Evolution of the Bacterial Flagellum

Cumulative evidence indicates that flagella developed as modular systems, with many components deriving from other systems

Tim Wong, Arezou Amidi, Alexandra Dodds, Sara Siddiqi, Jing Wang, Tracy Yep, Dorjee G. Tamang, and Milton H. Saier, Jr.

Proponents of the intelligent design (ID) explanation for how organisms developed claim that the bacterial flagellum (BF) is irreducibly complex. They argue that this structure is so complicated that it could not have emerged through random selection but had to be designed by an intelligent entity. One part of this claim is that each flagellar component is used solely for the purpose of making a flagellum that, in turn, is used only for motility. Further, each flagellar protein is assumed to have appeared independently of the other component proteins.

Here, we summarize evidence from hundreds of laboratories, including our own, showing that these assumptions are false. Instead of by design, BF developed as modular systems, with components deriving from many different sources. Each BF module evolved independently from various primordial systems, which, in most cases, had nothing to do with cell motility. Complexity within BF arose by domain and protein recruitment, by intragenic and extragenic duplication events, and by superimposition of various modules onto others. The net result was coevolution of many types of structurally and functionally distinct flagella in various bacterial species. Although these different flagella are all used for motility, they share only about half of their protein constituents.

Summary

- Bacterial flagella are modular structures consisting of (1) a basal body, (2) a filamentous propeller, (3) an interconnecting hook complex, (4) a rotary motor, (5) a secretion/assembly system, (6) a secretion energizing ATPase, and (7) various ancillary proteins.
- Each module probably evolved independently of the others from primordial systems having nothing to do with cell motility.
- Complexity arose by domain and protein recruitment as well as by intragenic and extragenic duplication events.
- Hundreds of structurally and functionally distinct flagella, present in various bacterial species, share only about half their protein constituents.

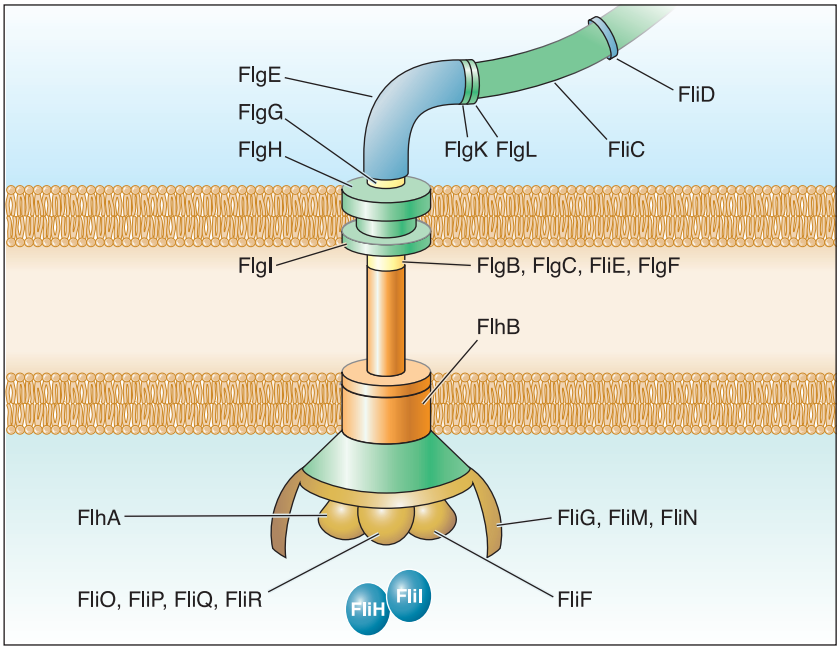
The Modular Construction of Flagella

Identifiable flagellar modules include (1) the basal body and its set of rings that anchor the flagellum to the bacterial cell envelope, (2) the hook complex that connects the basal body to the filament, (3) the filament that serves as the propeller for motility, (4) the motor for driving flagellar rotation, (5) the secretion system that exports flagellar subunits, and (6) the ATPase complex that energizes secretion (Fig. 1). Several ancillary proteins are also required for the proper synthesis and assembly of the BF.

We are learning how these modules appeared during evolutionary history. For instance, several functionally dissimilar proteins of the BF arose from a single common ancestral protein via gene duplications. Some additional constituents were constructed by fusing protein domains. Other constituents increased in size and complexity following intragenic duplications. Several modular BF proteins have homologues that serve functions unrelated to motility. Details describing the *Escherichia coli* flagellar protein components appear in a

Tim Wong, Arezou Amidi, Alexandra Dodds, Sara Siddiqi, Jing Wang, and Tracy Yep are undergraduate students; Dorjee G. Tamang is a Research Associate, and Milton H. Saier, Jr. is Professor of Molecular Biology, all in the Division of Biological Sciences, University of California at San Diego, La Jolla.

FIGURE 1



Physical structure of the *E. coli* flagellum showing the basal region embedded in the bacterial cell envelope with its rod and envelope-associated rings (bottom), the hook (middle) and the filament (top). The approximate positions of various flagellar proteins in the overall structure are shown (S.-I. Aizawa, FEMS Microbiol. Lett. **202:**157–164).

review published in 2006 by Mark Pallen and Nicholas Matzke of the University of Birmingham in England.

Variation in Flagellar Structures

Bacterial flagella are not uniform in construction. They include a range of related structures that differ from organism to organism, with different constituents that provide optimal motility in each particular organism. For example, while gram-negative bacteria have two membranes, gram-positive bacteria have only one membrane; consequently, the flagella of these two bacterial types differ in ways that reflect these two envelope types.

For instance, gram-negative flagella have basal bodies that include an M-ring that interacts with the inner membrane, a P-ring that interacts with the peptidoglycan cell wall, and an L-ring that interacts with the lipopolysaccharide-containing outer membrane (Fig. 2). Meanwhile, gram-positive bacteria, lacking an outer membrane, also lack the L-ring and its constituents. Having a structurally dissimilar cell wall, they also lack the typical gram-negative bacterial P-ring.

Many other constituents of the two flagella, including the proteins of the central rod and M-ring, are demonstrably homologous. Thus, evolving flagella apparently derived from a common ancestral system that was modified, allowing optimal embedment in various cell envelopes. Of the approximately 40 protein constituents of the enteric bacterial flagellum, only about 20 are common to all bacteria. What is essential for flagellar function in one bacterial species may not even be present in another!

Similarities Between Bacterial Flagella and Type III Protein Secretion Systems

Bacterial flagella and virulence-related type III protein secretion systems (T3SSs) derive from a common ancestral system and share at least nine homologous constituents. Apparently, both systems arose from a simpler primordial secretion system. These are just two of many distinct mechanisms for exporting proteins across gram-negative bacterial inner and outer membranes. Although most of these systems are thought to have evolved independently of each other, some of them share protein constituents, probably because homologous proteins were recruited for their construction.

T3SSs secrete proteins directly from the cytoplasm across the inner and outer membranes of the gram-negative bacterial envelope either into the external medium or directly into the cytoplasm of a host plant or animal cell with which the bacterium forms a symbiotic or pathogenic relationship. T3SSs consist of thin, rigid, hypodermic needle-like protein complexes anchored to the envelope by basal structures resembling flagellar basal bodies (Fig. 2). In some instances, both T3SSs and flagellar secretory systems export the same or similar proteins across two-membrane envelopes, showing that these systems overlap structurally and functionally.

T3SSs can be encoded on mobile plasmids and pathogenicity islands, both of which can be transferred horizontally between distinct gram-negative bacterial species. By contrast, flagellar systems are chromosomally encoded and appear to be transmitted to progeny bacteria largely by vertical descent. When the homologous protein constituents of these systems are examined, the proteins of flagellar basal bodies and those of T3SSs form distinct branches on the phylogenetic tree. This assignment to separate branches implies that the divergence of proteins comprising T3SSs and flagella occurred during the early evolutionary history of these systems, and that their constituents have not undergone shuffling between these two types of systems since their initial divergence.

Related Flagellar Structural Proteins

Several protein constituents of the flagellum share similar sequences, although not all share the same domains. Thus, the FlgB, FlgC, FlgF, and FlgG "transmission shaft" rod proteins all exhibit regions of sequence similarity with the FlgE "universal joint" hook protein (Fig. 3). For example, FlgG has three domains (N, C1, and C2) that are homologous to these domains in FlgE. However, the central domain of FlgE (M) is absent in FlgG. One simple explanation is that the evolutionary precursor of these proteins resembled the smaller ones (FlgG and FlgF), and that a novel domain was inserted into FlgE to generate the larger one. The differing subflagellar locations and functions of these homologous constituents resulted in part from domain insertion and in part from sequence divergence during evolution.

The Flg rod proteins, the FlgE hook protein, and the FlgK hook-associated protein-1 (HAP1) share a common C-terminal domain (C2; the DUF1078 domain in Fig. 3), and most also have similar N-terminal sequences (the Flg_bb_rod domain in Fig. 3). The function of the common C2 domain is unknown, but it could be involved in assembly. The complexity of the flagellar-rod/hook/filament assembly arose in part by gene duplications followed by sequence and domain divergence. Such mechanisms suggest a basis for the evolution of dissimilar protein-protein interactions since homo-multimeric proteins typically have self-associative properties.

The Hook and Flagellar Filament

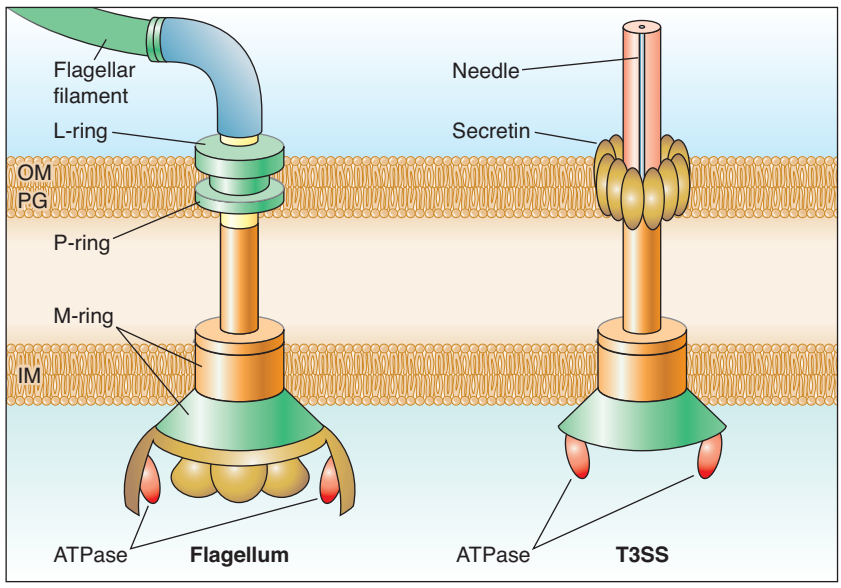
Subunits of bacterial flagellar filaments, called flagellins, vary tremendously in amino acid sequence as well as in quaternary structures of the assembled filament. These filaments can be curly or straight, right-handed or left-handed, and flexible or rigid. Some are modified by methylation or glycosylation. In different strains of *E. coli*, there are nearly 50 sequence divergent flagellins, a surprising observation since each *E. coli* strain usually has only one such protein. This observation suggests that the genes encoding these proteins, in contrast to most other flagellar constituents, were transferred between bacteria with high frequency.

Meanwhile, an Arctic sulfate-reducing bacterium, *Desulfotalea psychrophila*, contains dozens of flagellins encoded within its genome. In this case, we do not know what their functions are. However, we do know that while some flagellar filaments consist of a single flagellin,

Evidence for Internal Repeat Units in Flagellar Proteins

One common evolutionary mechanism for increasing biological complexity and generating diversity involves genetic duplication followed by sequence divergence. Our bioinformatic studies reveal that FlgK of *E. coli*, a hook-filament junction protein, shows significant sequence similarity throughout much of its length with internal regions of bacterial autotransporter-2 proteins, adhesins that contribute to bacterial virulence. The autotransporter-2 proteins contain multiple repeat units of from 7 to 60 amino acyl residues each. There are similar tandem repeats in FlgK homologues, and other flagellar proteins also exhibit repeated sequences. Several structural constituents of the BF thus apparently grew in size and complexity by multiplication of internal repeat units.

FIGURE 2



Comparative schematic depiction of the *E. coli* flagellum (left) and gram-negative bacterial type III secretion systems (T3SS, right). Homologous structures in the basal regions of both structures are similarly colored. Nonhomologous structures are shown as distinct structures. IM, inner (cytoplasmic) membrane; PG, peptidoglycan cell wall layer; OM outer (lipopolysaccharide-containing) membrane. The multilayered MS-ring interacts with the IM, the P-ring anchors the structure to the PG, and the L-ring interacts with the OM. ATP hydrolysis via the FlI ATPases, regulated by FliH, provides the energy required for protein export and assembly (M.H. Saier, Jr., TIM **12:**113–115).

Flagellar-Specific Auxiliary Proteins Derived from Nonflagellar Sources

The P-ring assembly chaperone protein, FlgA, is homologous to other assembly proteins in bacteria, such as pilus chaperone proteins. These proteins share the "β-clip fold" domain with many enzymes and accessory proteins. A common origin correlates with their shared folding and subfunctions.

Another example of a flagellar auxiliary protein sharing a common origin with nonflagellar proteins is FlgJ. It has a C-terminal amidase domain that locally hydrolyzes the bacterial cell wall, making a hole in preparation for flagellar construction. This domain is homologous to many type IV amidases serving a variety of functions unrelated to motility.

Finally, expression of flagellar genes involves transcription using a flagellum-specific sigma factor to initiate mRNA synthesis by RNA polymerase. This sigma factor (σ^F or FliA) is homologous to many other sigma factors in bacteria and could have evolved from a primordial vegetative σ.

others contain many. Surface-exposed residues in flagellins, which are strongly antigenic, are much more variable than the buried residues, which are poorly antigenic. This surface variability provides a mechanism for immune evasion by pathogenic bacteria and is easily explained by natural selection.

In spite of the tremendous sequence variation in flagellins, most if not all of them share a common ancestry. In fact, flagellins are also homologous to the hook-associated protein-3 (HAP3 or FlgL). Moreover, they share domains with the hook-associated protein-1 (HAP1; FlgK), which in turn shares domains with other flagellar proteins. Flagellins may also be partially homologous to the secreted needle complex proteins in T3SSs. This last observation might be expected since the BF and T3SSs share a common origin.

The BF Motor

The BF motor consists of two proteins (MotA and MotB) that are homologous to energizers of outer membrane receptors that concentrate large molecules such as vitamin B_{12} and iron complexes in the periplasm of the gram-negative bacterial cell, between the inner and outer envelope membranes. These proteins comprise transmembrane proton channels. We surmise that MotAB arose as a simple channel complex, allowing gated proton flow across the bacterial membrane, thereby dissipating H^+ and stabilizing the cytoplasmic pH. Once these channels existed, they could have been recruited for other functions such as motility and outer membrane transport energization.

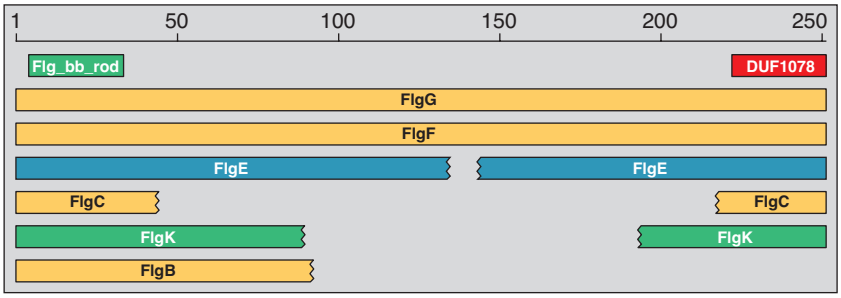
In contrast to the better-characterized flagella of enteric bacteria that rotate either clockwise or counterclockwise, other proteobacteria possess flagella that rotate only in one direction, but with two modes, rotating and stalling, or three modes: fast, slow, and stop. Different bacteria evidently use different motor mechanisms to achieve directed motility.

The Flagellar Assembly ATPase

Flagellar assembly ATPases, the FliI proteins, are homologous to (1) T3SS ATPases and (2) F-type ATPase α- and β-subunits of prokaryotes and eukaryotic organelles such as mitochondria and chloroplasts. Because ATPases energize numerous biological processes, FliI may have evolved independently of flagellar function, having later been recruited to energize flagellar assembly.

In addition to FliI, other constituents of the flagellar assembly system may have derived from a primordial F-type ATPase. Like the BF, F-type ATPases use a rotary motor. They are driven by the flow of protons (H^+) down their electrochemical gradient to make ATP. There are similarities in sequence and structure between the dimeric FliH protein, which interacts with and regulates the hexameric FliI ATPase, and the two subunits of the F-type ATPase, the

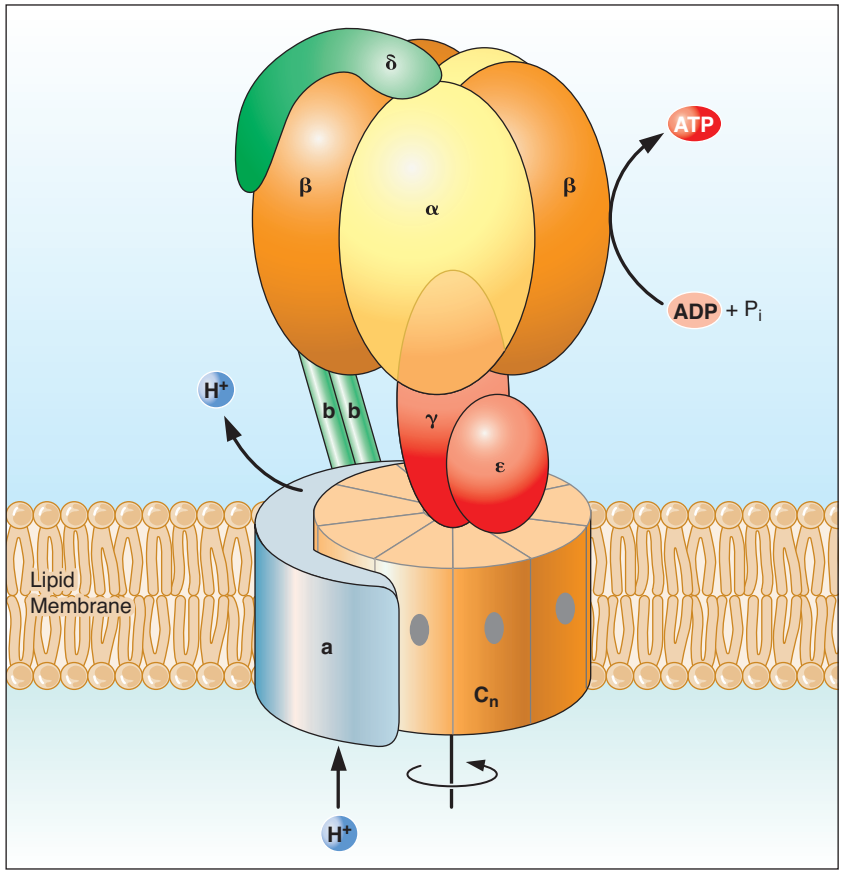
FIGURE 3



Results of a search of the National Center for Biotechnology Information (NCBI) Conserved Domain Database (CDD) using the *Escherichia coli* FlgG flagellar basal-body rod protein as the query sequence. Bars indicate regions of homology for different flagellar proteins: FlgF, FlgE, FlgC, FlgK and FlgB. The common N-terminal domain is labeled Flg_bb_rod while the common C-terminal domain is labeled DUF1078. Amino acyl residue position in FliG is provided at the top. The web address for the NCBI CDD is www.ncbi.nih.gov/structure/cdd/cdd.shtml.

dimeric β- and δ-subunits, which interact with the homologous hexameric α/β-subunit F-ATPase complex (Fig. 4). Moreover, our colleagues Jiwon Youm and Seul-a Shin noticed that in mycobacterial species, the β- and δ-subunits of F-type ATPases are fused, as in FliH. These observations suggest that both FliI and FliH coevolved from F-type ATPase-like subunits, while assuming the roles of energizing and regulating flagellar assembly.

FIGURE 4



Generalized structure of F-type ATPases of prokaryotes, mitochondria and chloroplasts. The heterohexameric $\alpha_3\beta_3$ ATPase complex in F-ATPases is homologous to the homohexameric $(FliI)_6$ ATPase of the flagellum. The dimeric FliH protein of the flagellum, which is known to interact with and regulate $(FliI)_6$, is equivalent of the fused b (N-terminal) and δ (C-terminal) subunits of the F-ATPase which interact with and regulate $\alpha_3\beta_3$ activity. The figure shows the membrane lipid bilayer as well as the movement of a proton from outside to inside the bacterium or organelle which drives rotation of the c-subunit ring and ATP synthesis. The close relationship of mitochondrial and chloroplast F-ATPase subunits with those of bacteria reflect the origin of these eukaryotic organelles from endosymbiotic α-proteobacteria and cyanobacteria, respectively (Iino et al., 2005).

Conclusions about the Evolutionary Development of Bacterial Flagella

Based on research conducted in hundreds of laboratories over several decades, we can outline how the components within the modular bacterial flagellum evolved from several different sources unrelated to an organelle of motility. Steps in this modular development include:

- The flagellar subunit secretion apparatus and T3SSs derived from an ancestral secretion system that used ATP and an ATPase to drive protein export.
- This ATPase and its regulatory protein share a common ancestry with and may have been derived from subunits of rotary F-type ATPases.
- The filament and parts of its connecting "hook complex" possibly arose from bacterial adhesins.
- The motor for flagellar rotation derived from a proton-conducting channel complex that also evolved into motors for molecular uptake into the periplasm of the gram-negative bacterial cell.
- Increased complexity from relatively simple homopolymeric structures resulted from both intragenic and extragenic duplication events, giving rise to multiply-interacting protein constituents.
- Sequence divergence and domain insertion resulted in functional specialization that rendered each protein irreplaceable.
- Flagellum-specific accessory apparatuses were recruited to facilitate flagellar synthesis and assembly.

Natural selection thus accounts for the development of flagellum-driven bacterial motility. We base these conclusions on that which is known, recognizing that much has yet to be understood. Religion and mythology, which both deal with the unknown, should not be confused with scientific hypotheses, which must be testable and cannot invoke supernatural pro-

cesses. The English playwright Oscar Wilde said, "Science is the record of dead religions." In terms of the intelligent design case regarding BF, the current factual analyses force this example to exit the realm of religion and return fully to the arena of science.

SUGGESTED READING

Aizawa, S. I. 2001. Bacterial flagella and type III secretion systems. FEMS Microbiol. Lett. **202**:157–164.
Darwin, C. 1859. On The Origin of Species (web address: http://www.esp.org/books/darwin/origin/facsimile/); see particularly Chapter VI (p. 186–194) where evolution of complex structures from structures that previously served different functions is considered.
Groisman, E. A. and H. Ochman. 1996. Pathogenicity islands: bacterial evolution in quantum leaps. Cell **87**:791–794.
Iino, R., Y. Rondelez, M. Yoshida, and H. Noji. 2005. Chemomechanical coupling in single-molecule F-type ATP synthase. J. Bioenerg. Biomembr. **37**:451–454.
Ito, M., N. Terahara, S. Fujinami, and T. A. Krulwich. 2005. Properties of motility in *Bacillus subtilis* powered by the H^+-coupled MotAB flagellar stator, Na^+-coupled MotPS or hybrid stators MotAS or MotPB. J. Mol. Biol. **352**:396–408.
Musgrave, I. 2004. Evolution of the bacterial flagellum, p. 72–84. *In* M. Young, and T. Edis (ed.), Why intelligent design fails: a scientific critique of the new creationism. Rutgers University Press, Piscataway, N.J.
Nguyen, L., I. T. Paulsen, J. Tchieu, C. J. Hueck, and M. H. Saier, Jr. 2000. Phylogenetic analyses of the constituents of type III protein secretion systems. J. Mol. Microbiol. Biotechnol. **2**:125–144.
Pallen, M. J. and N. J. Matzke. 2006. From *The Origin of Species* to the origin of bacterial flagella. Nature Rev. **4**:784–790.
Saier, M. H., Jr. 2006. Protein secretion systems in gram-negative bacteria. Microbe **1**:414–419.
Wilkins, A. S. 2006. 'Intelligent design' as both problem and symptom. Bioessays **28**:327–329.