Molecular characterization of the flagellar hook in *Bacillus subtilis*

Colleen R. Courtney1,2, Loralyn M. Cozy1,3, and Daniel B. Kearns*1

1Indiana University
Department of Biology
Bloomington, IN 47408

2Current address:
New York University
Sackler Institute of Microbiology-Parasitology
New York, NY 10016

3Current address:
University of Hawaii
Department of Microbiology
Honolulu, HI 96822

*Corresponding author
Email: dbkearns@indiana.edu
Phone: 1-812-856-2523
Fax: 1-812-855-6705

Running title: The *B. subtilis* flagellar hook.

Keywords: FlgE, hook, motility, flagella, FliK
The structure of the Gram-positive flagellum is poorly understood and *Bacillus subtilis* encodes three proteins homologous to the flagellar hook protein from *Salmonella enterica*. Here we generated a modified *B. subtilis* hook protein that could be fluorescently stained using a cysteine reactive dye. We used the fluorescently labeled hook to demonstrate that FlgE is the hook structural protein and that FliK regulated hook length. We further demonstrate that two proteins of unknown function FlhO and FlhP, and the putative hook cap FlgD, were required for hook assembly, such that when *flhO*, *flhP*, or *flgD* were mutated, hook protein was secreted into the supernatant. All mutants defective in hook completion resulted in homogeneously reduced σ^D^-dependent gene expression due to the action of the anti-sigma factor FlgM.
INTRODUCTION

Many bacteria are motile by rotating extracellular appendages called flagella but flagellar assembly and structure is best understood in the Gram-negative bacteria *Escherichia coli* and *Salmonella enterica* (45). Thrust is generated by rotating the filament, a long, hollow, helical structure polymerized from approximately 20,000 subunits of a single protein called flagellin (aka FliC, Hag) (44). The filament is assembled by secreting each flagellin subunit through the duct of the nascent structure such that polymerization occurs at the distal tip (31, 73). Flagellin secretion is driven by a type III protein secretion system housed within the flagellar basal body that is anchored to the cell envelope (16, 48). The basal body also interacts with proton channels to rotate a rod that ultimately turns the filament. Between the rotating rod and the helical filament is a short, curved, hollow, linker domain called the hook.

The hook is a flexible universal joint that transmits torque from the rod to the filament and changes the angle of rotation (6, 64). In *E. coli* and *S. enterica*, the hook is assembled from a single repeating monomer unit called FlgE (32, 35, 40, 66). The *flgE* gene (aka *flaK*, *flaFV*) that encodes the FlgE protein was identified by partial function alleles that altered the electrophoretic mobility of the hook structural subunit (1, 38). To build the hook, the hook cap protein FlgD must first be assembled on the end of the completed rod structure (59). Then, like flagellin, FlgE subunits are secreted through the basal body rod complex and inserted underneath FlgD and polymerized into a curved hollow cylinder (56, 59). The hook has an average length of 55 nm and hook length is controlled by the regulatory protein FliK and the flagellar secretion component FlhB.
Loss of FliK and particular alleles of FlhB result in elongated hook structures called polyhooks and the cells fail to produce a flagellar filament (27, 63, 67).

FliK is thought to act as a molecular ruler in which the length of the FliK primary sequence is proportional to the length of the hook (17, 34, 55). Prior to hook completion, FliK is secreted intermittently through the basal-body rod and the nascent hook structure (17, 48, 49). When the hook-basal body is the length of the extended FliK protein, the FliK N-terminus interacts with the hook cap, the middle section occupies the secretion channel, and the FliK C-terminus interacts with FlhB in the cytoplasm (34, 51, 52, 55). Interaction with FliK activates autoproteolysis of FlhB to cause a switch in the substrate specificity of the secretion apparatus (19, 20, 52, 71). After specificity has been switched, the secretion apparatus becomes proficient for the secretion of late class flagellar proteins.

One of the earliest proteins secreted after the substrate specificity switch is the negative regulator FlgM (23). FlgM is an anti-sigma factor that directly binds to and antagonizes the activity of the alternative sigma factor σ²⁸ (aka σ²⁸F, FliA) (33, 58). Prior to completion of the hook-basal body complex, FlgM accumulates in the cytoplasm and inhibits σ²⁸ activity. Once the hook is complete, FlgM is secreted and σ²⁸ is released to direct RNA polymerase to express the late class flagellar genes including *fliC*, encoding flagellin (30, 36, 41). Thus the length of the completed flagellar hook is not only regulated, but hook completion has profound regulatory effects on downstream gene expression and subsequent flagellar assembly.

The Gram-positive bacterium *Bacillus subtilis* is motile and encodes many flagellar structural proteins in the 31-gene fla/che operon but their putative functions are
largely inferred by homology to their *E. coli* and *S. enterica* counterparts (2, 3, 11, 60, 74). *B. subtilis* also encodes FlgM that binds to and inhibits the σ28 homolog σD (SigD) that directs flagellin expression (7, 10, 21, 54). In addition, σD activity is inhibited in the absence of completed basal bodies in a FlgM-dependent manner (5, 14, 29). FlgM has never been reported to be secreted in *B. subtilis* and how FlgM activity is coordinated with flagellar assembly is unknown. Finally, the *B. subtilis* genome contains a gene annotated as *fliK*, that is predicted to encode a FliK homolog, but the role of the *B. subtilis* FliK protein in hook-length regulation has not been explored.

*B. subtilis* encodes three homologs of the hook structural protein: FlgE, FlhO, and FlhP. The FlgE protein was predicted to constitute the primary structural subunit of the hook due to its high abundance in biochemical preparations of purified *B. subtilis* hook-basal body complexes (39). By contrast, FlhO and FlhP were present in much lower abundance and their functions are unknown (39). Whereas *flgE* gene is coexpressed with other hook-basal proteins in the *fla/che* operon, *flhO* and *flhP* are expressed as a separate and remote putative dicistron. Here we mutate each of the three genes encoding hook homologs and find that each is required for motility. Using genetic, biochemical, and cytological approaches, we conclude that FlgE is the primary structural subunit of the hook and demonstrate that FlhO and FlhP are required for hook assembly (39). We demonstrate that cells mutated for *fliK* produce polyhook structures and that cells mutated for the hook reduce σD-dependent gene expression in a FlgM-dependent manner.
MATERIALS AND METHODS

**Strains and growth conditions:** All strains used in this study are listed in Table 1. *B. subtilis* strains were grown in Luria-Bertani (LB) (10 g tryptone, 5 g yeast extract, 5 g NaCl per L) broth or on LB plates fortified with 1.5% Bacto agar at 37°C. When appropriate, antibiotics were included at the following concentrations: 10 µg/ml tetracycline, 100 µg/ml spectinomycin, 5 µg/ml chloramphenicol, 5 µg/ml kanamycin, and 1 µg/ml erythromycin plus 25 µg/ml lincomycin (MLS).

**Swarm expansion assay.** Cells were grown to mid-log phase at 37°C in LB broth and resuspended to 10 OD\(_{600}\) in pH 8.0 PBS buffer (137 mM NaCl, 2.7 mM KCl, 10 mM Na\(_2\)HPO\(_4\), and 2 mM KH\(_2\)PO\(_4\)) containing 0.5% India ink (Higgins). Freshly prepared LB containing 0.7% Bacto agar (25 ml/plate) was dried for 20 minutes in a laminar flow hood, centrally inoculated with 10 µl of the cell suspension, dried for another 10 minutes, and incubated at 37°C. The India ink demarks the origin of the colony and the swarm radius was measured relative to the origin. For consistency, an axis was drawn on the back of the plate and swarm radii measurements were taken along this transect.

**Microscopy.** For CFP and YFP microscopy, cells were grown at 37°C to OD\(_{600}\) 0.6-1.0, and 1 ml was washed once in pH 8.0 PBS buffer, pelleted, and resuspended in 50µl PBS buffer containing 5 µg/ml membrane strain FM 4-64 (Molecular Probes). For fluorescent microscopy of flagella, 0.5 ml of broth culture was harvested at 0.5-2.0 OD\(_{600}\), and washed once in 1.0 ml PBS buffer. The suspension was pelleted, resuspended in 50 µl of PBS buffer containing 5 µg/ml Alexa Fluor 488 C\(_5\) maleimide (Molecular Probes), and incubated for 5 min at room temperature (8). Cells were then washed twice...
with 500 µl PBS buffer and membranes were stained by resuspension in 50 µl of PBS buffer containing 5 µg/ml FM4-64. Three microliters of suspension were placed on a microscope slide and immobilized with a poly-L-lysine-treated coverslip.

Fluorescence microscopy was performed with a Nikon 80i microscope with a phase contrast objective Nikon Plan Apo 100X and an Excite 120 metal halide lamp. FM4-64 was visualized with a C-FL HYQ Texas Red Filter Cube (excitation filter 532-587 nm, barrier filter >590 nm). CFP fluorescent signals were viewed using a C-FL HYQ CFP Filter Cube (excitation filter 426-446 nm, barrier filter 460-500 nm). YFP was visualized using a C-FL HYQ YFP Filter Cube (excitation filter 490-510 nm, barrier filter 520-550 nm). Images were captured with a Photometrics Coolsnap HQ2 camera in black and white, false colored and superimposed using Metamorph image software.

**Strain construction.** All PCR products were amplified from purified chromosomal DNA from *B. subtilis* strain 3610. All constructs were first introduced into the domesticated strain PY79 by natural competence and then transferred to the 3610 background using SPP1 phage-mediated generalized phage transduction (72). All primers used in this study are listed in Supplemental Table S1. All plasmids used in this study are listed in Supplemental Table S2.

**In-frame deletions.** To generate the Δ*flgE* in-frame marker-less deletion construct pDP306, the region upstream of *flgE* was PCR amplified using the primer pair 1483/1484 and digested with EcoRI and XhoI, and the region downstream of *flgE* was PCR amplified using the primer pair 1485/1486 and digested with *XhoI* and *BamHI*. The two fragments were then simultaneously ligated into the EcoRI and BamHI sites of pMiniMAD which carries a temperature sensitive origin of replication and an
erythromycin resistance cassette (62). The plasmid pDP306 was introduced to PY79 by single cross-over integration by transformation at the restrictive temperature for plasmid replication (37°C) using mls resistance as a selection. The integrated plasmid was then transduced into 3610. To evict the plasmid, the strain was incubated in 3ml LB broth at a permissive temperature for plasmid replication (22°C) for 14 hours, diluted 30-fold in fresh LB broth, and incubated at 22°C for another 8 hours. Dilution and outgrowth was repeated 2 more times. Cells were then serially diluted and plated on LB agar at 37°C. Individual colonies were patched on LB plates and LB plates containing mls to identify mls sensitive colonies that had evicted the plasmid. Chromosomal DNA from colonies that had excised the plasmid was purified and screened by PCR using primers 1483/1486 to determine which isolate had retained the ΔflgE allele.

To generate the ΔflhO in-frame marker-less deletion construct pKB114, the region upstream of flhO was PCR amplified using the primer pair 1850/1524 and digested with EcoRI and SalI, and the region downstream of fliK was PCR amplified using the primer pair 1525/1526 and digested with SalI and BamHI. The two fragments were then simultaneously ligated into the EcoRI and BamHI sites of pMiniMAD. pKB114 was integrated into B. subtilis PY79 genome, transduced to strain 3610, and evicted as described above. Colonies were screened by PCR using primers 1850/1524 to determine which isolate had retained the ΔflhO allele.

To generate the ΔflhP in-frame marker-less deletion construct pCC12, the region upstream of flhP was PCR amplified using the primer pair 2347/2348 and digested with EcoRI and BamHI, and the region downstream of fliK was PCR amplified using the primer pair 2349/2350 and digested with BamHI and SalI. The two fragments were then
simultaneously ligated into the EcoRI and SalI sites of pMiniMAD. pCC12 was integrated into B. subtilis PY79 genome, transduced to strain 3610, and evicted as described above. Colonies were screened by PCR using primers 2347/2350 to determine which isolate had retained the ΔflhP allele.

To generate the ΔfliK in-frame marker-less deletion construct pKB93, the region upstream of fliK was PCR amplified using the primer pair 1387/1388 and digested with EcoRI and XhoI, and the region downstream of fliK was PCR amplified using the primer pair 1389/1390 and digested with SalI and BamHI. The two fragments were then simultaneously ligated into the EcoRI and BamHI sites of pMiniMAD (XhoI and SalI have compatible ends). pKB93 was integrated into B. subtilis PY79 genome, transduced to strain 3610, and evicted as described above. Colonies were screened by PCR using primers 1387/1390 to determine which isolate had retained the ΔfliK allele.

To generate the ΔflgD in-frame marker-less deletion construct pDP328, the region upstream of flgD was PCR amplified using the primer pair 2033/2034 and digested with EcoRI and XhoI, and the region downstream of flgD was PCR amplified using the primer pair 2035/2036 and digested with XhoI and BamHI. The two fragments were then simultaneously ligated into the EcoRI and BamHI sites of pMiniMAD. pDP328 was integrated into B. subtilis PY79 genome, transduced to strain 3610, and evicted as described above. Colonies were screened by PCR using primers 2033/2036 to determine which isolate had retained the ΔflgD allele.

ΔflhP::kan. The ΔflhP::kan insertion deletion allele was generated by long flanking homology PCR (using primers 624 and 625, 626 and 627), and DNA containing
a kanamycin drug resistance gene (pDG780) was used as a template for marker replacement (24, 69).

**Complementation constructs.** To generate the $P_{DsP_{fla/che}}flgE$ complementation construct pDP324, the $P_{DsP_{fla/che}}$ promoter was amplified using primer pair 2015/2016 and was digested with *EcoRI* and *XhoI*. The *flgE* region was amplified primer pair 2017/2018 and was digested with *XhoI* and *BamHI*. The two fragments were ligated simultaneously into the *EcoRI* and *BamHI* sites of pDG364 containing the polylinker and the chloramphenicol resistance cassette between the arms of the *amyE* gene (25). To generate the $P_{DsP_{fla/che}}flgE^{T123C}$ allele complementation construct pCC16, site directed mutagenesis was conducted using the Quickchange II kit (Stratagene) on pDP324 in two sequential steps. First, primer pair 2246/2247 was used to change codon 123 of *flgE* from ACT (threonine) to TCT. Next, primer pair 2378/2379 was used to change the TCT codon to TGT (cysteine). Sequences were verified by sequencing with primers 1853/1854.

To generate the $P_{DsP_{fla/che}}fliK$ complementation construct pCC14, the $P_{DsP_{fla/che}}$ promoter was amplified using primer pair 2015/2016 and was digested with *EcoRI* and *XhoI*. The *fliK* region was amplified using primer pair 2535/2536 and was digested with *XhoI* and *BamHI*. The two fragments were ligated simultaneously into the *EcoRI* and *BamHI* sites of pDG364.

To generate the $P_{flhO}flhO$ complementation construct pCC7, the *flhO* gene and DNA immediately upstream was amplified using primer pair 1251/1681. The PCR product was digested with *EcoRI* and *BamHI* and cloned into the *EcoRI* and *BamHI* sites.
of pDG1664 containing the polylinker and the erythromycin resistance cassette between
the arms of the *thrC* gene (25).

To generate the *P*flhO-flhP complementation construct pCC13, the *P*flhO promoter
was PCR amplified using primer pair 1251/2344, and the *flhP* gene was amplified using
primer pair 2344/2345. The *P*flhO containing PCR product was digested with *EcoRI* and
*XhoI*, the *flhP* containing product was digested with *XhoI* and *BamHI*, and the two
digested fragments were cloned simultaneously into the *EcoRI* and *BamHI* sites of
pDG364.

To generate the *P*D3Pfla/che-flgD complementation construct pDP403, the *PD3Pfla/che*
 promoter was amplified using primer pair 2460/2461 and was digested with *EcoRI* and
*XhoI*. The *flgD* region was amplified using primer pair 3163/3164 and was digested with
*XhoI* and *BamHI*. The two fragments were ligated simultaneously into the *EcoRI* and
*BamHI* sites of pDG364.

**LacZ reporter constructs:** To generate the β-galactosidase (*lacZ*) reporter
construct pCC1, the *P*flhO promoter region was PCR amplified using 3610 DNA as a
template and primer pair 1251/1252. The PCR product was digested with *EcoRI* and
*BamHI* and cloned into the *EcoRI* and *BamHI* sites of plasmid pDG268, which carries a
chloramphenicol-resistance marker and a polylinker upstream of the *lacZ* gene between
two arms of the *amyE* gene (4).

To generate the native site *lacZ* integrant *flhPΩlacZ* cat pDP405, the *flhP* gene
was PCR amplified using 3610 DNA as a template and primer pair 3210/3211. The PCR
product was digested with *EcoRI* and *BamHI* and cloned into the *EcoRI* and *BamHI* sites
of plasmid pEX44 which carries chloramphenicol-resistance marker and a polylinker upstream of the lacZ gene.

**MBP-FlgE protein expression construct:** To generate the MBP-FlgE fusion expression construct, the flgE gene was PCR amplified using primer pair 1992/1993, digested with EcoRI and BamHI and cloned into the EcoRI and BamHI sites of plasmid pMal-p2X (New England Biolabs)

**SPP1 phage transduction.** To 0.2 ml of dense culture grown in TY broth (LB broth supplemented after autoclaving with 10 mM MgSO_4 and 100 µM MnSO_4), serial dilutions of SPP1 phage stock were added and statically incubated for 15 minutes at 37°C. To each mixture, 3 ml TYSA (molten TY supplemented with 0.5% agar) was added, poured atop fresh TY plates, and incubated at 37°C overnight. Top agar from the plate containing near confluent plaques was harvested by scraping into a 50 ml conical tube, vortexed, and centrifuged at 5,000 x g for 10 minutes. The supernatant was treated with 25 µg/ml DNase final concentration before being passed through a 0.45 µm syringe filter and stored at 4°C.

Recipient cells were grown to stationary phase in 2 ml TY broth at 37°C. 0.9 ml cells were mixed with 5 µl of SPP1 donor phage stock. 9 ml of TY broth was added to the mixture and allowed to stand at 37°C for 30 minutes. The transduction mixture was then centrifuged at 5,000 x g for 10 minutes, the supernatant was discarded and the pellet was resuspended in the remaining volume. 100 µl of cell suspension was then plated on TY fortified with 1.5% agar, the appropriate antibiotic, and 10 mM sodium citrate.

**β-Galactosidase assay.** Cells were harvested from cultures growing at 37°C in LB broth. Cells were collected in 1 ml aliquots and suspended in equal volume of Z
buffer (40 mM NaH$_2$PO$_4$, 60 mM Na$_2$HPO$_4$, 1 mM MgSO$_4$, 10 mM KCl and 38 mM 2-mercaptoethanol). Lysozyme was added to each sample to a final concentration of 0.2 mg ml$^{-1}$ and incubated at 30°C for 15 min. Each sample was diluted in Z-buffer to a final volume of 500 µl and the reaction was started with 100 µl of 4 mg ml$^{-1}$ 2-nitrophenyl β-D-galactopyranoside in Z buffer and stopped with 250 µl of 1 M Na$_2$CO$_3$. The OD$_{420}$ of the reaction mixture was measured and the β-galactosidase-specific activity was calculated according to the equation: [OD$_{420}$/ (time x OD$_{600}$)] x dilution factor x 1000.

**FlgE protein purification.** The MBP-FlgE fusion protein expression vector pCC10 was transformed into *Rosetta GamE. coli*, grown to ~0.8 OD$_{600}$ in 750 ml of Luria-Bertani broth supplemented with 0.2% Glucose and 100µg/ml ampicillin, induced with 0.2 mM IPTG and grown at 16°C for 12 hours. Cells were pelleted and resuspended in Column Buffer (20mM Tris-HCl, 400mM NaCl, 1mM EDTA, and 0.2mM PMSF), disrupted on a French Press and lysed by sonication. Lysed cells were ultracentrifuged at 20,000 x g for 40 min. Cleared supernatant was combined with amylose resin (New England Biolabs) and incubated for overnight with gentle rocking at 4°C. The resin/lysate mixture was poured onto a 1 cm separation column (Bio-Rad), the resin were allowed to pack and was washed with Column Buffer. MBP-FlgE fusion protein bound to the resin was then eluted using Column buffer supplemented with 10mM Maltose. Elutions were separated by SDS-PAGE and Coomassie stained to verify purification of the MBP-FlgE fusion, and pure fractions were dialyzed into Dialysis buffer (300mm NaCl, 0.5mM EDTA, 10% Glycerol, 50mM Na$_3$PO$_4$) and stored at -20°C.

**FlgE antibody preparation.** 1 mg of purified MBP-FlgE protein was sent to Cocalico Biologicals Inc. for serial injection into a rabbit host for antibody generation.
Crude serum was sufficient for high-affinity, high-specificity detection of FlgE protein in Western blot analysis.

**Western blotting.** *B. subtilis* strains were grown in LB to OD$_{600}$ ~1.0. 1 ml was harvested by centrifugation, and resuspended to 10 OD$_{600}$ in Lysis buffer (20 mM Tris pH 7.0, 10 mM EDTA, 1 mg/ml lysozyme, 10 μg/ml DNAse I, 100 μg/ml RNAse I, 1 mM PMSF) and incubated 30 minutes at 37°C. 10 μl of lysate was mixed with 2 μl 6x SDS loading dye. Samples were separated by 12% Sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE). The proteins were electroblotted onto nitrocellulose and developed with anti-FlgE (1:20,000 dilution), anti-SigA (1:40,000 dilution), or anti-Hag (1:80,000 dilution) of primary antibody and a 1:10,000 dilution secondary antibody (horseradish peroxidase-conjugated goat anti-rabbit immunoglobulin G). Immunoblot was developed using the Immun-Star HRP developer kit (Bio-Rad).
RESULTS

Mutants defective in FlgE, FlhO, and FlhP fail to assemble the flagellar filament and are reduced for flagellin gene expression. The molecular composition of the flagellar hook is unknown in B. subtilis. The B. subtilis genome encodes three proteins, FlgE, FlhO, and FlhP, homologous to the single flagellar hook protein, FlgE, in S. enterica (Fig. 1A). To determine the contribution of each protein to flagellar-based motility, the gene encoding each homolog, flgE, flhO, and flhP respectively, was separately mutated by in-frame marker-less deletion (Fig. 1B). Mutation of any of the three genes abolished swimming in wet-mounts as observed by phase contrast microscopy, and abolished swarming motility, defined as the ability to spread rapidly atop the surface of a 0.7% agar Petri plate (Fig. 1C) (37). We conclude that FlgE, FlhO, and FlhP are required for motility.

To confirm that the loss-of-motility phenotypes resulted from the absence of the putative hook proteins, each individual gene was restored by complementation (Fig. 1B). To complement the flgE mutation, the flgE gene was fused to the P_D-3Pfla/che promoter of the fla/che operon and inserted at the ectopic amyE site (amyE::PD-3Pfla/che-flgE) (18, 70). To complement the flhO mutation, the flhO gene and 500 base pairs immediately upstream (P_flhO) was inserted at the amyE site (amyE::P_flhO-flhO). The flhP gene appeared to be part of an operon with flhO, and to complement the flhP mutation, the flhP gene was fused to the P_flhO promoter region and inserted at the amyE site (amyE::P_flhO-flhP). In each case, introduction of the complementation construct restored wild type swarming motility to the strain mutated for the corresponding gene (Fig. 1C). We conclude that each mutation resulted in a loss of motility due to the disruption of the
indicated gene and that the phenotype was not due to polar effects on downstream gene expression. We further infer that the region upstream of the \( flhO \) gene contains a promoter sufficient for expressing the \( flhO \) and \( flhP \) genes.

In \textit{E. coli} and \textit{S. enterica}, completion of the flagellar hook is required to assemble the flagellar filament. To determine whether mutation of the putative hook proteins affected filament assembly, flagellar filaments were stained with a cysteine-reactive dye in a strain expressing the flagellar filament protein Hag with a cysteine residue introduced on an exposed surface (\( amyE::\text{Phag-hag}^{T209C} \)) (8). Wild type cells expressing the modified filament produced many flagella on the cell surface (Fig. 2A). In contrast, cells mutated for \( \text{flgE, flhO, or flhP} \) failed to synthesize detectable filaments (Fig. 2A).

We conclude that FlgE, FlhO, and FlhP are each required for filament assembly as mutation of the genes that encodes each protein abolished filament assembly.

The failure to synthesize flagellar filaments may have either been due to a defect in flagellin protein synthesis or due to a defect in flagellin polymerization such that flagellin subunits were released into the supernatant. To measure cell-associated flagellin protein levels, cell pellets of the wild type and various mutants were harvested, resolved by SDS-PAGE, electroblotted, and probed with an anti-Hag antibody in Western blot analysis. Each lysate was also probed with an anti-SigA antibody to detect the constitutive vegetative housekeeping sigma factor \( \sigma^V \) as a loading control. Whereas flagellin levels were high in the wild type, flagellin was severely reduced in each of the \( \text{flgE, flhO, and flhP} \) backgrounds (Fig 3A). To measure the amount of secreted flagellin, cell supernatants were concentrated by TCA precipitation, resolved by SDS-PAGE, electroblotted, and probed with an anti-Hag antibody in Western blot analysis. Low
amounts of flagellin were found in the supernatant of wild type cells, presumably the result of secretion of unpolymerized flagellin and/or filament shearing (Fig. 3A). The flagellin found in the supernatant was unlikely the result of premature cell lysis as the cytoplasmic σ^A protein was undetectable (Fig. 3B). Finally, no flagellin was detected in the supernatants of the flgE, flhO, or flhP mutants suggesting that the reduction of flagellin in cell pellets was not due to secretion of flagellin into the supernatant (Fig 3A). We conclude that mutation of any of the putative hook components results in a dramatic reduction in total flagellin protein level.

Flagellin protein levels may be reduced due to a reduction in flagellin gene expression. To measure flagellin gene expression, the promoter of flagellin (Phag) was fused to the gene encoding yellow fluorescent protein and inserted at an ectopic site (amyE::Phag-YFP). Cells mutated for flgE, flhO, or flhP appeared to express YFP at a level lower than that of wild type and the cells grew in short chains (Fig. 4). Chaining may result from a decreased expression of the LytF cell-separating autolysin, which like flagellin, is expressed by the alternative sigma factor σ^D (12, 46). To measure LytF expression, the promoter of LytF (P_{lytF}) was fused to the gene encoding cyan fluorescent protein and inserted at an ectopic site (thrC::P_{lytF}-CFP). Cells mutated for flgE, flhO, and flhP appeared to express CFP from P_{lytF} at an undetectable level (Fig. 4). We conclude that mutation of the putative hook proteins reduces the expression of both the flagellin and LytF promoters uniformly in the population and likely acts at the level of the σ^D sigma factor.

The sigma factor σ^D is known to be antagonized by binding directly to the anti-sigma factor FlgM (7). To determine whether FlgM was antagonizing σ^D in the flgE,
flhO, and flhP mutants, σD-dependent gene expression was measured quantitatively by generating transcriptional fusions of the gene encoding β-galactosidase fused to either the Phag or PlytF promoter. Mutation of flgE, flhO, or flhP reduced the expression of both Phag and PlytF 10-fold relative to wild type (Fig. 5). Furthermore, expression was restored to levels above that of the wild type by the deletion of the gene encoding FlgM (Fig. 5). We conclude that σD-dependent gene expression (including flagellin and LytF expression) is low in the putative hook mutants due to the activity of the anti-sigma factor FlgM.

The FlgE protein is the structural subunit of the hook and FlhO, FlhP, and FlgD are required for hook assembly. To dissect the roles of the FlgE, FlhO, and FlhP proteins, we generated a strain to visualize the flagellar hook by introducing a cysteine residue on an exposed surface of a hook protein such that the hook could be labeled with fluorescent cysteine-reactive dye. We chose the FlgE protein for modification as previous work suggested that FlgE was substantially more abundant than either FlhO or FlhP in flagellar hook preparations (39). To identify candidate residues, the primary sequences of FlgE from B. subtilis and S. enterica were aligned, and surface-exposed serine and threonine residues were predicted by mapping on to the S. enterica FlgE three-dimensional structure (22, 64). Threonine at position 123 was predicted to be surface exposed and site-directed mutagenesis was used to replace the residue with a cysteine (FlgE^{T123C}) in the flgE complementation construct (amyE::P_{D-α}P_{flu/chf}flgE^{T123C}) (Fig. 1A). The ectopically integrated flgE^{T123C} complementation construct was functional and restored swarming motility to a ΔflgE deletion mutation (Fig. 1C). Furthermore, fluorescence microscopy of the flgE^{T123} allele introduced as a merodiploid to wild type cells resulted in
409 a faint punctate staining pattern when the cysteine-reactive dye was added (Fig. 2B).
410 Stronger fluorescence was detected in a ΔflgE mutant complemented by the ectopic
411 \( flgE^{T123C} \) presumably due to the lack of competition between the stainable and wild type
412 allele. We conclude that FlgE\(^{T123C} \) was functional for supporting motility and could be
413 fluorescently stained.
414 The puncta observed when staining the FlgE\(^{T123C} \) allele may represent flagellar
415 hooks but hooks are short and the curved shape is below the limit of resolution for
416 fluorescence microscopy. To increase the size the flagellar hooks, we mutated the
417 putative hook length regulator FliK because \( S. enterica fliK \) mutants fail to limit hook
418 length resulting in elongated “polyhooks” (27). An in-frame marker-less deletion of \( B. \)
419 subtilis fliK (Fig. 1B) abolished swimming and swarming motility (Fig 1C), abolished
420 flagellar filament assembly (Fig. 2A), reduced flagellin protein levels (Fig 3A), and
421 reduced σ\(^D \)-dependent gene expression in a FlgM-dependent manner (Figs. 4 and 5).
422 Thus, mutation of fliK appeared to phenocopy mutation of flgE, flhO, or flhP. When the
423 FlgE\(^{T123C} \) allele was expressed and fluorescently labeled in a ΔfliK ΔflgE double mutant
424 background, elongated structures consistent with polyhooks were observed (Fig. 2B, Fig.
425 S1). We conclude that FlgE\(^{T123C} \) puncta represent flagellar hooks, that FlgE is a primary
426 component of the hook structure, and FliK in \( B. subtilis \) is a hook length regulator.
427 To determine the roles of FlhO and FlhP in hook assembly, the FlgE\(^{T123C} \) allele
428 was introduced to ΔflhO ΔflgE or ΔflhP ΔflgE double mutants. No puncta were observed
429 when either strain was fluorescently labeled suggesting that both FlhO and FlhP were
430 required for assembly of the hook (Fig 2B). One way in which assembly of the hook
431 could be abrogated is by a reduction in the amount of FlgE protein synthesized. To
measure FlgE protein levels, cell pellets of the wild type and various mutants were harvested. Lysates from each pellet were resolved by SDS-PAGE, electroblotted, and probed with an anti-FlgE antibody in Western blot analysis. Whereas FlgE levels were high in the wild type and fliK mutant pellets, FlgE was severely reduced in the flhO and flhP mutant backgrounds (Fig 3C). Instead, FlgE was found in high levels secreted into the supernatant in the flhO and flhP mutants relative the other backgrounds (Fig 3C). We conclude that FlgE is secreted but not polymerized in the absence of FlhO and FlhP.

Accumulation of FlgE in the supernatant of cells lacking FlhO and FlhP was reminiscent of the phenotype reported for cells lacking the FlgD hook cap in S. enterica (59). The B. subtilis flgD gene (formerly ylxG) is predicted to encode a homolog of the hook capping protein FlgD from S. enterica and is located immediately upstream of the flgE gene in the B. subtilis fla/che operon (Fig. 1B). An in-frame markerless deletion of flgD abolished motility and was ectopically complemented by inserting the flgD gene fused to the P_{D,3P_{fla/che}} promoter at the ectopic amyE site (amyE:: P_{D,3P_{fla/che}} flgD) (Fig 1B, 1C). Like other mutants defective in hook completion, the flgD mutant grew in short chains (Fig 4), and showed a 10-fold reduction in σ^{D}-dependent gene expression that was dependent on FlgM (Fig 5). Furthermore, the flgD mutant secreted FlgE hook protein abundantly into the supernatant as had been observed with S. enterica flgD mutants (Fig 3) (53). Thus, with respect to FlgE secretion, cells mutated for fliO and fliP closely resembled cells mutated for flgD in B. subtilis. When a flgD flgE double mutant that encoded an ectopically integrated FlgE^{T123C} allele was stained with a maleimide dye and observed by fluorescence microscopy, fluorescence was observed as a haze over the cell with occasional puncta (Fig 2B). If the puncta represent hooks, we infer that the hooks
fail to be functionally complete as the flgD mutant was defective in assembling flagellar filaments (Fig 2A). We conclude that FlgD is required for hook completion.

*FlhO and FlhP are expressed from multiple promoters.* To characterize the regulation of flhO and flhP, the P_{flhO} promoter region was cloned upstream of the lacZ gene encoding β-galactosidase and inserted at an ectopic locus (amyE::P_{flhO}-lacZ). The P_{flhO} promoter region was proficient for driving lacZ gene expression supporting the presence of a functional promoter (Table 2). Some genes required for flagellar assembly are expressed under the control of σ^D, the alternative sigma factor that also expresses flagellin. The P_{flhO} promoter appeared to be σ^D-dependent as mutation of the sigD gene encoding σ^D abolished expression of the P_{flhO}-lacZ reporter (Table 2). Furthermore, P_{flhO} expression increased when the gene encoding FlgM, the anti-sigma factor that antagonizes σ^D activity, was mutated (Table 2). Finally, within the 500 bp region upstream of flhO there was a candidate promoter sequence, TTTA<15 bp spacer>TCCATAT, that differed at four positions (underlined) from the predicted σ^D consensus binding site TAAA<14-16 bp spacer>GCCGATAT (26). We conclude that upstream of flhO and flhP is a promoter, P_{flhO}, that is σ^D-dependent.

If the expression of flhO and flhP is under the control of a σ^D-dependent promoter, and FlhO and FlhP are required for hook assembly, then σ^D might also be required for hook assembly. A cell mutated for the gene encoding σ^D, sigD, grew in long chains due to the failure to express σ^D-dependent autolysins and was defective for filament biosynthesis due to the fact that hag is under exclusive σ^D control (Fig. 2A) (12, 53). When a sigD mutant encoding the FlgE^{T123C} allele was stained for fluorescence microscopy, faint puncta could be detected (Fig. 2B). Likewise, near wild type levels of
FlgE protein were found in the cell pellet of a sigD mutant (Fig 3). We conclude that σ^D is not essential for hook assembly.

If σ^D is not required for hook assembly, then σ^D must not be absolutely required for the expression of FlhO and FlhP. We hypothesized that flhO and flhP must be expressed from one or more additional promoters besides P_{flhO}. To detect additional promoter activity, the lacZ gene was integrated downstream of flhP at the native site in wild type cells and cells mutated for either sigD or flgM. Wild type cells expressed the native site flhPΩlacZ integrant at roughly the same level as the ectopically integrated P_{flhO}-lacZ fusion (Table 2). Likewise, expression of flhPΩlacZ decreased in the sigD mutant and increased in the flgM mutant (Table 2). Unlike the P_{flhO}-lacZ fusion, however, expression of flhPΩlacZ was not abolished in the absence of σ^D, and a low 4 Miller units of activity was detected. We conclude that a σ^D independent promoter is present upstream of flhP and we infer that the low level of expression from this promoter is sufficient to support hook assembly in a sigD mutant. We infer that it is important for cells to be able to complete hook assembly by a σ^D-independent mechanism because hook completion is necessary to activate σ^D by antagonizing FlgM (Fig. 5).

DISCUSSION

The hook is a critical component of the flagellum. The hook 1) connects the filament to the basal body, 2) acts as a universal joint to change the angle of rotation that is transmitted to the filament, 3) serves as a secretion conduit and polymerization platform for filament proteins, 4) instructs a change in specificity of the secretion apparatus when completed, and 5) indirectly controls late class flagellar gene
transcription. Finally, in *B. subtilis*, hook completion has recently been shown to be
required to relieve translation inhibition on the flagellin transcript (57). Despite its
importance, the hook has been relatively poorly studied in *B. subtilis*, and three different
proteins that are homologous to the hook structural subunit of *S. enterica* are encoded in
the *B. subtilis* genome. Here we introduce a unique cysteine residue into the FlgE
primary sequence such that FlgE could be fluorescently labeled: an approach that has
proved successful for labeling the filament (8, 68). We found that FlgE formed hooks
that appeared as puncta when stained for fluorescent microscopy and we support the
inference that FlgE is primary hook structural subunit (39). Furthermore, we used hook
staining to demonstrate that FliK was the *B. subtilis* hook length regulator and that the
FlgD hook cap and the FlgE homologs of unknown function, FlhO and FlhP, were
required for hook assembly.

The gene predicted to encode FliK in *B. subtilis* was shown to control flagellar
hook length because when fliK was mutated, long FlgE-containing structures resembling
polyhooks were observed (Fig 2B, Fig S1). FliK is thought to act like a molecular ruler
during secretion where the length of the FliK primary sequence is proportional to the
length of the hook (34, 65). The hook of *S. enterica* is approximately 55 nm long and the
FliK primary sequence is 405 amino acids (15, 27). The hook of *B. subtilis* is
approximately 71 nm long and the FliK primary sequence is 487 amino acids, 129% and
120% longer respectively than their *S. enterica* counterparts (15, 39). Although the
conservation between FliK homologs is poor, the two proteins nonetheless seem to
function similarly with lengths correlated to that of the hook they control.
FlgD, FlhO, and FlhP, are required for hook polymerization such that in their absence, FlgE is secreted into the supernatant. The role of FlgD is likely similar to that in *S. enterica*, where FlgD is loaded on the end of the rod and ushers FlgE subunits into the extending hook complex. Two models could account for the requirement of FlhO and FlhP (Fig. 6). In the first model, FlhO and FlhP form the distal rod (Fig 6A). The flagellar rod of *S. enterica* is complex and consists of four proteins, FlgB, FlgC, FlgF, and FlgG (28). FlgB and FlgC form the proximal rod with FlgB likely adjacent to the FliF/FliE basal body complex (50). FlgF and FlgG form the distal rod with FlgG likely being the most distal subunit involved in outer membrane penetration (13). In *B. subtilis*, the genes encoding the FlgB and FlgC putative proximal rod proteins are encoded at the 5' end of the long fla/che operon directly adjacent to the genes encoding FliF and FliE (Fig. 1B) (74). FlhO and FlhP could be the distal rod proteins because the hook and rod classes are evolutionarily related and difficult to distinguish at the sequence level (43, 61, 74). Further, FlhO and FlhP mutants may secrete FlgE into the supernatant like a FlgD mutant because in their absence, FlgD might be unable to load (Fig. 6). We note that mutations in the distal rod of *S. enterica* do not accumulate FlgE in the supernatant and instead secrete FlgE to the periplasm where FlgE is proteolytically degraded (9, 42, 59). Thus in *B. subtilis*, mutants defective in the rod and hook cap may resemble one another with respect to FlgE secretion due to the lack of an outer membrane.

In the second model, FlhO and FlhP could function with FlgD to form an extended cap (Fig. 6B). FlgD would first load on to the end of the rod as described for *S. enterica*. We note that FlhO and FlhP could still serve as the distal rod in this model and play a role in FlgD cap loading. If FlhO and FlhP are not part of the distal rod, however,
they would presumably be the first two proteins ushered underneath FlgD. In either case, once the extended cap is assembled, FlgE subunits would be polymerized, not underneath the FlgD protein, but underneath the entire FlgD/FlhO/FlhP complex. The distinction between the distal rod and extended cap models therefore differs in where FlgE is being inserted, either distal or proximal to FlhO and FlhP respectively. We tend to favor a model in which FlhO and FlhP form the distal rod but do not serve as an extended cap as this model is more parsimonious, requires fewer changes in insertion points, is supported by the homology of the proteins involved, and is more consistent with the S. enterica paradigm. We note, that if FlhO and FlhP do not form the distal rod, that no other rod/hook homologs are encoded in the chromosome and therefore the B. subtilis rod may be simpler than the rod assembled by S. enterica.

Although cells mutated for flhO and flhP resembled cells mutated for flgD in nearly all respects, the flgD mutant had one additional phenotype. A cell-associated fluorescent haze with rare puncta was observed when cells encoding the FlgET123C allele at an ectopic site were simultaneously mutated for flgD and flgE and were labeled with a maleimide stain. The fluorescent puncta in this case may either represent non-specific aggregation of FlgE protein on the cell surface or represent partially assembled hooks. If the puncta represent hooks, the hooks presumably remain incomplete due to the failure of the flgD mutant to assemble flagellar filaments (Fig 2A). The haze of fluorescence suggests that, in the absence of FlgD, FlgE was being secreted at a high level and retained non-specifically on the cell surface similar to the surface retention of Yop proteins that occurs during artificial high level secretion in Yersinia (47). Why the flgD mutant would result in expression or secretion of FlgE at greater levels than either the flhO or flhP
mutant is unknown but may suggest that FlgD has an additional regulatory effect on FlgE expression in *B. subtilis*. We note a regulatory connection between the two proteins has been previously reported insomuch as the translation of FlgE protein was impaired by deletions in the *flgD* gene of *S. enterica* (42).

The structural completion of the hook has regulatory consequences. Hook completion is required to antagonize FlgM and thereby activate σ^D^-dependent gene expression. Although *P_{flhO*}, the promoter that expresses FlhO and FlhP, is σ^D^-dependent, and FlhO and FlhP are required for hook synthesis, we found that a *sigD* mutant was nonetheless proficient for forming flagellar hooks (Table 2, Fig 2B). The ability to form hooks in the absence of σ^D^ is likely due to a low level of *flhO* and *flhP* transcription originating from an upstream promoter or incomplete termination of transcript from the upstream gene, *mbl* (Table 2). Sub-optimal expression of FlhO and FlhP may account for the faint FlgET123C puncta and the enhanced secretion of FlgE into the supernatant of the *sigD* mutant (Fig 2A, Fig 3). Furthermore, low levels of FlhO and FlhP may be sufficient for hook completion as the two proteins were present in purified hook-basal bodies at low abundance (39). The role of σ^D^ amplified expression of the *flhO* and *flhP* then may be to create a positive feedback loop for the antagonism of FlgM: the more hooks that are completed, the more FlgM is antagonized, and the more hooks are made. Similarly, we note that a σ^D^-dependent promoter (P_{pylxF3}) was recently identified upstream of the *fliK*, *flgD*, and *flgE* hook genes within the *fla/che* operon (14).

The hook is a critical structural intermediate in flagellar assembly the length of which length is regulated and the completion of which governs subsequent flagellar gene expression. Most flagellar studies have been conducted in Gram-negative bacteria, and
we presume that Gram-positive flagella may require different structures and regulatory systems to accommodate the substantial difference in envelope structure. In particular, the Gram-positive basal body has two fewer rings than its *S. enterica* counterpart as it transits a much thicker layer of peptidoglycan and lacks an outer membrane (15). Although many proteins in flagellar assembly are conserved, conservation is poor, and when dealing with such divergent bacteria, it is important to validate sequence-based predictions with structure-function studies. In total, we assign functional roles to five flagellar proteins previously unstudied in *B. subtilis*: the hook (FlgE), the hook length regulator (FliK), and hook assembly factors FlgD, FlhO, and FlhP. We note that if FlhO and FlhP represent distal rod proteins, they are interesting candidates for the study of Gram-positive-specific flagellar biology as the rod is the structure best situated to accommodate variations in the cell envelope.
We thank K. Blair, R. Calvo, S. Guttenplan, M. Marketon, S. Mukhopadhyay, S. Mukherjee, J. Patrick, B. Stein, and E. Vanderpool for reagents, assistance, and critical reading of the manuscript. This work was supported by NIH grant GM093030 to DBK.
<table>
<thead>
<tr>
<th>Strain</th>
<th>Genotype</th>
</tr>
</thead>
<tbody>
<tr>
<td>3610</td>
<td>Wild type</td>
</tr>
<tr>
<td>DK17</td>
<td>ΔflgDE amyE::P_fla/che-flgET123 cat</td>
</tr>
<tr>
<td>DK22</td>
<td>flhPΩlacZ cat</td>
</tr>
<tr>
<td>DK23</td>
<td>sigD::tet flhPΩlacZ cat</td>
</tr>
<tr>
<td>DK24</td>
<td>flgM::tet flhPΩlacZ cat</td>
</tr>
<tr>
<td>DS793</td>
<td>amyE::P_{phi-lacZ} cat</td>
</tr>
<tr>
<td>DS1916</td>
<td>amyE::P_{phi-hag}_200C spec</td>
</tr>
<tr>
<td>DS3159</td>
<td>sigD::tet amyE::P_{phi-hag}_200C spec</td>
</tr>
<tr>
<td>DS4536</td>
<td>AflK</td>
</tr>
<tr>
<td>DS4681</td>
<td>AflgE</td>
</tr>
<tr>
<td>DS5161</td>
<td>AflhO</td>
</tr>
<tr>
<td>DS5775</td>
<td>amyE::P_{phi-lacZ} cat</td>
</tr>
<tr>
<td>DS5776</td>
<td>amyE::P_{phi-lacZ} cat</td>
</tr>
<tr>
<td>DS5806</td>
<td>flhP:: kan amyE::P_{phi-lacZ} cat</td>
</tr>
<tr>
<td>DS5807</td>
<td>AflhO amyE::P_{phi-lacZ} cat</td>
</tr>
<tr>
<td>DS5808</td>
<td>AflgE amyE::P_{phi-lacZ} cat</td>
</tr>
<tr>
<td>DS5809</td>
<td>AflhK amyE::P_{phi-lacZ} cat</td>
</tr>
<tr>
<td>DS5812</td>
<td>flhP:: kan amyE::P_{phi-lacZ} cat</td>
</tr>
<tr>
<td>DS5821</td>
<td>AflhK amyE::P_{phi-lacZ} cat</td>
</tr>
<tr>
<td>DS5822</td>
<td>AflgE amyE::P_{phi-lacZ} cat</td>
</tr>
<tr>
<td>DS5823</td>
<td>AflhO amyE::P_{phi-lacZ} cat</td>
</tr>
<tr>
<td>DS5824</td>
<td>AflhP amyE::P_{phi-lacZ} cat</td>
</tr>
<tr>
<td>DS5894</td>
<td>AflhP amyE::P_{phi-hag}_200C spec</td>
</tr>
<tr>
<td>DS5895</td>
<td>AflhK amyE::P_{phi-hag}_200C spec</td>
</tr>
<tr>
<td>DS5896</td>
<td>AflhE amyE::P_{phi-hag}_200C spec</td>
</tr>
<tr>
<td>DS5897</td>
<td>AflhO amyE::P_{phi-hag}_200C spec</td>
</tr>
<tr>
<td>DS5944</td>
<td>AflhO thrC:: P_{phi-flhO} mls</td>
</tr>
<tr>
<td>DS6388</td>
<td>AflgE amyE::P_{phi-lacZ} cat</td>
</tr>
<tr>
<td>DS6555</td>
<td>AflgD</td>
</tr>
<tr>
<td>DS7006</td>
<td>thrC:: P_{phi-CFP} mls amyE::P_{phi-YFP} cat</td>
</tr>
<tr>
<td>DS7351</td>
<td>AflhP</td>
</tr>
<tr>
<td>DS7360</td>
<td>AflhP amyE::P_{phi-flhP} cat</td>
</tr>
<tr>
<td>DS7671</td>
<td>amyE::P_{phi-lacZ} cat</td>
</tr>
<tr>
<td>DS7673</td>
<td>amyE::P_{phi-lacZ} cat</td>
</tr>
<tr>
<td>DS7770</td>
<td>flgM:: tet amyE::P_{phi-lacZ} cat</td>
</tr>
<tr>
<td>DS8037</td>
<td>AflhK amyE::P_{phi-lacZ} cat</td>
</tr>
<tr>
<td>DS8704</td>
<td>AflhK thrC:: P_{phi-CFP} mls amyE::P_{phi-YFP} cat</td>
</tr>
<tr>
<td>DS8705</td>
<td>AflgE thrC:: P_{phi-CFP} mls amyE::P_{phi-YFP} cat</td>
</tr>
<tr>
<td>DS8706</td>
<td>AflhO thrC:: P_{phi-CFP} mls amyE::P_{phi-YFP} cat</td>
</tr>
<tr>
<td>DS8707</td>
<td>AflhP thrC:: P_{phi-CFP} mls amyE::P_{phi-YFP} cat</td>
</tr>
<tr>
<td>DS8795</td>
<td>flgM:: tet amyE::P_{phi-lacZ} cat</td>
</tr>
<tr>
<td>DS8800</td>
<td>flgM:: tet flhP:: kan amyE::P_{phi-lacZ} cat</td>
</tr>
<tr>
<td>DS8801</td>
<td>flgM:: tet AflhO amyE::P_{phi-lacZ} cat</td>
</tr>
<tr>
<td>DS8802</td>
<td>flgM:: tet AflgE amyE::P_{phi-lacZ} cat</td>
</tr>
<tr>
<td>DS8803</td>
<td>flgM:: tet AflhK amyE::P_{phi-lacZ} cat</td>
</tr>
<tr>
<td>DS8804</td>
<td>flgM:: tet flhP:: kan amyE::P_{phi-lacZ} cat</td>
</tr>
<tr>
<td>DS8805</td>
<td>flgM:: tet AflhK amyE::P_{phi-lacZ} cat</td>
</tr>
<tr>
<td>DS8806</td>
<td>flgM:: tet AflgE amyE::P_{phi-lacZ} cat</td>
</tr>
<tr>
<td>DS8807</td>
<td>flgM:: tet AflhO amyE::P_{phi-lacZ} cat</td>
</tr>
<tr>
<td>DS8808</td>
<td>flgM:: tet AflgE amyE::P_{phi-lacZ} cat</td>
</tr>
<tr>
<td>DS8839</td>
<td>AflhO AflgE amyE::P_{phi-flgET123C} cat</td>
</tr>
<tr>
<td>DS8840</td>
<td>AflhP AflgE amyE::P_{phi-flgET123C} cat</td>
</tr>
<tr>
<td>DS9840</td>
<td>AflgD amyE:: P_{phi-lacZ} cat</td>
</tr>
<tr>
<td>DS9841</td>
<td>AflgD amyE:: P_{phi-lacZ} cat</td>
</tr>
<tr>
<td>DS9852</td>
<td>flgM:: tet AflgD amyE:: P_{phi-lacZ} cat</td>
</tr>
<tr>
<td>Strain</td>
<td>FlgM::tet</td>
</tr>
<tr>
<td>----------</td>
<td>------------</td>
</tr>
<tr>
<td>DS9853</td>
<td></td>
</tr>
<tr>
<td>DS9892</td>
<td></td>
</tr>
<tr>
<td>DS9893</td>
<td></td>
</tr>
<tr>
<td>DS9894</td>
<td></td>
</tr>
<tr>
<td>DS9915</td>
<td>sigD::tet</td>
</tr>
<tr>
<td>DS9916</td>
<td></td>
</tr>
<tr>
<td>DS9920</td>
<td></td>
</tr>
</tbody>
</table>

- flgM::tet  ΔflgD  amyE::PlytF-lacZ cat
- sigD::tet  ΔflgE amyE::PD-3Pfla/che-flgET123C cat
- ΔflgD amyE::Pflag-hag2066 spec
- sigD::tet  amyE::Pfla/che-flgET123C cat
- flgM::tet  amyE::Pfla/che-flgET123C cat
- ΔflgD amyE::Pfla/che-flgET123C cat
Table 2: $P_{flhO}$ is a $\sigma^D$-dependent promoter

<table>
<thead>
<tr>
<th>Genotype</th>
<th>$amyE::P_{flhO}$-lacZ$^\prime$</th>
<th>$flhP$-lacZ$^\prime$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wild type</td>
<td>51 ± 2 (DS5776)</td>
<td>46 ± 4 (DK22)</td>
</tr>
<tr>
<td>$sigD$</td>
<td>1 ± 0 (DS9915)</td>
<td>4 ± 1 (DK23)</td>
</tr>
<tr>
<td>$flgM$</td>
<td>107 ± 5 (DS9916)</td>
<td>93 ± 9 (DK24)</td>
</tr>
</tbody>
</table>

All cultures were grown to 1.0 OD$_{600}$ in LB medium. All values are the average of three replicas (Miller units ± standard deviation). Strains used to generate the data are indicated in parentheses.
REFERENCES


**Figure 1:** *B. subtilis* encodes three proteins homologous to the hook structural protein and each is required for motility. A) Multiple sequence alignment of FlgE from *Salmonella enterica* subsp. Typhimurium (Sen) and FlgE, FlhO, and FlhP proteins from and *Bacillus subtilis* (Bsu). B) Genetic map of regions encoding the *flgE*, *flhO*, *flhP*, and *fliK* genes, including location of mutations used in the study and the organization of ectopic complementation constructs. Note that *flgE* and *fliK* are encoded within the long 25 kb *fla/che* operon whereas *flhO* and *flhP* are expressed as a separate dicistron. C) Quantitative swarm expansion assays. Each graph contains wild type 3610 (solid circles), a strain containing a mutation in the gene indicated in the upper left hand corner (open circles), and a strain containing the indicated mutation and ectopic complementation construct of the gene indicated in the upper left hand corner (grey circles). In the leftmost graph, the grey triangles indicate complementation with the *flgET123C* allele. The following strains were used to generate the swarm expansion assays: 3610 (wild type), DS4681 (*AflgE*), DS6388 (*AflgE amyE::P_D3_Pfla/che-flgE*), DS5161 (*AflhO*), DS5944 (*AflhO thrC::P_flhO-flhO*), DS7351 (*AflhP*), DS7360 (*AflhP amyE::P_flhO-flhP*), DS4536 (*Aflk*), DS8037 (*Aflk amyE::P_D3_Pfla/che-fliK*), DS7673 (*AflgE amyE::P_D3_Pfla/che-flgE*), DS6555 (*AflgD*), DS9920 (*AflgD amyE::P_D3_Pfla/che-flgD*).

**Figure 2:** FlgE is the primary structural subunit of the hook. A) Fluorescence micrographs in which membranes have been stained with FM4-64 (false colored red) overlaid with the HagT209C allele of flagellin that was stained with a maleimide dye.
Figure 3: FlgE is secreted into the supernatant of cells mutated for FlhO, FlhP, and FlgD. Cells were grown to mid-log phase, and cell pellets and supernatants were separated. Cell pellets were lysed and supernatants were concentrated by TCA precipitation of dissolved proteins. Lysates and supernatants were separately resolved by SDS-PAGE, electroblotted and probed with anti-Hag antibody (Panel A), anti-SigA antibody (Panel B), or anti-FlgE antibody (Panel C). All three panels were generated using the same samples. The following strains were used to generate the samples: wild type (3610), flgE (DS4681), flhO (DS5161), flhP (DS7351), fliK (DS4536), flgD (DS6555), and sigD (DS323).

Figure 4: Cells defective in hook completion grow in short chains. Fluorescence micrographs of cells of the indicated genotype were membrane stained with FM4-64 (false colored red) and contained a reporter for flagellin (Phag-YFP, false colored green).
and autolysin ($P_{\text{flg}}$-CFP, false colored blue) expression. Merge indicates an overlay of the three channels of fluorescence. Scale bar is 2 μm. The following strains were used to generate this figure: wild type (DS7006), ΔflgE (DS8705), ΔflhO (DS8706), ΔflhP (DS8707) ΔfliK (DS8704), and ΔflgD (DS9894).

Figure 5. Cells defective in hook completion are reduced for expression of flagellin ($P_{\text{phag}}$) and autolysins ($P_{\text{flg}}$) in a FlgM-dependent manner. β-galactosidase assays of strains of the indicated genotype that either express a $P_{\text{phag}}$-lacZ reporter (Panel A) or a $P_{\text{flg}}$-lacZ reporter (Panel B). In each panel, black bars represent the genotype indicated on the X-axis (Ø) and gray bars represent the genotype indicated on the X-axis that also contains a mutation in flgM (flgM). Error bars are the standard deviation of three replicates. The raw data used to generate this figure is included in Supplemental Table S3.

Figure 6. Two models for the function of FlhO and FlhP. A) FlhO and FlhP may form the distal rod. In this model, FlgD is inserted on the end of FlhO and FlhP and FlgE subunits are inserted underneath FlgD. B) FlhO and FlhP may form an extended cap. In this model, FlhO and FlhP either form the distal rod or are the first proteins inserted underneath the FlgD cap. In either case, FlgE subunits are inserted underneath FlhO and FlhP. Basal body structure is shown in pink. Hook (FlgE) is shown in blue. The predicted locations of FlhO and FlhP are indicated. Dark grey indicates membrane. Light grey indicates peptidoglycan.
Figure 1

A) Sequence alignment of different strains of the fla/che operon.

B) Diagram of the fla/che operon, showing regulatory elements and gene expression.

C) Graphs showing the swarm radius over time for different genes involved in motility.
Wild type flgE flhO flhP

Figure 2

fliK

FlgET123C

HagT209C

Wild type flgE flhO flhP flgE fliK flgE

A) B)

flgE

flhO flgE

flhP

flhO flgE

fliK

flhP flgE

flgK flgE

flgD flgE

flgD

sigD

sigD flgE

flhP

flgE
pellet
supernatant
pellet
supernatant
pellet
supernatant
Hag
Hag
SigA
SigA
FlgE
FlgE
WT flgE flhO flhP fliK flgD sigD
WT flgE flhO flhP fliK flgD sigD
WT flgE flhO flhP fliK flgD sigD
A
B
C
Figure 5

A) 

B) 

β-gal activity (MU)
Figure 6

A) Model 1: Distal rod
B) Model 2: Extended cap