Structural Characterization of the Flagellar Stator Motor Protein motA and relatives

Name: Jiahe Zhang
Student ID: 287351052
Supervisors: Caroline Puente-Lelievre
Nicholas Matzke
October 2023
Total word counts: 3122

Abstract

The bacterial flagellum is a motility organelle that serves as a model system for studying complex evolutionary adaptations. The MotA protein is an essential component of the stator motor complex of the flagellum because it creates the ion flow necessary to generate the torque for the filament to rotate. MotA can use different power sources (Na+ or H+) to propel bacteria through various environments. The evolutionary history of this protein can be inferred by comparing sequence and structural differences. The recent breakthrough of structure prediction AI, AlphaFold, has successfully addressed the challenge of complex protein structure prediction. AlphaFold's accurate ability to predict protein structures offers a novel perspective for gaining deeper insights into the MotA protein family. The aim of this study is to determine whether there are structural differences between the clades that include proteins that use H+ and/or Na+, and to document the structure of unknown homologous domains. We generate protein structure predictions of the MotA protein family using AlphaFold to systematically compare protein structure across clades. Our results lead to a more consistent and accurate classification of the flagellar proteins that identifies and documents the high protein diversity in this group.

Key words: Bacterial flagellum, MotA protein, Stator motor complex, Power sources, AlphaFold, Phylogenetic tree

Contents

1 Background

1.1 Overview of bacterial flagella

The bacterial flagellum is a slender, spiral appendage on the surface of bacterial cells. This long, filamentous structure can vary across different bacteria(Haiko & Westerlund-Wikström, 2013). Flagellated bacteria can have multiple flagella dispersed across their cell surface, others may have only one positioned at a specific location on the cell(Schuhmacher et al., 2015). These flagella provide the driving force for bacterial movement, propelling bacteria in search of nutrients. They also play a pivotal role as sensory organs, allowing bacteria to interact with and respond to external stimuli(Wang et al., 2005).

The bacterial flagellum is an intricate supramolecular complex containing more than 30 that self-assemble into three major parts: the basal body, the hook, and the external flagellar filament(Zhao et al., 2014). The basal body is a complex structure that spans the bacterial cell membrane and cell wall, supplying the energy for flagellar rotation and generating torque. The hook connects the basal body and the flagellar filament, creating an angle that allows the flagellum to rotate in different directions and subsequently change its course of motion. The flagellar filament, the outermost part, typically assumes a helical shape, with its length, form, and helical density varying among different bacterial species(Vonderviszt & Namba, 2013).

1.2 Rotation mechanism of the bacterial flagellum

The rotation of bacterial flagellum is a collaborative effort between the helical filament and the motor at the base of the flagellum. This motor is powered by protons or other ions that pass through the cytoplasmic membrane(Larsen et al., 1974). The rotation of the motor results in the filament propelling or pulling the bacterium through a liquid environment.

The direction of rotation determines the type of movement of the cell. When the flagellum rotates counterclockwise, it forms a bundle, propelling the cell in a straight line. This is known as "running." When the flagellum rotates clockwise, it disrupts the cellular bundle, causing the cell to tumble and change direction. These chemical concentration-mediated, flagellum-driven properties become the bacterial chemotaxis. It is one of the most thoroughly studied bacterial properties, bacterial flagella are closely related to this movement(Colin et al., 2021).

When protons flow through the flagellar motor, they first pass through specific protein channels, such as Motility protein A(MotA) and Motility protein B (MotB). These channels are located on the bacterial cell membrane and are key for the bacterium to harness the difference in proton concentration inside and outside. As protons pass through these channels, proteins within the motor undergo conformational changes, thereby generating torque(K. I. Ito et al., 2021). This torque is transmitted to the base of the flagellum, driving its rotation. The entire process is highly efficient, enabling the bacterium to move purposefully in liquid environment, responding to external chemical signals.

1.3 Stator motor complex motA, function and structure

Cryo-electron microscopy observations by Deme et al. (2020) indicate that the structure of the bacterial flagellar motor possesses a stoichiometry of 5:2. This structure involves five MotA monomers

surrounding two MotB monomers forming a heteroheptameric complex. This rotary motor derives its power from smaller units, specifically the MotAB stator-rotor complex, which spans across the entire inner cell membrane and the peptidoglycan layer.

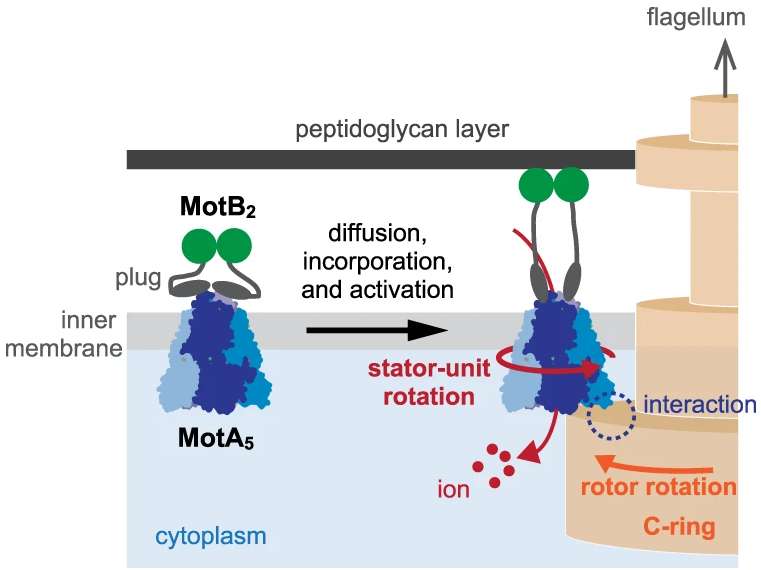

Figure 1: Ions flow through ion channels in the stator, creating torque(K. I. Ito et al., 2021).

MotA is a crucial transmembrane protein within the bacterial flagellar motor, serving as a linchpin for its effective functionality. Forming a pentameric wheel, it intricately envelops the MotB in the bacteria's cytoplasmic membrane. Structurally, MotA comprises four transmembrane helices (TM1-TM4) and cytoplasmic domain(Dean et al., 1984). This domain extend into the cytoplasm, engaging actively with other vital flagellar motor components, including FliG, FliM, and FliN. MotA's functions are multifaceted, encompassing the transmission of rotational force from proton flow to the rotor and the meticulous regulation of both the direction and speed of flagellar rotation(Rieu et al., 2022). Moreover, it can switch between two distinctive conformations: a locked state, providing a tight connection with MotB to prevent proton leakage, and a power stroke state, releasing MotB and advancing the rotor by a single step. This conformational flexibility of MotA is initiated and influenced by the proton movement from the periplasm to a binding site on MotB, inducing a shift in TM3 relative to TM2 and affecting the interaction with FliG on the rotor, ultimately determining the direction of rotation(K. I. Ito et al., 2021).

1.4 MotAB stator complex homologs

1.4.1 Flagellar proteins

The bacterial flagellar motor system is a highly complex and refined apparatus for movement, incorporating multiple complexes driven by either protons or sodium ions. In Vibrio alginolyticus, the proteins PomA and PomB collectively establish a sodium ion channel complex crucial for sustaining and facilitating the movement of bacterial flagella(Steiner et al., 2020). In the extremely alkalophilic Bacillus, the cell interior houses a sodium-driven stator complex composed

of MotP and MotS. Interestingly, in Bacillus subtilis, there is not only the presence of a proton-powered MotAB stator complex but also a concurrent existence of a sodium-driven MotPS stator complex(M. Ito et al., 2004). This unique scenario unveils potential sophisticated and diversified mechanisms of energy conversion and transmission within these organisms. Further studies indicate that Pseudomonas aeruginosa can precisely regulate its movement functions by finely tuning the activities of two sets of proton-dependent stators—MotAB and MotCD. During this regulatory process, MotAB primarily acts to prevent cluster movement, while MotCD plays a vital role in promoting it(Toutain et al., 2007).

1.5 Homolog Proteins

In biology, homology meticulously denotes a similitude in structure and function, attributed to genes or proteins that stem from a unified origin(Wagner, 1989). Amidst the evolutionary trajectory, these genes or proteins, while preserving similarities, undergo mutations, thereby engendering novel functionalities(Camps et al., 2007).

Homologous structures, wherein two entities represent corresponding parts built per an identical biological blueprint, have consistently been a linchpin in reconstructing phylogenetic histories(Wagner, 1989). Distinctions amongst bacterial flagella, irrespective of being quantitative or positional, can all be found on the bacterial surface, thereby facilitating an assessment of their homology.

Sequence homology is emblematic of a similarity in the arrangement of biological molecules, for instance, the nucleotide sequences in DNA or amino acid sequences in proteins. Conventionally, such sequential similarities are construed to emanate from a shared evolutionary genesis. Engaging in sequence alignment to study this homology facilitates speculative inferences regarding their functional or structural correlative aspects, and by extension, their evolutionary kinships(Pearson, 2013).

Conversely, structural homology is predominantly concerned with the three-dimensional conformational similarities of molecules, with a particular focus on protein tertiary structures(Russell et al., 1997). Notably, even amidst scenarios where two proteins exhibit negligible or nil sequence similarity, their three-dimensional conformations may bear striking resemblances (Balaji & Srinivasan, 2007). In a general context, analogous protein structures infer a likelihood of analogous functionalities or properties. Undertaking structural homology comparisons through protein conformations can glean insights into the evolutionary trajectory of protein functionalities.

1.6 Proteins structure to Homologs

The relationship between the structure and function of proteins is profoundly intertwined, the structure of a protein directly influences the architecture of the entire component as well as its biological function(Alberts et al., 2002). With the advent of X-ray crystallography and the emergence of various protein structure insight tools, a new understanding of the relationship between protein structure and function has been attained(Zheng et al., 2015). The study of homology has been extended from sequence to structural level.

Illergård et al. (2009) found that structure evolves three to ten times slower than sequence, indicating that structure is more conserved than sequence. Throughout the lengthy evolutionary process, under various selection pressures, the three-dimensional structure of proteins maintains

a more stable conformation. This phenomenon provides evidence for our understanding of the homology of proteins through structure, offering a better direction for evolutionary research.

Moreover, for proteins with low similarity at the sequence level but high similarity in structure and function, the evolutionary relationship of these proteins cannot be determined. Even if the amino acid sequence changes over the long evolutionary process, its core functional area and three-dimensional structure are still retained to meet the physiological needs of the organism(Rivoire et al., 2016). It is this highly conserved characteristic of structure that makes the study of protein structure homology crucial.

1.6.1 Non-flagellar proteins

Although ExbB and TolQ exhibit some differences in sequence similarity with MotA, they collectively utilize proton-driving forces to execute distinct key physiological functions(Cascales et al., 2001; Saier, 2000). ExbBD primarily concentrates on the active transport of iron and vitamin B12, facilitating the uptake of these two nutrients through the bacterial outer membrane(Braun et al., 2023). This process is crucial for bacteria to absorb and utilize essential trace elements and nutrients, thereby sustaining their physiological activities and metabolic requirements. Concurrently, TolQR primarily functions to maintain the integrity and stability of the bacterial outer membrane, playing a role in various significant biological processes and cellular signaling pathways(Cascales et al., 2001). Additionally, it's noteworthy that the AlgR protein is closely associated with the gliding mobility of mucous bacteria. AlgR is also a component of the proton channels in mucous bacteria, influencing their gliding and migration in physical environments and playing a key role in their adaptation and response to environmental changes(Jermy, 2011). The channel proteins within these complexes exhibit substantial structural similarity, indicating a sequence-structure resemblance whether in bacterial flagellar proteins or non-flagellar proteins, showcasing the potential for further research.

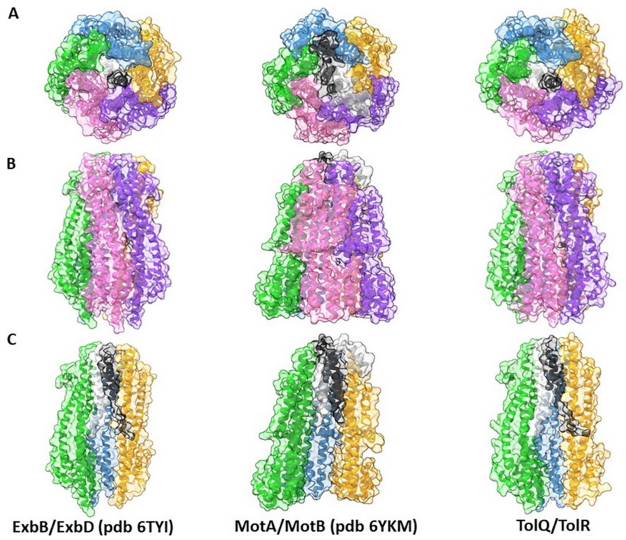

Figure 2: The ExbB/MotA/TolQ subunits are purple, pink, green, blue, and orange. The pentameric subunits form a central pore in which a dimer of ExbD/MotB/TolR subunits is shown in black and white(Braun et al., 2023).

1.7 Alphafold for structure prediction

AlphaFold is a revolutionary protein structure prediction tool developed by DeepMind. The software's core technology is based on deep learning and artificial intelligence algorithms, enabling high-precision predictions of three-dimensional protein structures(Jumper et al., 2021a). By analyzing amino acid sequences, AlphaFold employs advanced techniques, including convolutional neural networks and graph neural networks, to precisely model and predict the spatial conformation of proteins. In the Critical Assessment of Structure Prediction (CASP) competition, this approach demonstrated unprecedented predictive accuracy, marking a new era for computational and structural biology(Varadi et al., 2022). The advent of AlphaFold has significantly advanced protein-related research and is invaluable for making inferences about the phylogenetic processes of homologous proteins. With its deep learning algorithms, AlphaFold can rapidly and accurately decipher protein structures, providing researchers with a powerful tool for deeply exploring the evolutionary history of proteins throughout the process of biological evolution.

2 Research question and general aims

2.1 Structure and function of stator motor proteins and their homologs

The research focuses on the structural and functional diversity of the stator motor protein motA and its homologs. We aim to investigate the structural characteristics of these proteins, the functions of homologous and non-homologous structural domains, and how they execute their roles in various biological environments. By comparing the three-dimensional features, we intend to understand their evolutionary history, focusing on the conservation and changes in these structures. This in-depth comparative analysis will elucidate the evolutionary mechanisms responsible for the observed structural and functional alterations, enriching our comprehension of these structurally similar and functionally identical phenomena in proteins.

Although we have a deep understanding of the structure of the flagellum, knowledge about its origin and evolution remains relatively limited.

2.2 General aims

The general aim is to systematically study the homology of proteins constituting the flagellum to better understand the evolution of the bacterial flagellar system. We will focus on comparing the structural and functional attributes of flagellar proteins across various bacterial species to uncover evolutionary patterns and mechanisms driving the flagellar apparatus's diversification. This endeavor will provide a deeper understanding of bacterial flagella's evolutionary intricacies, shedding light on the molecular components underpinning bacterial motility.

3 Methodology

3.1 Sampling

A minimum of 80 protein sequences belonging to the flagellar protein clade (Figure 3) will be selected for structural analysis we strategically selected branches within the phylogenetic tree where MotA proteins are predominant.

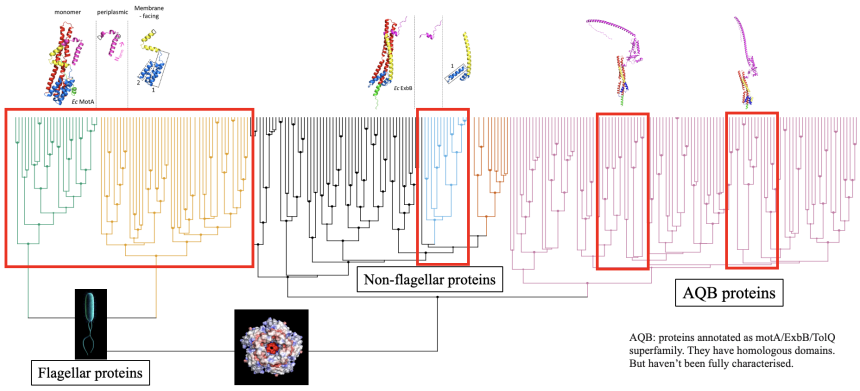

Figure 3: Phylogenetic tree of representative bacterial MotA proteins and their homologs(Puente-Lelievre et al, unpublished).

For each selected sample, we will retrieve the corresponding sequences through their GenBank (www.ncbi.nlm.nih.gov/genbank) accession numbers. Upon retrieval, the sequences will be formatted and prepared for subsequent structural prediction analysis.

3.2 Protein structure prediction

Protein structure predictions will be retrieved from the AlphaFold database(Jumper et al., 2021b). However, given AlphaFold's ongoing model updates, some of the protein structures in the database may be based on earlier model predictions. In order to ensure prediction accuracy and research rigour, we chose to repredict all FASTA sequences, including those proteins already in the AlphaFold database, to ensure that all research data are based on AlphaFold's latest model.

Structure predictions will be generated using AlphaFold (AlphaFold2DB/2023-04) on the New Zealand e-Science Infrastructure (NeSI)(https://www.nesi.org.nz).

3.3 Structural characterisation

Protein structures will be visualized and analysed using PyMOL (The PyMOL Molecular Graphics System, Version 1.2r3pre, Schro¨dinger, LLC). Diagnostic and highly conserved structural features will be identified for each protein. For Alphafold predictions, only the regions of the protein with confidence scores above 90% will be characterized.

3.4 Deep homology search

Sequence-based homology searches will be run using HMMER (hmmer v3.4)(Finn et al., 2015). HMMER applies Hidden Markov Models to analyze protein sequences to pinpoint proteins similar to the sequences of the conserved domains. Through these homology searches protein members related to these conserved domains can be identified. Structure-based homology searches for the Alphafold predictions will be run using Foldseek(van Kempen et al., 2023). Foldseek aims to identify homologous proteins based on the structure similarity of the conserved domains.

3.5 Map structural and functional traits into the phylogeny

In order to infer the evolutionary history and phylogenetic relationships of the observed structural features of these proteins, they will be manually mapped onto a phylogenetic tree (figure 4). Firstly, we identify the systematic similarities between each obtained structural feature and different branches (clades) within the phylogenetic tree. For this purpose, we meticulously analyze the protein structural features of each branch, identifying the main and unique structural characteristics of each. These characteristics are then matched with the corresponding branches in the phylogenetic tree. By manually positioning the identified conserved structural domains at their respective branches on the phylogenetic tree, we ensured that these locations accurately represented the principal structural features of each branch.

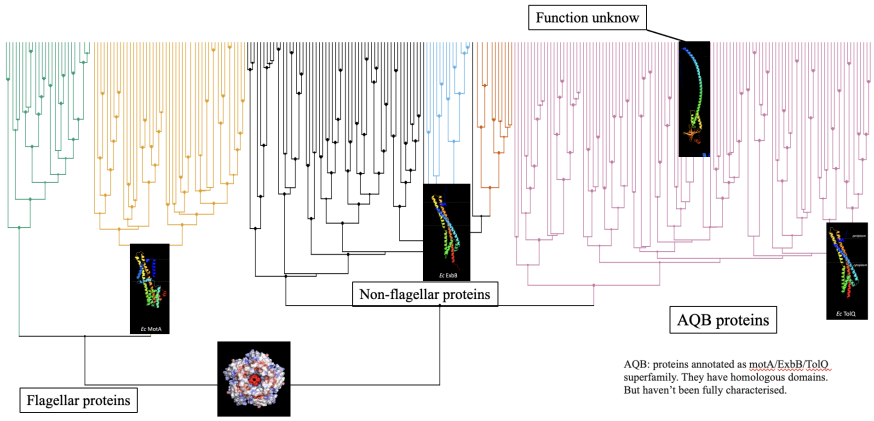

Figure 4: Phylogenetic tree with structural traits(Puente-Lelievre et al, unpublished).

4 Expected results

The expected outcomes of this project are:

-Accurate structural predictions and visualization analyses

-Comprehensive comparison and analysis of the three-dimensional structures of proton-driven and sodium-driven proteins.

-Identification of diagnostic structural traits and potential functions of MotA homologs

-Determine the structure, homology and function of additional domains of unknown function.

-Insights into the evolutionary adaptation and diversity of MotA proteins and their variants

5 Implications and significance

Through employing systematic comparison and construction based on protein structures, this study will advance our understanding of the MotA protein family. The refined classification features derived from this research will provide a tool for the scientific community to navigate the complexity of this diverse protein family with enhanced precision and clarity. This development is particularly significant as it facilitates a more nuanced comprehension of the diverse array of proteins within this family, fostering further research and discovery in Phylogenetic and taxonomic fields

Moreover, the unveiling of key evolutionary events within the MotA protein family constitutes a landmark understanding in evolutionary biology. These events, integral to the narrative of adaptive evolution, act as guideposts illuminating the evolutionary trajectory of these proteins. As such, this newfound knowledge significantly contributes to the broader understanding of adaptive evolution mechanisms and processes. It enables researchers and scientists to craft informed hypotheses and models that more accurately reflect the dynamic and complex nature of protein evolution, serving as a robust framework for future studies.

Furthermore, the insights gained into the evolutionary history of the stator, a critical component of the flagellum, will provide a coherent picture of the evolution of the flagellum. This comprehensive perspective is instrumental in piecing together the intricate puzzle of the flagellum's evolutionary development and functionality. The understanding furnished through this study, therefore, not only fills existing knowledge gaps but also serves as a catalyst for future explorations into the mysteries of bacterial flagella, their evolution, and their multifaceted roles and applications in biology and medicine. Prediction of protein structure using Alpha and subsequent alignment provides a high-resolution structure of the protein. Bacterial flagella serve as virulence factors, and research based on these flagellins is driving breakthroughs in fundamental questions in life sciences and is expected to change research and accelerate drug discovery(Nussinov et al., 2022).

6 Limitation

This study is a systematic comparison of protein structures predicted using AlphaFold rather than experimentally obtained and verified. Although the accuracy of AlphaFold has been widely recognized, predicted structures are not 100% accurate(Jumper et al., 2021b). Discrepancies exist between predicted and actual protein structures, and these discrepancies may lead to potential misjudgments in our findings.

Homology searches primarily rely on existing databases and analysis tools. Sequence-based homology search tools that use Hidden Markov Models (HMMs) such as HMMER may not be exhaustive enough due to the possibility of unknown or unincorporated protein sequences and structures within the databases, thereby limiting the completeness of our results.

Although sampling for this study aims to capture most of the bacterial diversity, it does not comprise all of the diversity and complexity of bacterial species. Therefore, the results of this study may not be entirely generalizable to all types of bacteria. Unique protein structures and functions may exist in other species of bacteria that have not been included in this study .

References

Alberts, B., Johnson, A., & Lewis, J. (2002). Protein Function - Molecular Biology of the Cell - NCBI Bookshelf. Retrieved October 15, 2023, from https://www-ncbi-nlm-nih-gov.ezproxy.auckland.ac.nz/books/NBK26911/

Balaji, S., & Srinivasan, N. (2007). Comparison of sequence-based and structure-based phylogenetic trees of homologous proteins: Inferences on protein evolution. *Journal of Biosciences*, *32*(1), 83–96. https://doi.org/10.1007/s12038-007-0008-1

Braun, V., Ratliff, A. C., Celia, H., & Buchanan, S. K. (2023). Energization of Outer Membrane Transport by the ExbB ExbD Molecular Motor [Publisher: American Society for Microbiology]. *Journal of Bacteriology*, *205*(6), e00035–23. https://doi.org/10.1128/jb.00035-23

Camps, M., Herman, A., Loh, E., & Loeb, L. A. (2007). Genetic Constraints on Protein Evolution. *Critical reviews in biochemistry and molecular biology*, *42*(5), 10.1080/10409230701597642. https://doi.org/10.1080/10409230701597642

Cascales, E., Lloubès, R., & Sturgis, J. N. (2001). The TolQ-TolR proteins energize TolA and share homologies with the flagellar motor proteins MotA-MotB. *Molecular Microbiology*, *42*(3), 795–807. https://doi.org/10.1046/j.1365-2958.2001.02673.x

Colin, R., Ni, B., Laganenka, L., & Sourjik, V. (2021). Multiple functions of flagellar motility and chemotaxis in bacterial physiology. *FEMS Microbiology Reviews*, *45*(6), fuab038. https://doi.org/10.1093/femsre/fuab038

Dean, G. E., Macnab, R. M., Stader, J., Matsumura, P., & Burks, C. (1984). Gene sequence and predicted amino acid sequence of the motA protein, a membrane-associated protein required for flagellar rotation in Escherichia coli [Publisher: American Society for Microbiology]. *Journal of Bacteriology*, *159*(3), 991–999. https://doi.org/10.1128/jb.159.3.991-999.1984

Deme, J. C., Johnson, S., Vickery, O., Aron, A., Monkhouse, H., Griffiths, T., James, R. H., Berks, B. C., Coulton, J. W., Stansfeld, P. J., & Lea, S. M. (2020). Structures of the stator complex that drives rotation of the bacterial flagellum [Number: 12 Publisher: Nature Publishing Group]. *Nature Microbiology*, *5*(12), 1553–1564. https://doi.org/10.1038/s41564-020-0788-8

Finn, R. D., Clements, J., Arndt, W., Miller, B. L., Wheeler, T. J., Schreiber, F., Bateman, A., & Eddy, S. R. (2015). HMMER web server: 2015 update. *Nucleic Acids Research*, *43*(W1), W30–W38. https://doi.org/10.1093/nar/gkv397

Haiko, J., & Westerlund-Wikström, B. (2013). The Role of the Bacterial Flagellum in Adhesion and Virulence. *Biology*, *2*(4), 1242–1267. https://doi.org/10.3390/biology2041242

Illergård, K., Ardell, D. H., & Elofsson, A. (2009). Structure is three to ten times more conserved than sequence–a study of structural response in protein cores. *Proteins*, *77*(3), 499–508. https://doi.org/10.1002/prot.22458

Ito, K. I., Nakamura, S., & Toyabe, S. (2021). Cooperative stator assembly of bacterial flagellar motor mediated by rotation [Number: 1 Publisher: Nature Publishing Group]. *Nature Communications*, *12*(1), 3218. https://doi.org/10.1038/s41467-021-23516-y

Ito, M., Hicks, D. B., Henkin, T. M., Guffanti, A. A., Powers, B. D., Zvi, L., Uematsu, K., & 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 [_eprint: https://onlinelibrary.wiley.com/d 2958.2004.04173.x]. *Molecular Microbiology*, *53*(4), 1035–1049. https://doi.org/10.1111/j.1365-2958.2004.04173.x

Jermy, A. (2011). Motor helps gliders to gain traction [Number: 6 Publisher: Nature Publishing Group]. *Nature Reviews Microbiology*, *9*(6), 398–398. https://doi.org/10.1038/nrmicro2579

Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Žídek, A., Potapenko, A., Bridgland, A., Meyer, C., Kohl, S. A. A., Ballard, A. J., Cowie, A., Romera-Paredes, B., Nikolov, S., Jain, R., Adler, J., … Hassabis, D. (2021a). Highly accurate protein structure prediction with AlphaFold. *Nature*, *596*(7873), 583–589. https://doi.org/10.1038/s41586-021-03819-2

Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Žídek, A., Potapenko, A., Bridgland, A., Meyer, C., Kohl, S. A. A., Ballard, A. J., Cowie, A., Romera-Paredes, B., Nikolov, S., Jain, R., Adler, J., … Hassabis, D. (2021b). Highly accurate protein structure prediction with AlphaFold [Number: 7873 Publisher: Nature Publishing Group]. *Nature*, *596*(7873), 583–589. https://doi.org/10.1038/s41586-021-03819-2

Larsen, S. H., Adler, J., Gargus, J. J., & Hogg, R. W. (1974). Chemomechanical coupling without ATP: The source of energy for motility and chemotaxis in bacteria. *Proceedings of the National Academy of Sciences of the United States of America*, *71*(4), 1239–1243. https://doi.org/10.1073/pnas.71.4.1239

Nussinov, R., Zhang, M., Liu, Y., & Jang, H. (2022). AlphaFold, Artificial Intelligence (AI), and Allostery [Publisher: American Chemical Society]. *The Journal of Physical Chemistry B*, *126*(34), 6372–6383. https://doi.org/10.1021/acs.jpcb.2c04346

Pearson, W. R. (2013). An Introduction to Sequence Similarity ("Homology") Searching. *Current protocols in bioinformatics / editoral board, Andreas D. Baxevanis ... [et al.]*, *0 3*, 10.1002/0471250953.bi0301s42. https://doi.org/10.1002/0471250953.bi0301s42

Rieu, M., Krutyholowa, R., Taylor, N. M. I., & Berry, R. M. (2022). A new class of biological ion-driven rotary molecular motors with 5:2 symmetry. *Frontiers in Microbiology*, *13*, 948383. https://doi.org/10.3389/fmicb.2022.948383

Rivoire, O., Reynolds, K. A., & Ranganathan, R. (2016). Evolution-Based Functional Decomposition of Proteins. *PLoS Computational Biology*, *12*(6), e1004817. https://doi.org/10.1371/journal.pcbi.1004817

Russell, R. B., Saqi, M. A. S., Sayle, R. A., Bates, P. A., & Sternberg, M. J. E. (1997). Recognition of analogous and homologous protein folds: Analysis of sequence and structure conservation11Edited by F. E. Cohen. *Journal of Molecular Biology*, *269*(3), 423–439. https://doi.org/10.1006/jmbi.1997.1019

Saier, M. H. (2000). A functional-phylogenetic classification system for transmembrane solute transporters. *Microbiology and molecular biology reviews: MMBR*, *64*(2), 354–411. https://doi.org/10.1128/MMBR.64.2.354-411.2000

Schuhmacher, J. S., Thormann, K. M., & Bange, G. (2015). How bacteria maintain location and number of flagella? *FEMS Microbiology Reviews*, *39*(6), 812–822. https://doi.org/10.1093/femsre/fuv034

Steiner, T. J., Stovner, L. J., Jensen, R., Uluduz, D., Katsarava, Z., & Lifting The Burden: the Global Campaign against Headache. (2020). Migraine remains second among the world's causes of disability, and first among young women: Findings from GBD2019. *The Journal of Headache and Pain*, *21*(1), 137. https://doi.org/10.1186/s10194-020-01208-0

Toutain, C. M., Caizza, N. C., Zegans, M. E., & O'Toole, G. A. (2007). Roles for flagellar stators in biofilm formation by Pseudomonas aeruginosa. *Research in Microbiology*, *158*(5), 471–477. https://doi.org/10.1016/j.resmic.2007.04.001

van Kempen, M., Kim, S. S., Tumescheit, C., Mirdita, M., Lee, J., Gilchrist, C. L. M., Söding, J., & Steinegger, M. (2023). Fast and accurate protein structure search with Foldseek [Publisher: Nature Publishing Group]. *Nature Biotechnology*, 1–4. https://doi.org/10.1038/s41587-023-01773-0

Varadi, M., Anyango, S., Deshpande, M., Nair, S., Natassia, C., Yordanova, G., Yuan, D., Stroe, O., Wood, G., Laydon, A., Žídek, A., Green, T., Tunyasuvunakool, K., Petersen, S., Jumper, J., Clancy, E., Green, R., Vora, A., Lutfi, M., … Velankar, S. (2022). AlphaFold Protein Structure Database: Massively expanding the structural coverage of protein-sequence space with high-accuracy models. *Nucleic Acids Research*, *50*(D1), D439–D444. https://doi.org/10.1093/nar/gkab1061

Vonderviszt, F., & Namba, K. (2013). Structure, Function and Assembly of Flagellar Axial Proteins. In *Madame Curie Bioscience Database [Internet]*. Landes Bioscience. Retrieved October 12, 2023, from https://www.ncbi.nlm.nih.gov/books/NBK6250/

Wagner, G. P. (1989). The Biological Homology Concept [Publisher: Annual Reviews]. *Annual Review of Ecology and Systematics*, *20*, 51–69. Retrieved October 15, 2023, from https://www.jstor.org/stable/2097084

Wang, Q., Suzuki, A., Mariconda, S., Porwollik, S., & Harshey, R. M. (2005). Sensing wetness: A new role for the bacterial flagellum. *The EMBO Journal*, *24*(11), 2034–2042. https://doi.org/10.1038/sj.emboj.7600668

Zhao, X., Norris, S. J., & Liu, J. (2014). Molecular Architecture of the Bacterial Flagellar Motor in Cells [Publisher: American Chemical Society]. *Biochemistry*, *53*(27), 4323–4333. https://doi.org/10.1021/bi500059y

Zheng, H., Handing, K. B., Zimmerman, M. D., Shabalin, I. G., Almo, S. C., & Minor, W. (2015). X-ray crystallography over the past decade for novel drug discovery – where are we heading next? *Expert opinion on drug discovery*, *10*(9), 975–989. https://doi.org/10.1517/17460441.2015.1061991