OPEN ACCESS

BIO-complexity

Research Article

An Engineering Perspective on the Bacterial Flagellum: Part 2—Analytic View

Waldean A Schulz*

Conceptual Assets, Spokane, Washington, USA

Abstract

This paper, the second of three, takes a *systems biology* view of the bacterial flagellum. The flagellum is the organelle imparting motility to common bacteria. The first paper, "An Engineering Perspective on the Bacterial Flagellum, Part 1—Constructive View," was a *constructive* or *top-down view* from a *systems engineering* viewpoint. It detailed the typical environment, the purpose, the required existing and new resources, the necessary functional requirements, various constraints, the control means, and the requirements for assembly of any kind of bacterial motility organelle. The specification of these requirements was intended to be independent of knowledge about the actual flagellum. A converse approach is detailed in this Part 2. It is an *analytical*, *reductionist*, or *bottom-up view*, which discusses the known 40+ protein components and the observed and inferred structure, control, and assembly of a typical bacterial flagellum. This cellular subsystem is well-researched. Much of that research is reviewed herein. However, the assembly orchestration is illustrated in a form and detail not found elsewhere. The final paper in the series, Part 3, will compare the two views and will conclude with original observations. Those include an ontology of the exceedingly specific protein binding relationships in the flagellum. The latter observation is new and significant.

Cite as: Schulz W (2021) An Engineering Perspective on the Bacterial Flagellum: Part 2—Analytic View. BIO-Complexity 2021 (2):1–16. **doi:**10.5048/BIO-C.2021.2.

Editor: Robert J. Marks II

Received: December 4, 2020; **Accepted:** March 27, 2021; **Published:** September 6, 2021

Copyright: © 2021 Schulz. This open-access article is published under the terms of the Creative Commons Attribution License, which permits free distribution and reuse in derivative works provided the original author(s) and source are credited.

Notes: A *Critique* of this paper, when available, will be assigned **doi:**10.5048/BIO-C.2021.2.c.

*dean@conceptual-assets.com

INTRODUCTION

Systems biology employs methodology and techniques typical of systems engineering. Similarly, reverse engineering of the features of biological organisms leverages both biology and engineering disciplines. The systems perspective on the bacterial flagellum detailed below studies the purpose, functions, components, and structure of a typical bacterial flagellum. The dynamic operation and control of this organelle and the flagellum's assembly stages are also studied.

The bacterial flagellum is a well-researched bacterial subsystem [1,2,3,4] in biology. However, this three-part engineering study takes two essentially independent approaches. First was a *constructive approach*, which was discussed in detail in Part 1 [5]; the other is an *analytical approach*, which is discussed in detail herein.

The first, constructive approach was a *top-down specification*. That is, Part 1 started with specifying the purpose of a bacterial motility organelle, the environment of a bacterium, its existing resources, its existing constitution, and its physical limits, all within the relevant aspects of physics and molecular chemistry. From that, the constructive approach derived the logically necessary functional requirements, the constraints, the assembly needs, and the hierarchical relationships within the functionality. The functionality included a required control subsystem to properly direct the operation of a propulsion subsystem. Those functional requirements and constraints then suggested the few—and very limited—viable implementation schemata for a bacterial propulsion system. The details of one schema were then set forth. A sincere attempt was made to keep the elaboration of this constructive approach logical and as independent as possible from knowledge of the actual flagellar structure.

The second, analytical approach employed here in Part 2 is the converse of the first approach; it is a bottom-up analysis. This Part 2 presents the constituent proteins, observed structure, assembly, and resultant behavior of a typical bacterium. This knowledge has been acquired by microscopic observation, by gene sequencing, by disabling component proteins

(gene "knock-out" experiments), and by other experimental methods. Higher-level organization, functionality, mechanism, and assembly orchestration are hypothesized and inferred from those basic low-level details, though our understanding is incomplete and much still remains unclear.

The account of the flagellum presented here has been accumulated by piecing together the details of bacterial flagella from the cited research literature, as coherently as possible, and by introducing material not found elsewhere.

BOTTOM-UP ANALYSIS OF AN ARCHETYPAL FLAGELLUM

This analytical study includes (a) a diagram of an archetypical flagellum and its perceived configuration into a coherent set of subassemblies, (b) the chemotaxis control subsystem, (c) a listing of the proteins typically included in the subassemblies, and (d) the orchestration of their assembly. The last item is depicted as a large, hierarchal, multi-page graphical network, which is more detailed than found elsewhere.

The specifics herein relate to the two most studied species of bacteria: *Salmonella enterica* (for the structure and assembly details of the flagellum) and *Escherichia coli* (for chemotaxis as the operational control means).

Systems Biology and Reverse Bioengineering

The following reductionist description is consistent with the developing disciplines of reverse engineering of biological systems and systems biology [6]. Both disciplines often start with the molecular, low-level details about existing cellular entities and then organize those details systematically into an inferred hierarchical, structural, and functional model (thus, a bottom-up methodology).

Among the flagella of various species of bacteria, there are differences in the constituent proteins and assembly control [7]; nevertheless, they have comparable, identifiable subassemblies. Herein, the focus is on *Salmonella enterica*. The following description strives for accuracy but will surely contain some debatable details. One legitimate reason for the debatable details is that the referenced papers themselves disagree at some points with each other, especially as later research has shed new light. The goal of this section, however, is not to provide an authoritative compendium of knowledge about flagella, but to portray the exquisite coherence of the protein constituents, their implicit hierarchical structure, and the very intricate assembly process.

STRUCTURE OF AN ARCHETYPAL FLAGELLUM

Figure 1 is a cross-section illustration of the observed and inferred structural subassemblies of a common flagellum, as typically depicted in the references and their illustrations [1,9,10]. The name or description given to each subassembly concisely describes the inferred purpose thereof. The structure and functions of subassemblies are surmised from x-ray crystallography, mutants, cryo-electron microscopy images [11] [12], and gene/protein "knock-out" experiments [13].

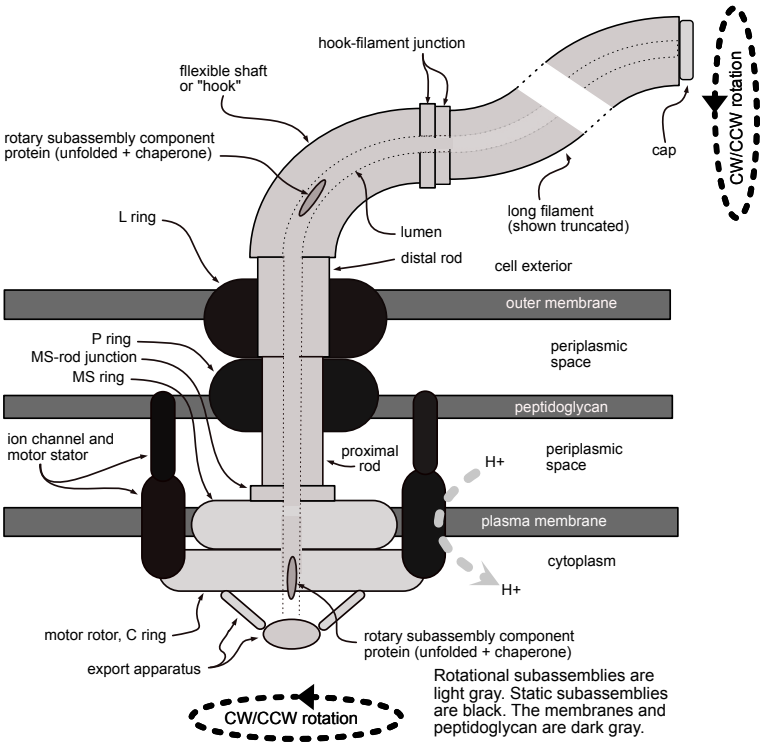

Figure 1: Structure of an archetypical flagellum. doi:10.5048/BIO-C.2021.2.f1

The use of the knock-out experiments in hampering operation of the flagellum to identify protein functionality suggests the generally irreducible nature of the flagellum. Cohen et al. state:

> The bacterial flagellum exemplifies a system where even small deviations from the highly regulated flagellar assembly process can abolish motility and cause negative physiological outcomes. Consequently, bacteria … [possess] robust regulatory mechanisms to ensure that flagellar morphogenesis follows a defined path, with each component self-assembling to predetermined dimensions. [14]

In Figure 1 the motor stator and the motor rotor (the C ring) form an ion-powered electric motor. During operation the stator induces torque on the C ring (see section 3.5.2 below). The induced torque and ensuing rotation of the C-ring is transmitted through the MS ring and MS-rod junction to the rod, which comprises two sections: a proximal rod and a distal rod. The proximal rod transects the peptidoglycan layer; the distal rod transects the outer

cell membrane. Clearly the C ring, the MS ring, the proximal rod, and the distal rod all must be essentially rigid, especially circumferentially, so that the torque is efficiently transmitted along the rotational axis formed by those subassemblies. The static P ring stabilizes the rotational axis by keeping the proximal rod in place like a bushing, and it forms a seal or barrier so no material can pass between the P ring and the rotating proximal rod. The static L ring provides the same two functions, except for stabilizing the distal rod instead. The diagrams in some papers depict binding between the P ring and L ring [10]. That binding would add further stability to the rotation axis within the cell wall. Binding between the P and L rings would precisely locate their locations with respect to each other.

The distal rod extends through the cell wall and binds to the proximal end of the flexible shaft. The flexible shaft is often called the hook (inappropriately) because of its curved shape. It redirects the rotational axis rearward and is sometimes likened to a mechanical U-joint. (A flexible shaft is superior to a U-joint, because a U-joint locks up if the rotational axis bends 90°.) The shaft is flexible axially, like a rubber tube. However, the flexible shaft must transmit torque efficiently "downstream" to the propeller-like helical filament to which it must confer rotation. (There is an adaptor in between the flexible shaft and filament, presumably because the flexible shaft proteins and filament proteins would not otherwise bind strongly.) So the shaft must be rigid circumferentially [15], like a flexible electrical conduit or steel-braided plumbing supply line. This property dictates very special attributes for the copies of the protein which compose the flexible shaft. The subunits of this protein can adopt at least eleven conformations as the shaft rotates [11].

The purpose of the flexible shaft is to redirect the filament so that all the flagella of a bacterium (a bacterium may have many flagella) extend in the same rearward direction and thus can contribute to forward motion in the same direction.

There is a pair of short annuli (ringlike structures) forming the hook-filament junction between the distal end of the flexible shaft and the proximal end of the filament—presumably acting as an adaptor to bind the proteins of the shaft to those of the filament.

The helical filament is the subassembly that converts the (counterclockwise) axial rotation to forward motion, much like a mechanical worm gear, a corkscrew, or a propeller. The helical filament must be sufficiently rigid to maintain its helical shape during forward motion. However, when the rotation is reversed (as is a common case), the helical shape collapses, which induces the bacterium to tumble and reorient to a random new direction of forward travel [1]. The redirection contributes to a random search for nutrients or an escape path from a toxin. The filament proteins therefore are configured very particularly to provide helical mechanical rigidity and transmit torque in one rotational direction but to become flaccid in the opposite direction.

Note that all proteins in each rotary subassembly need to (non-covalently) bind tightly to themselves and to the proteins of the adjacent subassemblies. This is noteworthy: the combinatorial configurations of the ensemble of proteins must be very specifically orchestrated.

Component Proteins of the Subassemblies

Tables 1a through 3 list the plethora of proteins constituting the archetypical flagellum and its control subsystem, expanded from Macnab [1]. The tables hierarchically group the proteins by the subassembly they constitute, as identified in the leftmost column. The proteins of the rotary subassembly of the propulsion/redirection system are shown in Tables 1a-c. The rotary subassembly is further subdivided into the MS ring, the Type III Export Apparatus, the C Ring (motor rotor), the MS-rod junction, the rod cap, the proximal and distal rods, the hook cap, the hook, the hook-filament junction, the filament cap, and the filament. Similarly, the static subassembly proteins, shown in Table 2, make up the P Ring, L Ring, and motor stator. Lastly, the chemotaxis subsystem comprises sensors, control logic, and signal proteins, as shown in Table 3.

For each protein, its usual symbol, its estimated stoichiometry, its cellular location, and its gene/operon template are listed. The means of insertion or assembly path (if known) and the observed order of assembly are included. Finally, the purpose of the assembly or protein is noted.

OPERATIONAL CONTROL OF AN ARCHETYPAL FLAGELLUM

The discussion above detailed the structural aspects of the flagellum, but there is much to be learned from its operation as well: specifically, the control subsystem and the torque generation mechanism.

Chemotaxis: the control subsystem for the flagellum

Auletta [17] describes the chemotaxis subsystem of *Escherichia coli* as an instance of a general control system, very familiar to engineers. Auletta's "Figure 1" is reproduced here as Figure 2. The subsystem comprises an input "physical-chemical signal", the control logic itself, and an output "action." The control logic itself comprises a "processor," a "regulator," and a "decider." The logic includes decoding the input signal to trigger a control state, overriding a default state. Feedback within the logic provides regulation of the state and "feedforward" generation of an output action. Finally, "action feedback" is the result of all the foregoing, which in turn modifies the environment, which then induces a new "physical-chemical signal."

Krell [18], Bren [19], Sarkar [20], and U. of Utah [21]

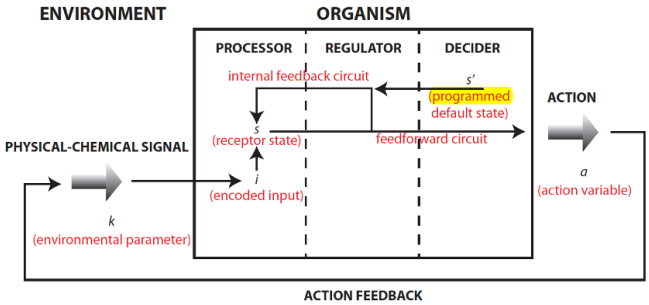

Figure 2: Auletta's diagram of the chemotaxis control system feedback. (Reproduced here under open access Creative Commons CC BY 4.0 license from Auletta, Figure [17]) **doi:**10.5048/BIO-C.2021.2.f2

Table 1a: The rotary subassembly proteins—MS ring and export apparatus

sub-assembly	protein symbol	stoichi-ometry	location	operon	assembly / insertion	assembly order	function / purpose
propulsion / redirection subsystem							
rotary subassembly							
MS ring			cytoplasmic membrane			assembles first	armature
	FliF	24..27		*fliFGHIJK*	Sec pathway	first	armature
Type III Export Apparatus			cytoplasm			after MS ring	protein export apparatus, T3EA
	FlhA	≥2	MS ring center	*flhBAE*	?	after MS ring	target for soluble export complex
	FlhB	≥2	MS ring center	*flhBAE*	?	after MS ring	substrate specificity switch
	FliO	≥1	MS ring center	*fliLMNOPQR*	?	after MS ring	component of T3EA
	FliP	4	MS ring center	*fliLMNOPQR*	?	after MS ring	component of T3EA
	FliQ	≥1	MS ring center	*fliLMNOPQR*	?	after MS ring	component of T3EA
	FliR	≥1	MS ring center	*fliLMNOPQR*	?	after MS ring	component of T3EA
	FliI	?	cytoplasm	*fliFGHIJK*	-	after MS ring	ATPase driver protein
	FliH	?	cytoplasm	*fliFGHIJK*	-	after MS ring	the negative regulator of FliI
	FliJ	?	cytoplasm	*fliFGHIJK*	-	after MS ring	general export chaperone for the Type III export apparatus (T3EA)

Table 1b: The rotary subassembly proteins—C ring and rod

sub-assembly	protein symbol	stoichi-ometry	location	operon	assembly / insertion	assembly order	function / purpose
propulsion / redirection subsystem							
rotary subassembly							
C ring (motor rotor)			cytoplasm			after MS ring	torque generator along with stator
	FliM	32…37	on MS ring	*fliLMNOPQR*	self-assembly	after MS ring	switch: target for chemotaxis CheY-P
	FliN	~110	on MS ring	*fliLMNOPQR*	self-assembly	after MS ring	switch protein
	FliG	24…26	on MS ring	*fliFGHIJK*	self-assembly	after MS ring	torque-generation: rotor protein
MS-rod junction			peptidoglycan			after T3EA	binds MS ring to rod; export gate
	FliE	9	peptidoglycan	*fliE*	T3EA	after T3EA	binds MS ring to rod; export gate
rod cap (temporary jig)						after MS-rod junction'	temporary for assembling rod; punctures peptidoglycan and outer membrane
	FlgJ	5	peptidoglycan to outside cell	*flgBCDEFGHIJ*	T3EA	after MS-rod junction'	(miramidase) punctures peptidoglycan and outer membrane; assembles rod
proximal rod			peptidoglycan			after rod cap	through peptidoglycan layer
	FlgB	7	peptidoglycan	*flgBCDEFGHIJ*	T3EA	after rod cap	part of hollow tube with lumen
	FlgC	6	peptidoglycan	*flgBCDEFGHIJ*	T3EA	after rod cap	part of hollow tube with lumen
	FlgF	6	peptidoglycan	*flgBCDEFGHIJ*	T3EA	after rod cap	part of hollow tube with lumen
distal rod			peptidoglycan to outside cell			after prox rod	through outer cell membrane
	FlgG	26	cell membrane	*flgBCDEFGHIJ*	T3EA	after prox rod	forms tube in helical protein arrangement

Table 1c: The rotary subassembly proteins—hook and filament

sub-assembly	protein symbol	stoichi-ometry	location	operon	assembly / insertion	assembly order	function / purpose
propulsion / redirection subsystem							
rotary subassembly							
hook cap (temporary jig)			outside cell			after rod	temporarily caps hook; helps assemble
	FlgD	5	outside cell	*flgBCDEFGHIJ*	T3EA	after rod	temporarily caps hook; helps assemble
hook			outside cell			after hook cap	much like a u-joint, but does not lock up
	FlgE	132	outside cell	*flgBCDEFGHIJ*	T3EA	after hook cap	forms hollow tube with lumen & helical protein arrangement
	FliK	?	outside cell	*fliFGHIJK*	T3EA	after hook cap	hook length control
hook-filament junction			outside cell			after hook	binds hook to filament
	FlgK	13	outside cell	*flgKL*	T3EA	after hook	binds with FlgE of hook
	FlgL	10	outside cell	*flgKL*	T3EA	after hook	binds with FliC of filament
	FlgN	23?	cytoplasm	*flgMN*	-	after hook	chaperone for FlgK and FlgL
filament cap (jig)			outside cell			after hook-fila. junction	folds the unfolded FliC monomer and places it
	FliD	5	outside cell	*fliDST*	T3EA	after hook-fila. junction	permanent cap; folding chaperone
	FliT	5	cytoplasm	*fliDST*	-	after hook-fila. junction	chaperone: one FliT per FliD
filament			outside cell			after fila. cap	helical rotary-to-linear forward propeller
	FliC	20000	outside cell	*fliC*	T3EA	after fila. cap	forms tube in helical protein arrangement
	FliS	20000	cytoplasm	*fliDST*	T3EA	after fila. cap	chaperone for FliC

Table 2: The static subassembly proteins of an archetypal flagellum

sub-assembly	protein symbol	stoichi-ometry	location	operon	assembly / insertion	assembly order	function / purpose
static subassembly							
P ring			peptidoglycan			after prox rod	peptidoglycan bearing
	FlgI	24	peptidoglycan	*flgBCDEFGHIJ*	Sec pathway	after prox rod	bearing around proximal rod
	FlgA	24	peptidoglycan	*flgAMN*	Sec pathway	after prox rod	chaperone for FlgI
L ring			outer membrane			after distal rod	outer membrane bearing
	FlgH	28	outer membrane	*flgBCDEFGHIJ*	Sec pathway	after distal rod	bearing around distal rod
motor stator			cytoplasmic membrane			after C ring	torque generator along with rotor
	MotB	can vary	cytoplasmic membrane	*motAB-cheAW*	Sec pathway	after C ring	form a piston-like mechanism or a gear
	MotA	can vary	cytoplasmic membrane	*motAB-cheAW*	Sec pathway	after C ring	form a piston-like mechanism or a gear

Table 3: The control subsystem proteins of an archetypal flagellum

sub-assembly	protein symbol	location	function / purpose
chemotaxis control subsystem			
sensor			
	Tar	trans-membrane	an MCP that senses for aspartate and maltose, against nickel and cobalt
	Tsr	trans-membrane	an MCP that senses for serine, against leucine, indole and weak acids
	Trg	trans-membrane	an MCP that senses for galactose and ribose
	Tap or Tcp	trans-membrane	an MCP that senses for dipeptides
	Aer	trans-membrane	an MCP that senses for oxygen
	CheA	cytoplasm	decreasing ligand binding increases phosphorylation; methylation sites
	CheW	cytoplasm	linkage between multiple MCPs and CheA
control logic			
	CheZ	cytoplasm	dephosphorylates CheY-P: provides adaptation (hysteresis) with negative feedback
	CheB	cytoplasm	receptor demethylating enzyme resets the state of the MCP-CheA-CheW complex
signal			
	CheY-P	cytoplasm	binds to motor switch to reverse rotation of flagellum for redirection (tumbling)

describe the implementation of that chemotaxis control in terms of the specific protein interactions and the known specifics of the protein configurations and binding sites. While there are diverse chemotaxis systems for differing bacteria, generally the concentrations of environmental chemistry are sensed by transmembrane methyl-accepting chemotaxis proteins (MCP). In the case of our archetypical bacterium, these MCPs include Tar, Tsar, Trg, Tap (*E. coli*) or Tcp (*Salmonella*), and Aer. They have a periplasmic ligand binding region and a cytosolic signalling region. These are bound to CheA by CheW to form clusters located at one or both polar ends. CheA auto-phosphorylates according to the methylation of the MCPs. Presence of nutrients increases methylation; toxins or repellants decrease it. The methylation state implements the short-term memory of the sensor system [22]. Rebbapragada [23] states, "Repellent binding to a chemotaxis receptor induces a conformational change in the signalling domain [of the MCP] that increases the rate of CheA autophosphorylation. The phosphoryl residue from CheA is transferred to CheY." CheY-P in its phosphorylated state, diffuses to a flagellum, and binds to the flagellar rotor. That causes the rotor to switch the direction of rotation of the filament so that the bacterium tumbles. Tumbling causes a random new direction for forward travel after CheZ dephosphorylates CheY-P. Then CheY unbinds from the switch, default (counterclockwise) rotation ensues, and forward travel resumes. This periodic tumbling occurs about every second [24]. The upshot is a biased random "climb" up an attractant gradient [22].

Meanwhile the methyl esterase protein CheB demethylates the MCP-Che-CheW complex, eventually resetting it back to its non-signalling state. The response of CheB is slower than the transfer of phosphorylation to CheY, so CheY-P can interact with the flagellar motor rotor before CheW responds (as a kind of delayed negative feedback). In parallel, CheZ removes the phosphorylated state of CheY. So, CheB and CheZ provide adaptation (hysteresis) by a time-delayed negative feedback. If stimuli are present in abundance, the phosphorylation of CheA outdoes the negative feedback effect of CheB. If stimuli decrease, then the effect of CheB starts winning out, too few CheY proteins are phosphorylated, and the flagellum rotation reverts to its default rotation.

This control system alters the period between tumbles, with longer periods occurring in the presence of increasing attractants [22] to prevent unneeded redirection.

Torque Generation

In Figure 1 the rotation is powered by ions passing from the lower pH periplasmic space to the higher pH cytoplasm. Typically, these are hydrogen ions, but some bacterial species use sodium ions instead [9].

One hypothesis says the ions reconfigure the MotA proteins to cooperatively interact with ratchet-like cogs in the C-ring to induce a local tangential force resulting in torque. Typically, multiple stator MotA complexes provide combined torque [25]. It has been suggested that these form a microscopic electric stepper motor with MotA forming a piston-like plunger [26]. To convert a tangential force on the C ring to rotation requires the stator and the axis of rotation both to be held firmly in place. The direction of rotation is controlled by configuration changes in the C ring, which act like a direction switch (or a brake/clutch in some bacteria) [24,27].

However, the most recent research, based on electron cryotomography, hypothesizes that each MotA subunit forms a tube that rotates clockwise around MotB and generates torque in response to the ion flow:

> The CW rotation of MotA could generate the torque for C ring rotation in either CCW or CW directions through different interaction modes with FliG at the top of the C ring, that is, either with the inner side of the tube (closer to the motor axis) or with the opposite outer side… This is just like a two-cogwheel gear system, composed of small and large cogwheels, with cogs at the cylinder edges that can switch their relative positions; the small cogwheel (MotA) always generates torque in one direction, but it can mesh with the large cogwheel (FliG in the C ring) via either its internal or external edges, hence driving rotation of the large cogwheel in either direction. [28]

That is, the MotA in effect forms a "gear" and always rotates in the same CW direction. During normal CCW rotation for forward motion, the C-ring's FliG engages with the side of MotA that is nearer to the C-ring axis. During CW rotation for tumbling, the MotA reconfigures so the C-ring engages with the side of the MotA cog gear that is farther from the C-ring axis. If this hypothesis is correct, the protein configurations of FliG and of MotA are exquisitely matched.

The power output is nearly 100% efficient. Further, the torque must be continuous and relatively large to keep the flagellum rotating at up to 100,000 rpm [28]. How much does this also need to overcome Van der Waals force and stiction?

Research by Nord et al. [25] indicates the relationship between flagellum speed, number of stator units (which can vary dynamically depending on the required torque needed), and load.

ASSEMBLY OF AN ARCHETYPAL FLAGELLUM

Usage of Systems Biology Graphical Notation

Figures 3 and following use the Systems Biology Graphical Notation, SBGN [29], to depict the intricate interdependencies of the assembly process of a typical flagellum. The assembly details may differ among various bacterial species, but the Figures are elucidatory and visually demonstrate the coherent orchestration of the assembly process.

Figures 3 through 24 are portions of one large assembly dependency graph, subdivided for convenience of discussion and document layout. Figures 3 through 6 can be interpreted as the *fabrication* of the base materials. Figures 7 and following show the dependency ordering of assembly stages, each adding another flagellum subassembly. Each stage of a nascent flagellum is a separate Figure depicting a subgraph, which itself shows a set of static and/or rotary subassemblies.

Before describing the following dependency network graph, note its glyphs (nodes with meaningful shapes) and their attributes, which are part of the SBGN. The following explains the particular usage of SBGN here. First, the dark bar at the bottom of many of the glyphs means that that entity appears in multiple places in the overall graph. The first appearance of a glyph does not have the bar, but subsequent instances of that

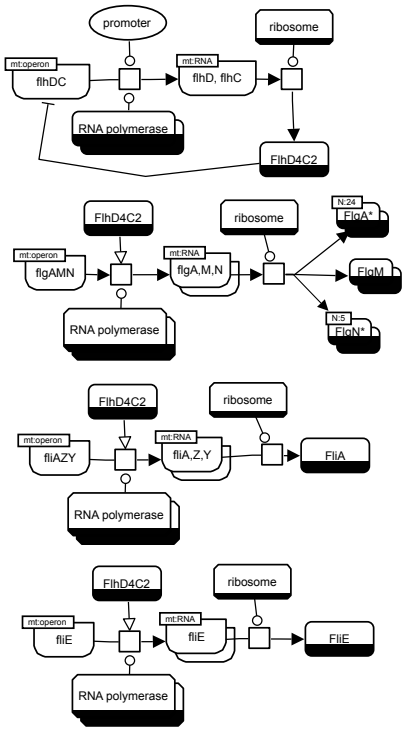

Figure 3: Transcription of initial flagellar proteins. Each glyph (node) represents a protein complex, one or more instances of protein, a whole subassembly, or a process. Chaperones or scaffolding proteins (not part of the eventual structure) are marked with an asterisk. The rectangles in the upper left of a node show the count of protein copies represented by it or identify the class of the entity represented by it. The dark bar at the bottom of a node means that it appeared previously. The small unnamed square glyphs are processes; the arrow with a black head through the square glyph indicates the major input and output of the process. An arrow with a white head indicates stimulation of the process by some entity. An arc with an open circle head indicates the influence of a catalyst. **doi:**10.5048/BIO-C.2021.2.f3

glyph do. If what the glyph represents is undergoing change then the bar is not present. Each glyph may represent a protein complex, one or more instances of a protein, a whole subassembly, or a process. Oval glyphs represent conditions or attributes that must hold. The rectangles in the upper left of a glyph provide the count of protein copies represented by it or identify the class of the entity represented by it. Most glyphs are labeled with the standard cell component names. Names of chaperones or scaffolding proteins (not part of the eventual structure) are marked with an asterisk.

The small unnamed square glyphs are processes, and the arrow (with a filled head) directed through the square glyph indicates the major input and output of the process. An arrow with an open triangular head indicates stimulation of the process by some entity. An arrow with an open round head indicates the influence of a catalyst. An arrow with a bar as the head indicates inhibition by some entity. Thin-walled rectangles enclose subgraphs which are expanding in subsequent Figures. Rectangles with beveled corners indicate subassemblies. Thick-walled rounded rectangles with round corners indicate physical containers—herein, the cell membranes. Within the latter container rectangles, each subassembly is so placed to represent

the relative location of the subassembly to the cell membranes. Further, adjacent subassemblies, which need to bind to each other, abut.

The bindings between the constituent proteins are hydrogen or ionic—not covalent. In the case of chaperones and scaffolding proteins, weak Van der Waals forces might do the temporary binding, but they may be too weak given that Brownian motion is significant at the molecular scale. In any case, chaperones and scaffolding proteins need to temporarily bind to other proteins but then unbind at a specific location or assembly step. This clearly implies specificity in the proteins' configurations.

Flagellum Assembly Dependency Network

All flagellar systems coordinate flagellar gene expression through a transcriptional hierarchy central to an integrated regulatory network of multiple regulatory components. These networks exhibit a number of conserved circuit architectures reflective of the strong conservation found within the structural components of the flagellum [30].

Figures 3 and following are an amalgam of the assembly dependencies from a number of sources: Fig 1 of Macnab [1], Fig 1 of Minnich [2], Fig 9A of Fitzgerald [13], Fig 1 of Karlinsey [33], and Fig. 1 of Cohen [14].

Figure 3 illustrates the fabrication of flagellar proteins from specific operons using the normal gene expression machinery of the bacterium. In particular, assembly is triggered when transcribing flhDC and so forming four FlhD proteins and two FlhC proteins, which are the Class I proteins translated from the flhDC operon. This complex in turn activates the genetic expression of the Class II proteins shown and auto-represses the transcription of flhDC [31].

Figure 4 illustrates the fabrication of further proteins. Details related to the formation of subassemblies are hidden inside the various subgraph rectangles. The particular protein components of each subassembly are shown. Note the elliptical graphs labeled "sequencing." That represents the ordering in which the subassemblies seem to form. The mechanism is partly, but not completely, understood; part of the explanation is the order in which operon translation activated or inhibited. Nevertheless, the proteins of the same class, apparently available at the same time, still are assembled on the flagellum in the correct number and correct relative relationship to each other.

Kalir et al. [32] found that "the observed temporal program of transcription was much more detailed than was previously

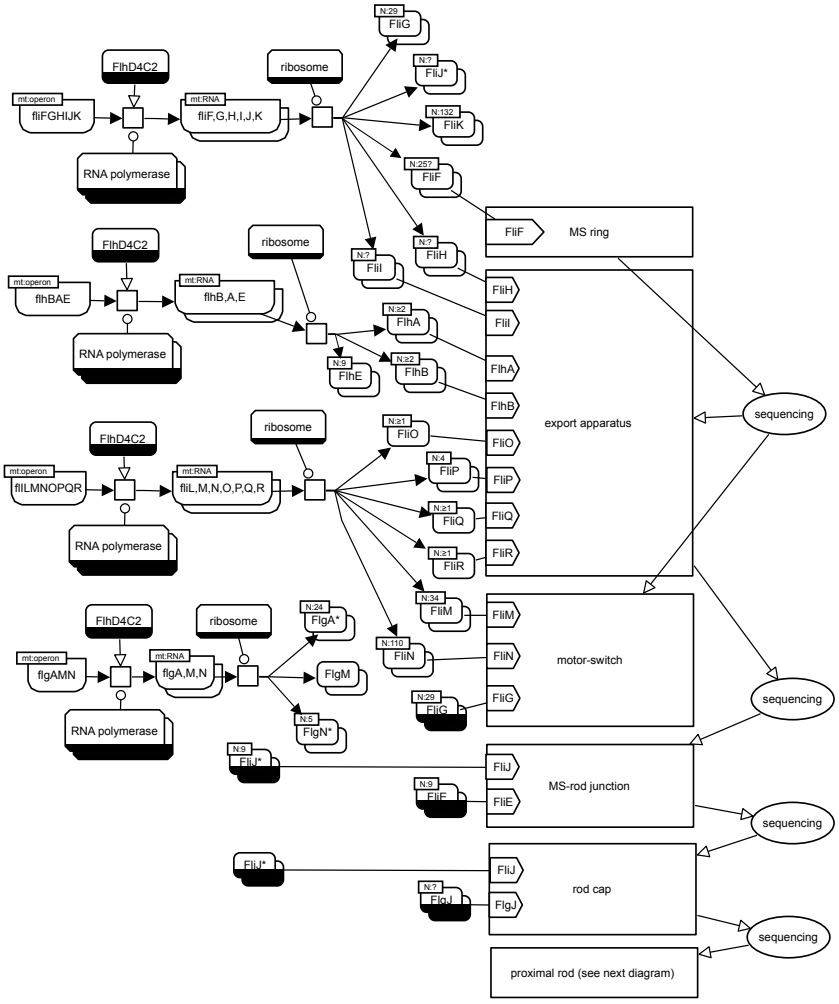

Figure 4: Transcription of more flagellar proteins. Thin-walled rectangles enclose subgraphs which are expanded in subsequent Figures.
doi:10.5048/BIO-C.2021.2.f4

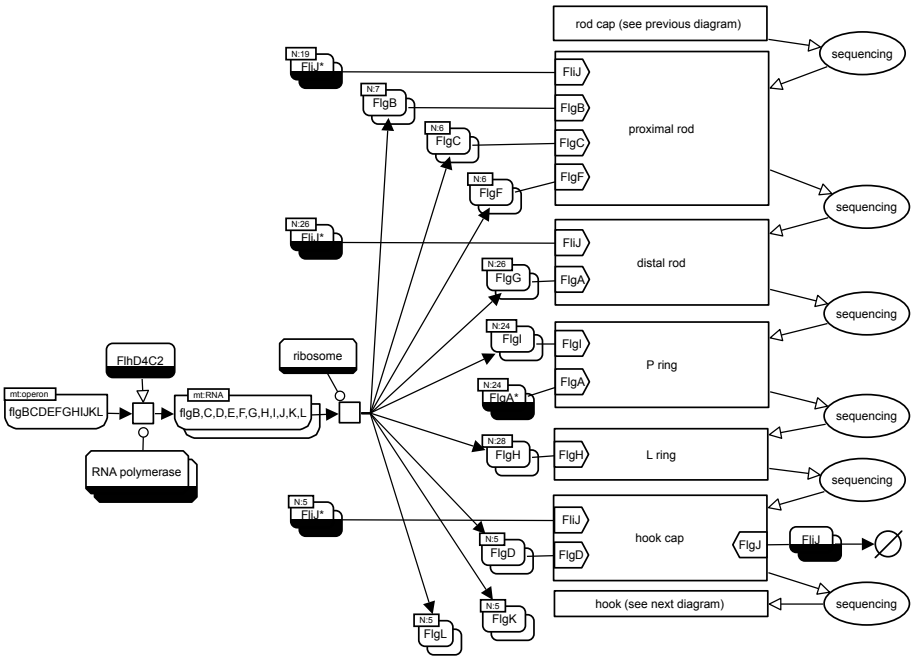

Figure 5: Transcription of yet more flagellar proteins. Thin-walled rectangles enclose subgraphs which are expanded in subsequent Figures.
doi:10.5048/BIO-C.2021.2.f5

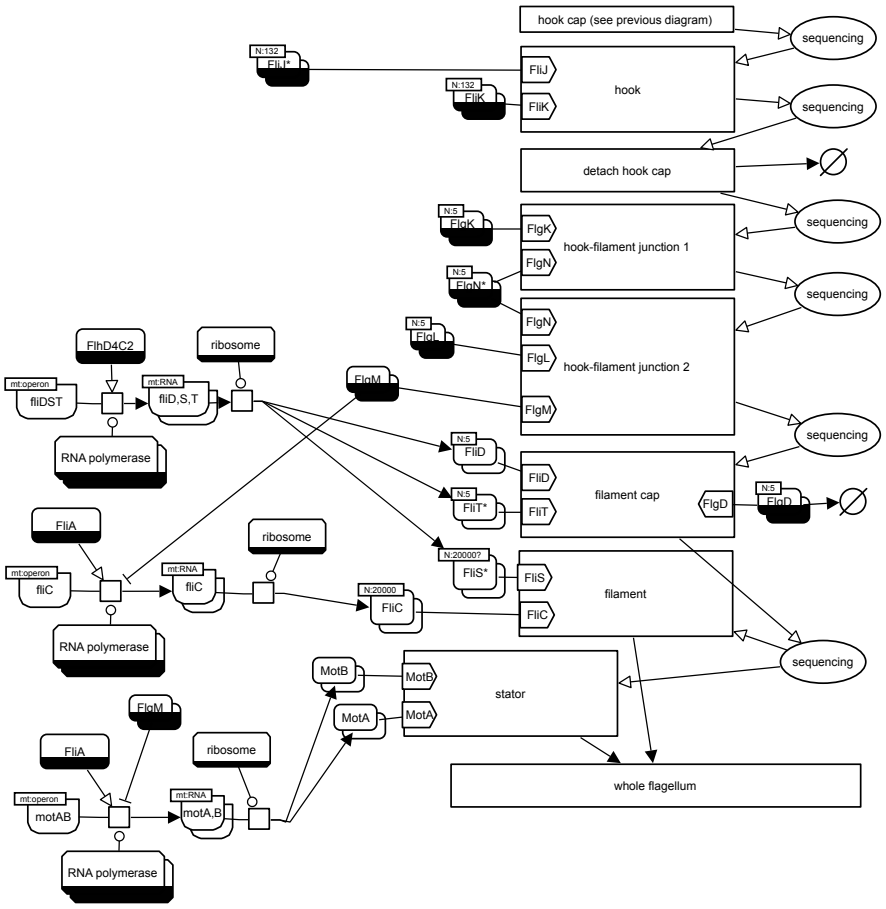

Figure 6: Transcription of final flagellar proteins. Thin-walled rectangles enclose subgraphs which are expanded in subsequent Figures.
doi:10.5048/BIO-C.2021.2.f6

thought and was associated with multiple steps of flagella assembly." They temporally refined the classic three classes of operon expression—specifically for *E. coli* in this case—as follows (in order of time):

Cluster 1:	Class I: *flhDC*
Cluster 2:	Class 2: *fliL*
	Class 2: *fliE, fliF*
	Class 2: *flgA, flgB, flhB*
	Class 2: *fliA*
	Class 3: *fliD, flgK*
	Class 3: *fliC*
Cluster 3:	Class 3: *meche, mocka*
	Class 3: *flgM*

Similar refinement would surely also apply to *Salmonella enterica*. Kalir et al. [32] then suggest how the sequencing might be accomplished:

> A simple hypothesis for the mechanism underlying the temporal order of promoter activation within classes 2 and 3 is that the DNA regulatory sites in the promoter regions of the operons are ranked in affinity. As the concentration of the relevant transcription factor (FlhDC, FliA) gradually increases in the cell, it first binds and activates the operons with the highest affinity sites, and only later does it bind and activate operons with lower affinity sites.

Figures 5 and 6 continue the assembly steps following Figure 4. Note that insertion and assembly of the P and L rings surely must respectively follow assembly of the proximal and distal parts of the rod. How else could the ring proteins locate properly? The question is how the P and L rings make their way to and become inserted around the proximal and distal portions of the rod. The details of how the ring proteins are precisely positioned needs further explication. This is interesting because the P and L ring proteins do not bind with the rod proteins. Therefore, the Figures 3 to 6 and following presume a later assembly [2] for the L and P rings and stator than was shown in Macnab [1].

The above observation about the P and L rings surely holds for each stator MotB-MotA complex, which form around the MS ring. The actual number of MotB-MotA complexes appears to be dynamic and relate to the torque needed for motility [25]. The latter would relate to the viscosity of the fluid.

Figure 7 illustrates that the MS ring seemingly is the first subassembly to assemble [1] What determines the location where it begins to form? That unknown is signified by the square process glyph with a question mark. The location does not seem to be random, but it differs among bacterial species [34]. The MS ring proteins must each have a folded configuration with non-covalent binding in just the right places so that the proteins bind tightly to each other. Further, each of the about 25 FliF proteins forms a segment of an annulus (ring), providing a functional hole in the center. The hole is the entry to a lumen forming a delivery channel for subsequent assembly. The process with a question mark inside the rotary assembly indicates the unexplained source of design for such necessary features. The

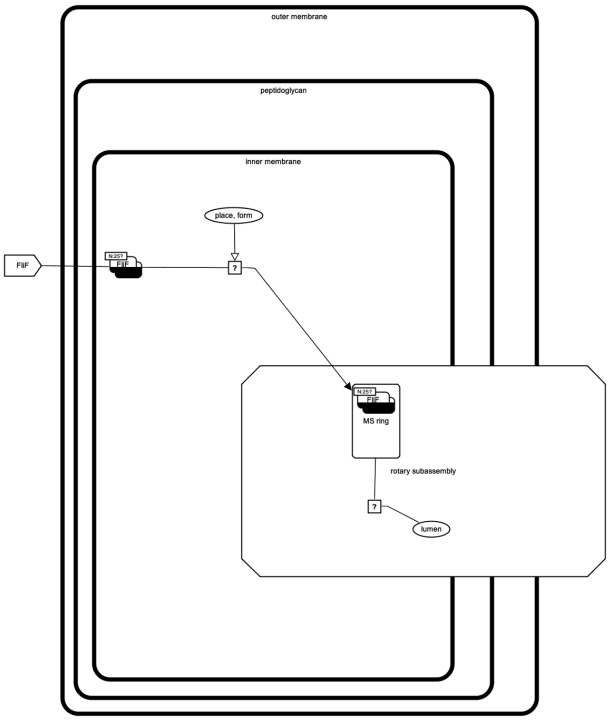

Figure 7: Assembly of the MS Ring. Thick-walled rectangles with round corners indicate physical containers—herein, the cell membranes. Rectangles with beveled corners indicate subassemblies. Within the membranes, the subassemblies are so placed to represent the relative location of the subassembly to the membranes. **doi:**10.5048/BIO-C.2021.2.f7

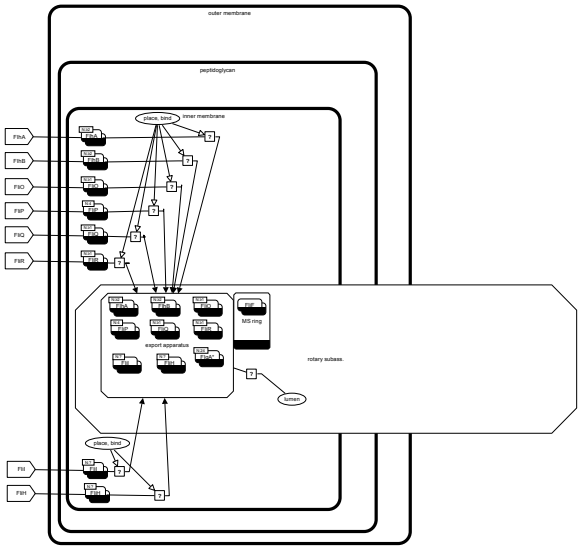

Figure 8: Assembly of the export apparatus. Rectangles with beveled corners indicate subassemblies. Thick-walled rectangles with round corners indicate physical containers—herein, the cell membranes. Within the membranes, the subassemblies are so placed to represent the relative location of the subassembly to the membranes. **doi:**10.5048/BIO-C.2021.2.f8

hole in the MS ring is also where the Type III export apparatus forms, as shown in Figure 8. The export apparatus sequences and inserts proteins in the lumen, through which the proteins will pass and form each rotary subassembly from the inside. The

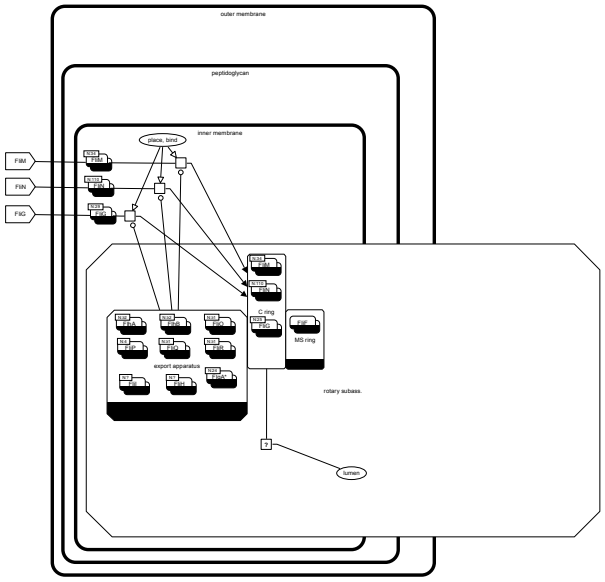

Figure 9: Assembly of the C Ring. Adjacent subassemblies, which need to bind to each other, abut. **doi:**10.5048/BIO-C.2021.2.f9

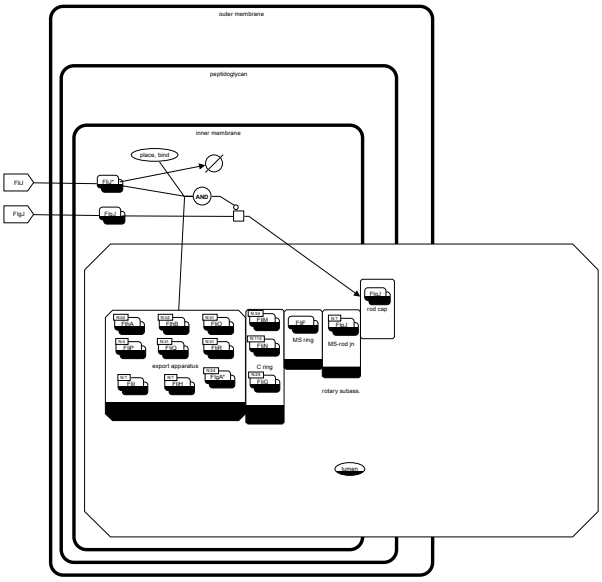

Figure 11: Assembly of the rod cap. The rod cap is the "scaffolding" needed to assemble the rod. **doi:**10.5048/BIO-C.2021.2.f11

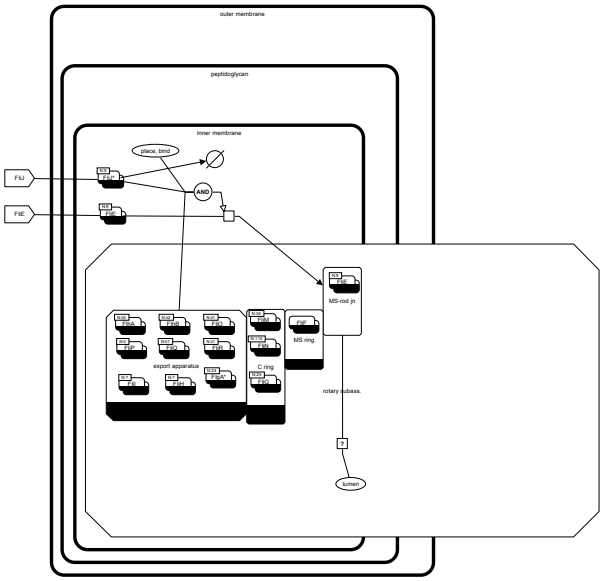

Figure 10: Assembly of the junction between MS ring and the proximal rod. Adjacent subassemblies, which need to bind to each other, abut. **doi:**10.5048/BIO-C.2021.2.f10

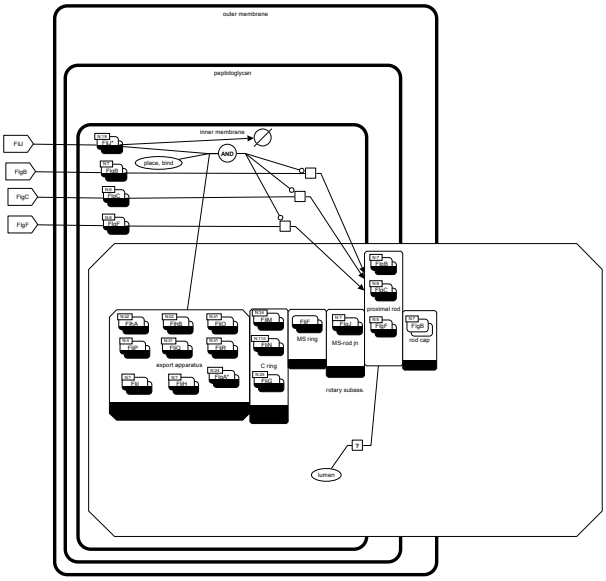

Figure 12: Assembly of the proximal rod. **doi:**10.5048/BIO-C.2021.2.f12

export apparatus involves more proteins because it forms an active machine powered by ATP proteins. Chaperones are used to help ferry proteins through the lumen to their destination in a linear, unfolded state and to prevent polymerization [35].

In Figure 8 and subsequent figures, the square glyph associated with the lumen and labeled with "?" signifies not so much a process but the requirement that each of the rotary subassemblies manifests the remarkable characteristic of providing one segment of the continuous lumen needed for guiding proteins to their destination during flagellar assembly.

Figure 9 shows the assembly step of forming the C ring, which is the rotor part of the motor. Because the C ring, like the export apparatus, binds to the MS ring, it can form independently of the export apparatus. Again, note that the C ring proteins have a unique configuration [36], which causes tight binding to each other and to the MS ring.

Figure 10 depicts the assembly of the junction between the MS ring and the proximal rod. As shown, it comprises nine proteins, which form an adaptor between the proteins of the MS ring and those of the proximal rod. Without this adaptor, the binding sites on the MS ring proteins would not likely match well with those of the proximal rod.

Note that the various Figures approximately depict the locations of the sundry subassemblies within the cell's membranes. Further, the Figures show an unknown process (indicated by the question mark), which was responsible for the inclusion of the lumen, a feature which all of the rotary subassemblies must possess.

Figure 11 shows the stage adding the rod cap, which is a temporary protein complex accomplishing two functions. It prevents leakage of assembly components to the space beyond where the rod is being assembled. It also aids as a scaffolding or jig, which folds, directs, and binds the next supplied rod protein to its assembly location on the nascent rod. Figures 12 and 13 show the stages for which the rod cap does its work. The difference between the two is that the proximal rod must pierce and transect the peptidoglycan while the distal rod must pierce and transect the outer cell membrane. Cohen et al. [14] say that when the distal rod cap hits the outer membrane the P and L rings form around the shaft and the L ring "forms an outer-membrane pore and results in the initiation of hook polymerization."

Figure 14 is the subgraph showing the assembly stage of the P ring transecting the peptidoglycan. Presumably this assembles after the proximal rod is in place in order to be correctly located. Figure 15 is the subgraph showing the assembly of the L ring bridging the outer cell membrane. Because the L ring is more distal than the P ring, it seems that the L ring might

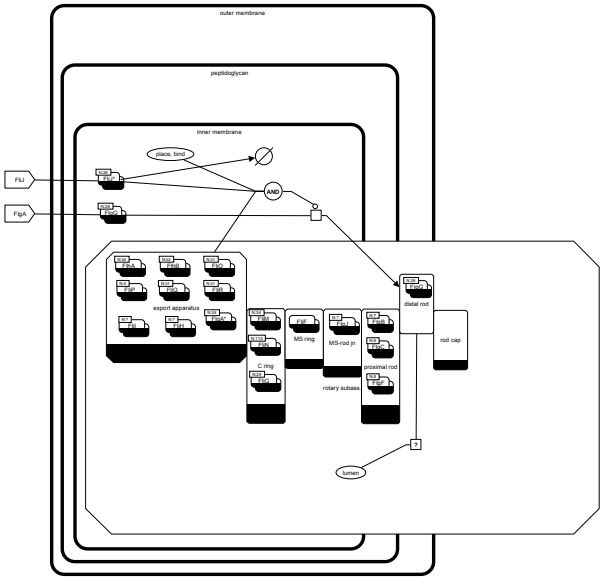

Figure 13: Assembly of the distal rod. doi:10.5048/BIO-C.2021.2.f13

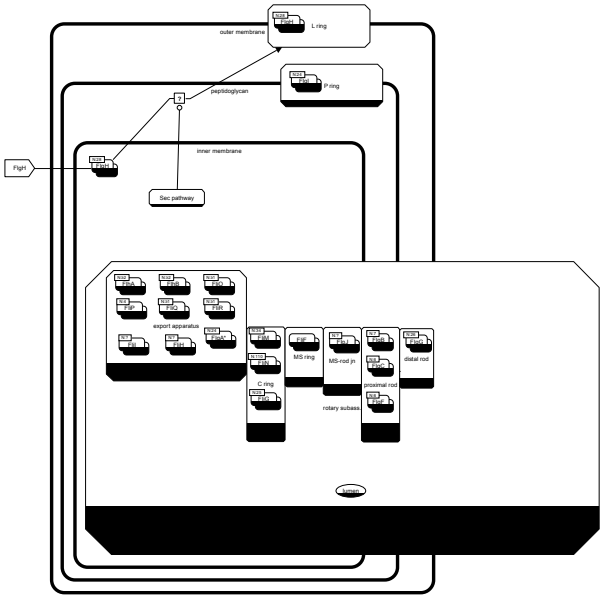

Figure 15: Assembly of the L Ring. doi:10.5048/BIO-C.2021.2.f15

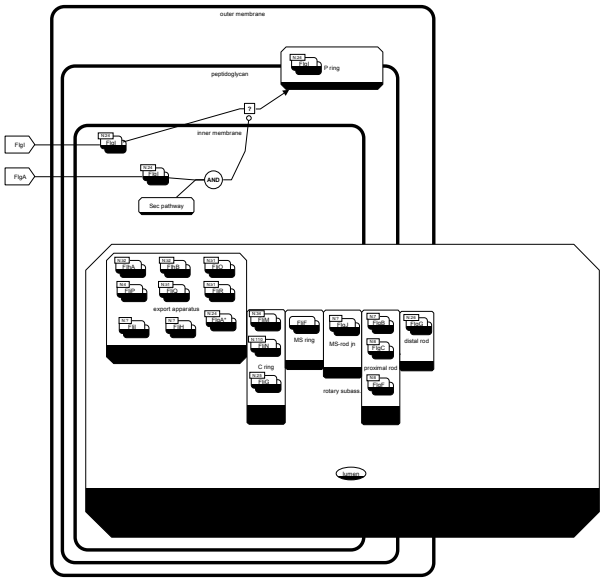

Figure 14: Assembly of the P Ring. doi:10.5048/BIO-C.2021.2.f14

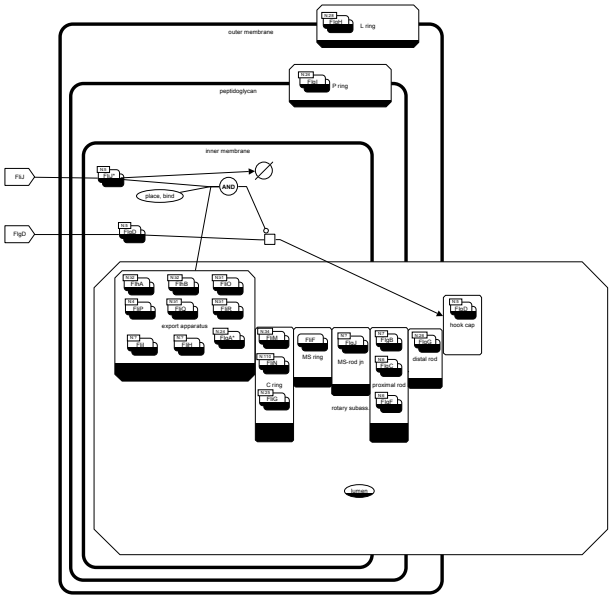

Figure 16: Assembly of the hook cap. doi:10.5048/BIO-C.2021.2.f16

form first. However, the L ring must form after the distal rod is formed, because the ring must position itself around the rod and form the outer-membrane pore. The Sec pathway, a ubiquitous means used elsewhere in a bacterium, is the tool for inserting the P and L ring proteins into the peptidoglycan and outer cell wall respectively [1].

Figure 16 shows the assembly of the hook cap initially onto the distal rod. The cap is the scaffolding protein that acts as an assembly jig. During assembly the hook cap helps fold each unfolded hook protein coming through the lumen and directs it to the vacant next position. The formation of the hook itself is shown in Figure 17. The cap also prevents the hook proteins from escaping out into the external environment. When the hook is complete the hook cap is released into the environment, as shown by the arrow to the SBGN symbol ∅ in Figure 18.

Figures 19 and 20 are the protein dependency networks for assembling the two hook-filament joint rings. They apparently act as a junction or adapter between the hook and the filament. Presumably this is because the proteins of the hook and those of the filament would not bind together directly. So, the amino acids of the junction rings provide non-covalent binding regions in the right locations to bind hook to junction 1, junction 1 to

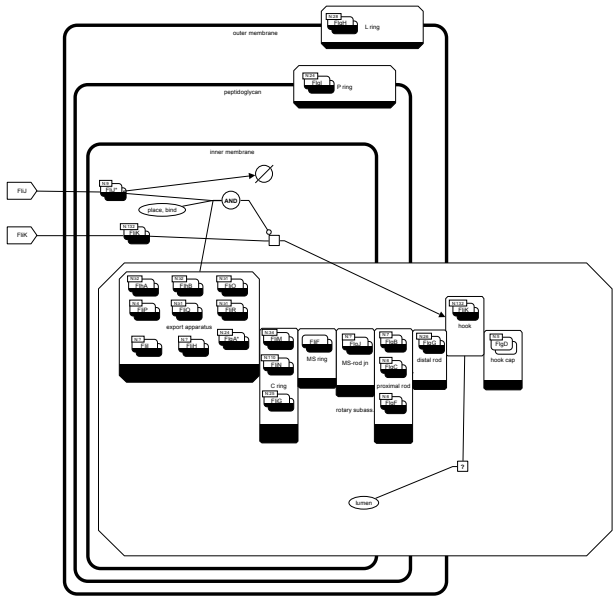

Figure 17: Assembly of the hook. **doi:**10.5048/BIO-C.2021.2.f17

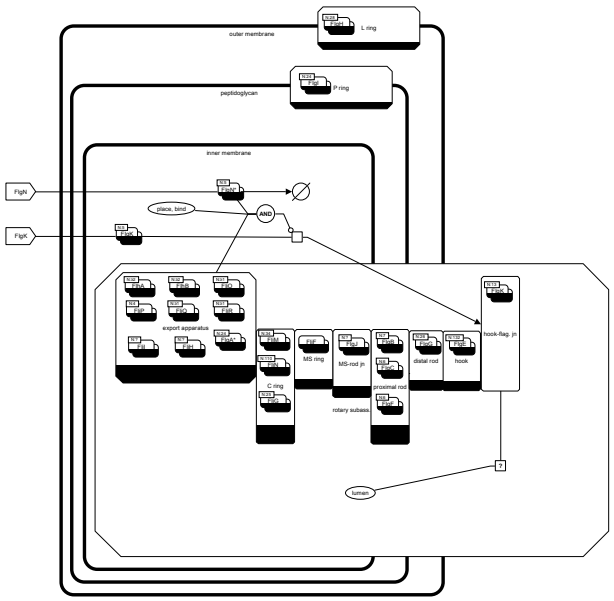

Figure 19: Assembly of the hook-filament junction 1. **doi:**10.5048/BIO-C.2021.2.f19

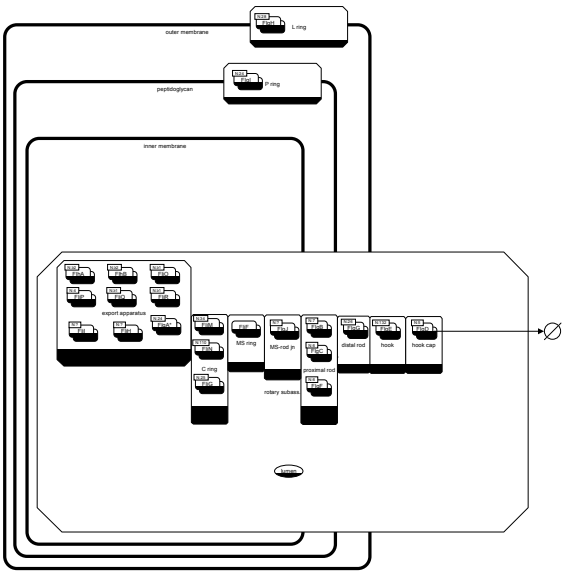

Figure 18: Detachment of the hook cap. **doi:**10.5048/BIO-C.2021.2.f18

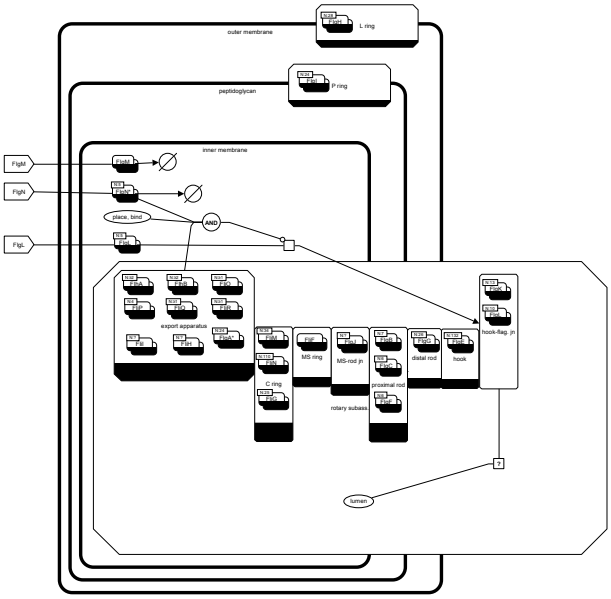

Figure 20: Assembly of the hook-filament junction 2. **doi:**10.5048/BIO-C.2021.2.f20

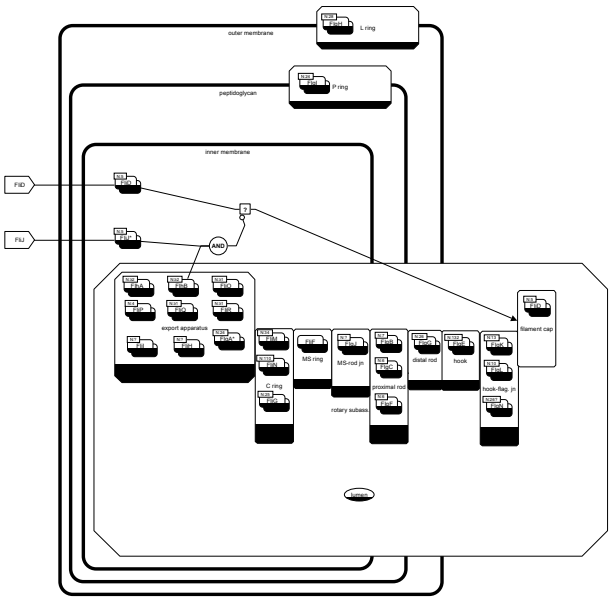

Figure 21: Assembly of the filament cap. **doi:**10.5048/BIO-C.2021.2.f21

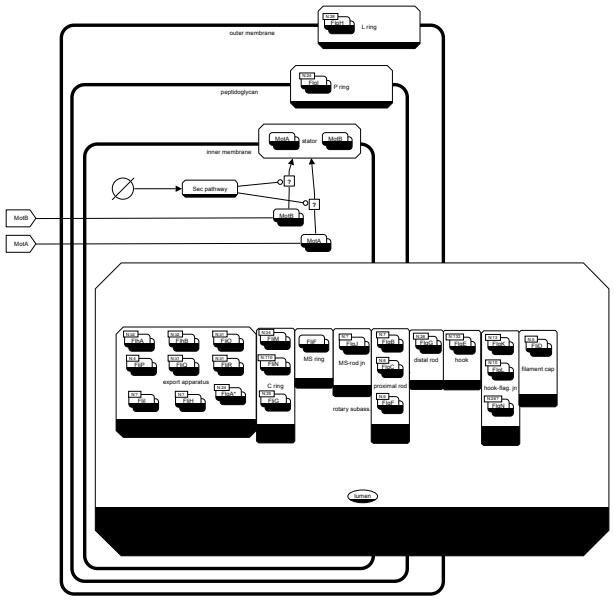

Figure 23: Assembly of the stator complexes. **doi:**10.5048/BIO-C.2021.2.f23

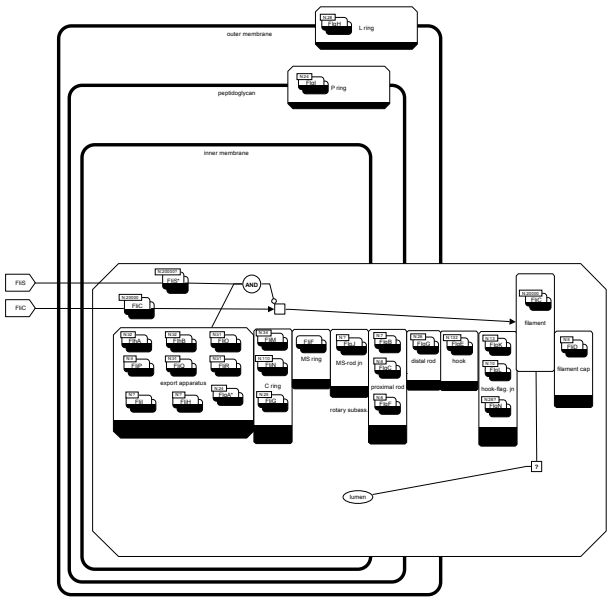

Figure 22: Assembly of the filament. **doi:**10.5048/BIO-C.2021.2.f22

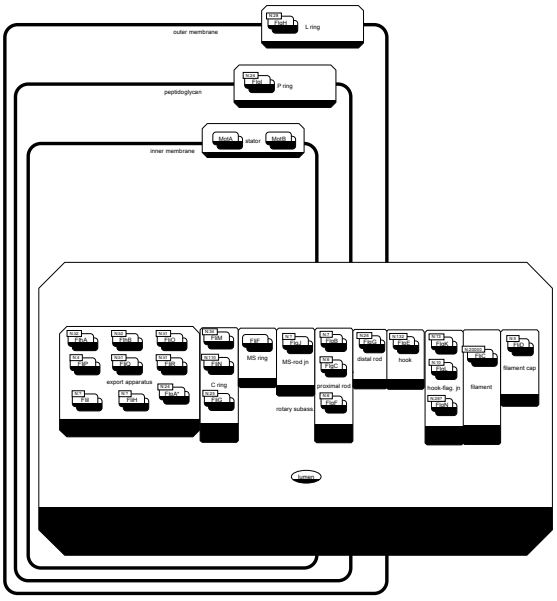

Figure 24: Assembly complete. **doi:**10.5048/BIO-C.2021.2.f24

junction 2, and junction 2 to filament.

Figure 21 shows the assembly of the filament cap, another scaffolding complex. It seals the lumen from the external environment—as did the other scaffolding protein complexes. It also helps fold and position each filament protein arriving at the next assembly location after passing through the lumen unfolded. That is represented in Figure 22.

Figure 23 shows the formation of the stator complexes. They presumably can form simultaneously with the assembly of the more distal rotary subassemblies.

Finally, Figure 24 illustrates the assembled flagellum.

FUTURE WORK

The future work for an evolutionary biologist is twofold: (1) to provide a *detailed* explanation for how all the tightly constrained interlocking coherence described above could have evolved stepwise and naturalistically under real-world constraints; (2) to show evidence that such a scenario actually occurred in the past.

For the molecular biologist, the control and sequencing of protein fabrication and assembly begs for further elucidation. Major unknowns include what controls the stoichiometry of the various subassemblies and how the correct sequence of

proteins is gated through the type 3 export apparatus. Further, a study of the binding sites of the proteins to each other and to their immediate neighbors would be a helpful endeavor, the principles of which could have much wider application. This is nontrivial, because there are quite a few protein-protein pair bindings to consider—both those pairs that bind and the pairs that should *not* adhere to each other (lest they impede the assembly process and later operation). Even more, how, when, and why do flagellar chaperone proteins attach and then detach from the proteins they guard or direct? What controls that?

For the molecular modeller, future work could simulate in detail the geometry and specific binding loci of each pair of bound proteins. How do they become oriented to each other by electrostatic attraction? Can the binding force be estimated? Further simulation might illustrate how the chaperones and scaffolding proteins bind and then release. Still further, the putative interaction between the gear-like C-ring and rotating MotA proteins needs verification and further elaboration. Lastly, one might model the interface between the L and P ring proteins to illustrate (1) how they, together with the rod, form a seal which does not allow any but the smallest molecules past their interface, and (2) why the rings still do not bind with the rod.

CONCLUSIONS

The above facts about the intricacy of the structure, control, function, and assembly of the flagellum are objective. Engineers and patent offices always attribute functional devices—even trivial ones—to some intelligent designer(s), implementor(s), or inventor(s). Meanwhile, the evolutionary biological community has yet to hypothesize a *likely, detailed, step-by-step* scenario to explain how the flagellum and its control system could have been blindly engineered naturalistically. Yet even that would still fall short of real evidence that such a thing actually happened, given real-world constraints. The flagellum seemingly is irreducible. How would portions of an incomplete, nascent flagellum be protected from degradation for generations while the remainder was yet to be gradually added? If some of the subassemblies discussed above could be omitted, what function would result?

These are real questions, and the challenge is to answer them. The questions are hard. Knowledge is still too limited to answer them. Meanwhile, it seems disingenuous to dismiss teleology and intelligent causation, when so much is already known about the apparently ingenious, coordinated hierarchical assembly, control, and function of the flagellum. Yet, the dominant explanatory framework assumes such a mindless, unimaginative, undirected process.

REFERENCES

1. Macnab RM (2003) How Bacteria Assemble Flagella. Annu Rev Microbiol 57:77–100. **doi:**10.1146/annurev.micro.57.030502.090832
2. Minnich SA, Meyer SC (2004) Genetic analysis of coordinate flagellar and type III regulatory circuits in pathogenic bacteria. Design and Nature II. Collins MW, and Brebbia CA, eds. WIT Press (Southampton, UK).
3. Beeby M, Ferreira JL, Tripp P, Albers S-V (2020) Propulsive nanomachines: the convergent evolution of archaella, flagella and cilia. FEMS Microbiology Reviews 44, 253–304. **doi:**10.1093/femsre/fuaa006
4. Jarrell K, ed. (2009) Pili and Flagella: Current Research and Future Trends. Caister Academic Press (Poole, UK).
5. Schulz W (2021) An Engineering Perspective on the Bacterial Flagellum: Part 1—Constructive View. BIO-Complexity 2021 (1):1–14. **doi:**10.5048/BIO-C.2021.1
6. Alon U (2007) An Introduction to Systems Biology: Design Principles of Biological Circuits. Chapman & Hall/CRC (Boca Raton, FL).
7. Chaban B, Coleman I, Beeby M (2018) Evolution of higher torque in *Campylobacter*-type bacterial flagellar motors. Scientific Reports 8:97. **doi:**10.1038/s41598-017-18115-1
8. Macnab RM (2003) How Bacteria Assemble Flagella: Annu. Rev. Microbiol 57, 77–100. **doi:**10.1146/annurev.micro.57.030502.090832
9. McLatchie J (2017) The Bacterial Flagellum: A Motorized Nanomachine. March 9, 2013. https://www.scribd.com/document/106728402/The-Bacterial-Flagellum
10. Imada K (2018) Bacterial flagellar axial structure and its construction. Biophysical Reviews 10:559–570. **doi:**10.1007/s12551-017-0378-z
11. Okinawa Institute of Science and Technology (2019) The Flagellar Hook: Making Sense of Bacterial Motility. www.oist.jp/news-center/press-releases/flagellar-hook-making-sense-bacterial-motility
12. DeRosier DJ (1998) The Turn of the Screw: The Bacterial Flagellar Motor. Cell 93:17–20. **doi:**10.1016/s0092-8674(00)81141-1
13. Fitzgerald DM, Bonocora RP, Wade JT (2014) Comprehensive Mapping of the *Escherichia coli* Flagellar Regulatory Network. PLoS Genet 10, 10:1–18. **doi:**10.1371/journal.pgen.1004649
14. Cohen EJ, Ferreira JL, Ladinsky MS, Beeby M, Hughes KT (2017) Nanoscale-length control of the flagellar driveshaft requires hitting the tethered outer membrane. Science 356 (April 14, 2017), 197–200. **doi:**10.1126/science.aam6512
15. Shibata S, Matsunami H, Aizawa S-I, Wolf M (2019) Torque transmission mechanism of the curved bacterial flagellar hook revealed by cryo-EM. Nat Struct Mol Biol 26:941–945. **doi:**10.1038/s41594-019-0301-3
16. Kato T, Makino F, Miyata T, Horvath P, Namba K (2019) Structure of the native supercoiled flagellar hook as a universal joint. Nature Communications 10, 5295. **doi:**10.1101/748384
17. Auletta G (2013) Information and Metabolism in Bacterial Chemotaxis. Entropy 2013:311–326. **doi:**10.3390/e15010311
18. Krell T, Lacal J, Muñoz-Martinez F, Reyes-Darias JA, Cadirci BH, Garcia-Fontana C, Ramos JL (2011) Diversity at its best: bacterial taxis. Environmental Microbiology 13:1115–1124. **doi:**10.1111/j.1462-2920.2010.02383.x
19. Bren A, Eisenbach M (2000) How Signals Are Heard during Bacterial Chemotaxis: Protein-Protein Interactions in Sensory Signal Propagation. Journal of Bacteriology, Dec. 2000, 6865–6873. **doi:**10.1128/jb.182.24.6865-6873.2000

20. Sarkar MK, Koushik P, Blair D (2010) Chemotaxis signalling protein CheY binds to the rotor protein FliN to control the direction of flagellar rotation in *Escherichia coli*. PNAS 107, 20 (May 18, 2010):9370–9375. **doi:**10.1073/pnas.1000935107
21. Parkinson Lab, Department of Biology, University of Utah (2019) An overview of *E. coli* chemotaxis. http://chemotaxis.biology.utah.edu/Parkinson_Lab/projects/ecolichemotaxis/ecolichemotaxis.html
22. Sourjik V, Wingreen NS (2012) Responding to Chemical Gradients: Bacterial Chemotaxis. Current Opinion in Cell Biology 24:262–268. **doi:**10.1016/j.ceb.2011.11.008
23. Rebbapragada A, Johnson MS, Harding GP, Zuiccarelli AJ, Fletcher HM, Zhulin IB, Taylor BL (1997) The Aer protein and the serine chemoreceptor Tsar independently sense intracellular energy levels and transduce oxygen, redox, and energy signals for *Escherichia coli* behavior. Proc. Natl. Acad. Sci. USA 94:10541–10546. **doi:**10.1073/pnas.94.20.10541
24. Berg HC (2003) Bacterial Flagellar Motor. Current Biology 18, 16:689–691. **doi:**10.1016/j.cub.2008.07.015
25. Nord AL, Gachon E, Perez-Carrasco R, Nirody JA, Barducci A, Berry RM, Pedaci F (2017) Catch bond drives stator mechanosensitivity in the bacterial flagellar motor. PNAS 114, 49 (December 5, 2017):12952–12957. https://www.pnas.org/content/pnas/114/49/12952.full.pdf
26. Fan B, Lo C-J, Berry RM, Xing J (2009) Model Studies of the Dynamics of Bacterial Flagellar Motors. Biophys J 96:3154–3167. **doi:**10.1016/j.bpj.2009.01.023
27. Subramanian S, Gao X, Dann CE, Kearns DB (2017) MotI (DgrA) acts as a molecular clutch on the flagellar stator protein MotA in *Bacillus subtilis*. PNAS 114, 51 (December 19, 2017):13537–13542. **doi:**10.1073/pnas.1716231114
28. Namba, K (2020) A proposed gear mechanism for torque generation in the flagellar motor. Nat Struct Mol Biol 27:1004–1006. **doi:**10.1038/s41594-020-00514-0
29. SBGN (2019) Systems Biology Graphical Notation: Process Description Language Level 1: Systems Biology Graphical Notation. https://sbgn.github.io/downloads/specifications/pd_level1_version2.pdf
30. Brown J, Faulds-Pain A, Aldridge P (2009) Chapter 7: The Coordination of Flagellar Gene Expression and the Flagellar Assembly Pathway. In: Pili and Flagella: Current Research and Future Trends. Caister Academic Press (Poole, UK) 99–120.
31. Chevance FFV, Hughes KT (2008) Coordinating assembly of a bacterial macromolecular machine. Nature Reviews Microbiology 6:455–465. **doi:**10.1038/nrmicro1887
32. Kalir S, McClure J, Pabbaraju K, Southward M, Ronen M, Leibler S, Surette MG, Alon U (2001) Ordering Genes in a Flagella Pathway by Analysis of Expression Kinetics from Living Bacteria. Science 292, 5524 (June 15, 2001): 2080–2083. **doi:**10.1126/science.1058758
33. Karlinsey JE et al. (2000) Completion of the hook-basal body complex of the *Salmonella typhimurium* flagellum is coupled to FlgM secretion and *FliC* transcription. Molecular Microbiology 37(5):1220–1231. **doi:**10.1046/j.1365-2958.2000.02081.x
34. Schumacher JS, Thormann KM, Bange G (2015) How bacteria maintain location and number of flagella? FEMS Microbiology Reviews 39:812–822. **doi:**10.1093/femsre/fuv034
35. Ozin AJ, Claret L, Auvray F, Hughes C (2003) The FliS chaperone selectively binds the disordered flagellin C-terminal D0 domain central to polymerization. FEMS Microbiology Letters 219:219–224. **doi:**10.1016/s0378-1097(02)01208-9
36. Blair DF (2009) Chapter 8: Structure and Mechanism of the Flagellar Rotary Motor. In: Pili and Flagella: Current Research and Future Trends. Caister Academic Press (Poole, UK), 121–136.