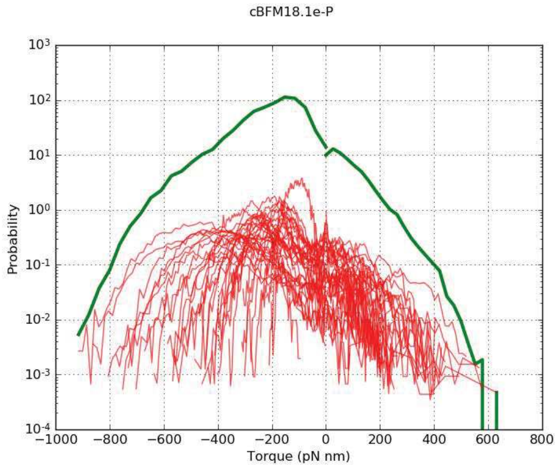


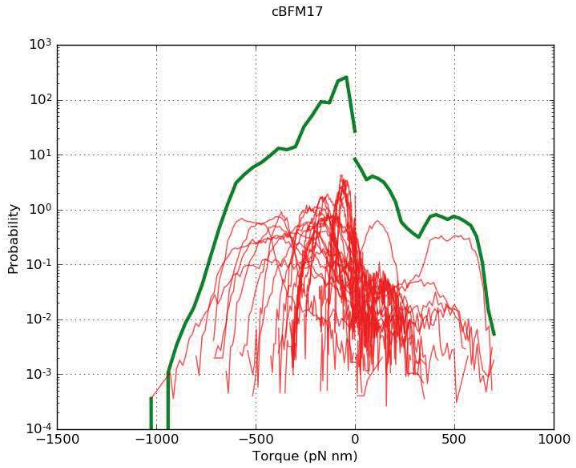


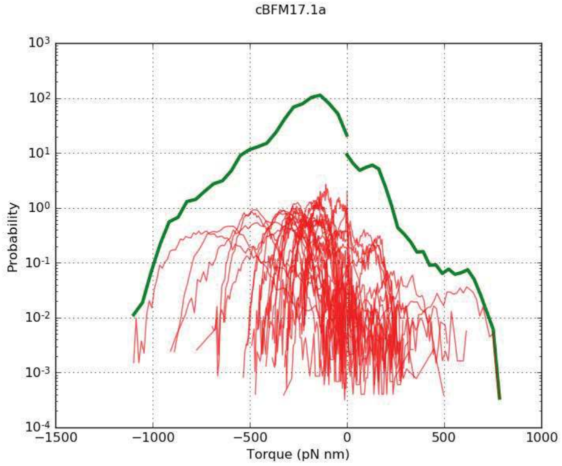


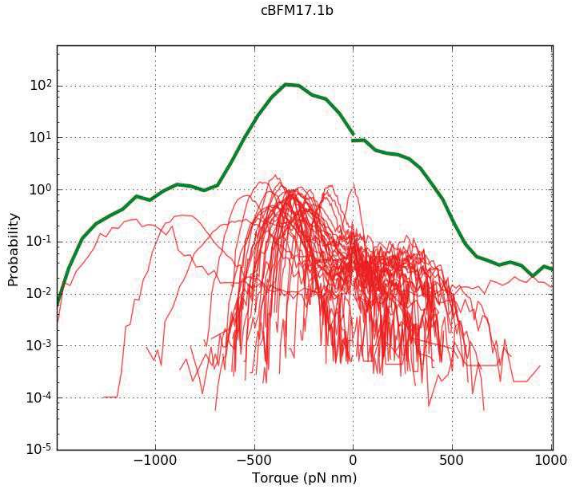


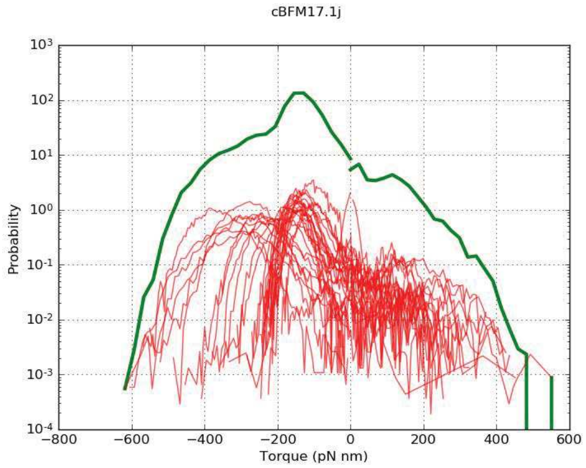


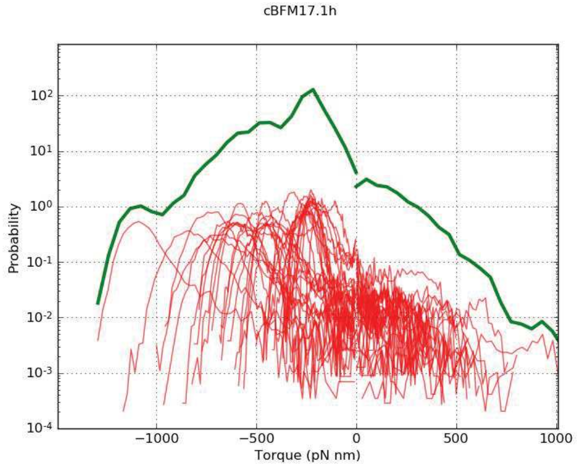

cBFM18.1e-P

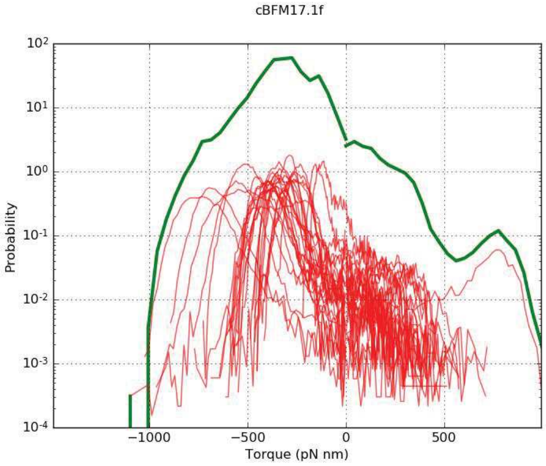

cBFM17

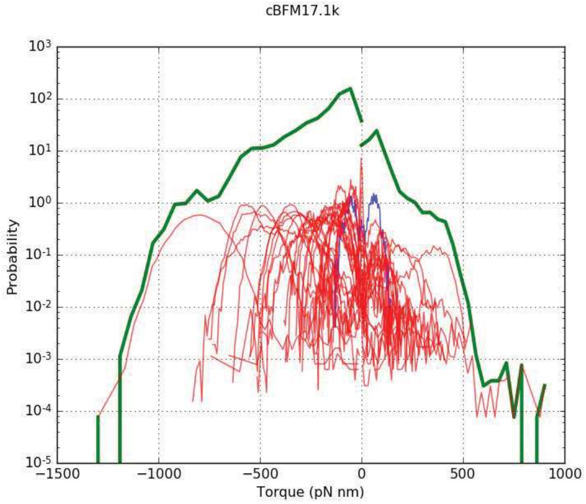

cBFM17.1a

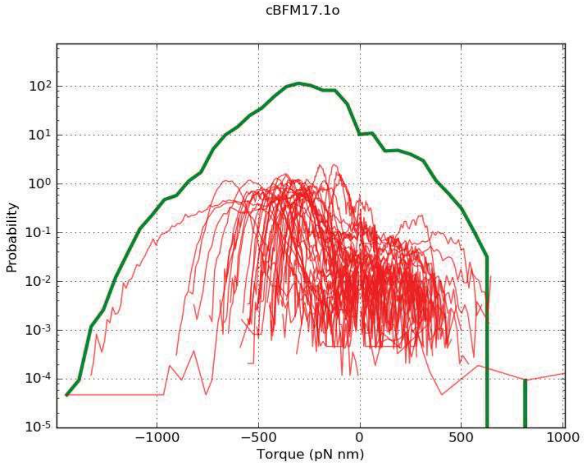

cBFM17.1b


cBFM17.1j


cBFM17.1h

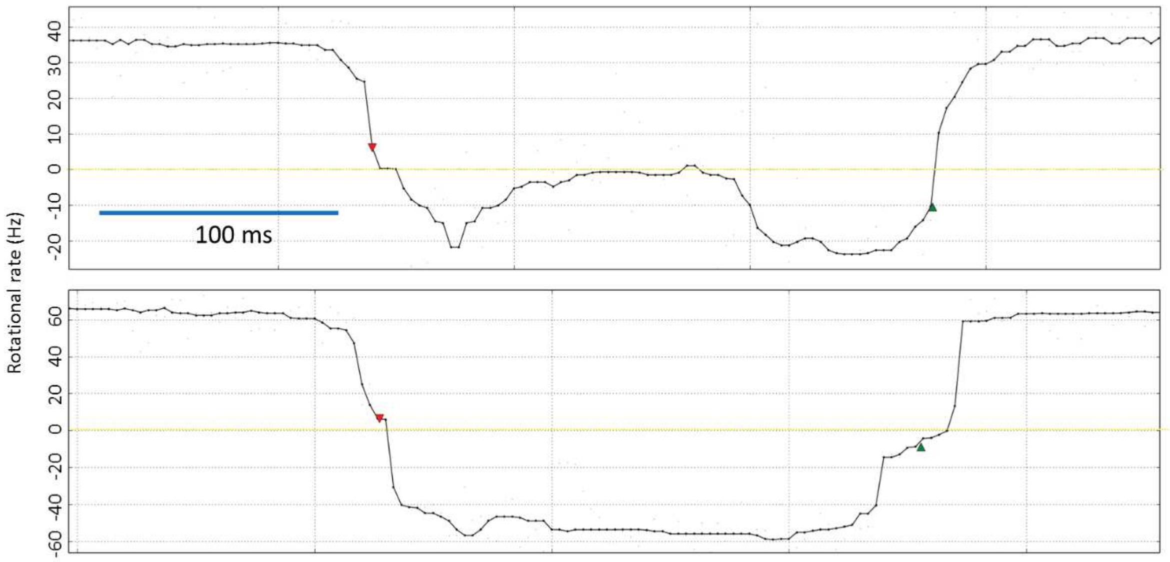

cBFM17.1f


cBFM17.1k

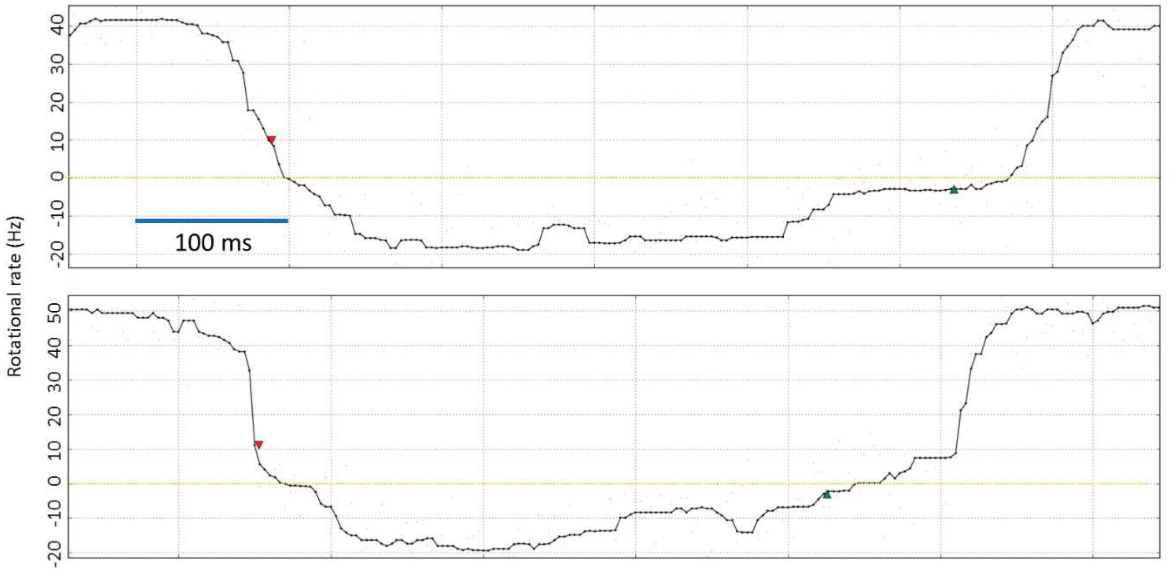

cBFM17.1o

D. Switching events of the BFMs often slowed down around 0 Hz (1.1 µm bead assay), in relation to the figure 3.14. Switching events from the two independent motors (top and bottom) are shown per each motor.

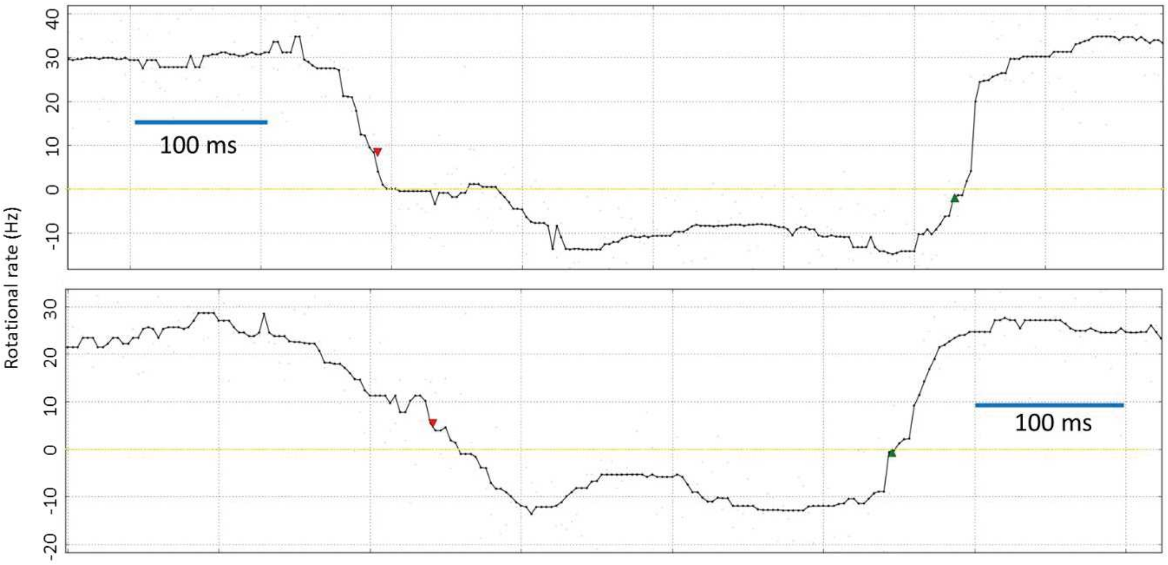
WT



YPet



eGFP



The bacterial flagellar motor (BFM) is a macromolecular complex which allows bacteria to swim in liquid media. Located at the base of the flagellum, this remarkably small (~45nm) yet powerful rotary motor rotates each flagellum up to rotation speeds of ~1000 Hz in both counterclockwise (CCW) and clockwise (CW) direction. The motor rotation is generated at the interface between the two key components of the motor: the stator protein complexes (each composed of 4 MotA and 2 MotB proteins) and the C- ring protein complex at the base of the rotor. The stator complexes are structurally and functionally discernible modules of the motor, and their dynamical association and dissociation around the rotor controls the torque generation. Previously, dynamic nature of the stators has been demonstrated by the fluorescence detection of the stators fused to a fluorescent protein (FP). When a FP is fused to MotB, the motor is functional, but a reduced motility of the cells has been observed. The precise reasons for such reduction in motility is yet to be determined.

The first project aims to investigate how the FP tag on the stator protein modifies the torque generation and switching of the motor. This is important because the fluorescent protein tag lies at the interface between stator and rotor, where torque and switching are produced. Three different FPs (eGFP, YPet, Dendra2) were fused to MotB. Interestingly, despite the high similarity of these three FPs' structures, our analysis revealed that the three fusion stators generate different torques by single stator. Furthermore, the motors driven by the fusion stators showed significantly impaired switching abilities. When switching direction of the rotation, the absolute value of the speed of WT motors does not change, whereas this symmetry of speed upon switching is not observed in the fusion stator motors and switching can be accompanied with a significant (~30%) decrease in absolute speed. Both the impaired torque generation and the switching ability were improved by introducing a rigid linker between the stator and the FP tag. Taken together, this study provides a further insight into the dynamics of the stator and rotor interaction at its interface.

When the cells carrying the fluorescently labeled stators were observed in a custom made TIRF-fluorescence microscope with single molecule capability, the fluorescence signals were detected as concentrated clusters in the membrane as expected for the stator proteins around the motors, together with a population of stators diffusing in the membrane. The fluorescent clusters were also observed at the center of rotating cells tethered to the glass slide by a single flagellum, confirming that the fluorescent spots observed were attributable to the functioning motors.

In a second project developed in Bertus Beaumont lab at TU Delft, taking BFM as an experimental evolutionary model system, its modularity and evolvability have been explored to learn the molecular details of the evolution of molecular machines. When the stators of *E.coli* (K-12) have been replaced by a set of 21 homologue foreign stators, some of the foreign stators were immediately compatible with the E.coli motor, while some of the non-compatible stators were positively modified by a chemotaxis evolution experiment. More than half of those evolved motors accumulated beneficial mutations in the functional domains of their foreign stator genes. Motilities of the evolved motors were investigated and compared at the level of population (chemotaxis), single cell (swimming) and single motor. This three-levels of functional investigation enabled detailed functional characterizations of the evolved or foreign motors. Especially, the single motor level assays revealed that those beneficial mutations improved the torque generation and/or the switching ability. The detailed genotype and phenotype investigations of the evolutionary modified BFM may bring an insight into how molecular machines such as BFM have evolved as well as the functional effects of the beneficial mutations that facilitate functional integration.