JOURNAL OF BACTERIOLOGY, Oct. 1997, p. 6092–6099 Vol. 179, No. 19
0021-9193/97/$04.00+0

The FliO, FliP, FliQ, and FliR Proteins of *Salmonella typhimurium*: Putative Components for Flagellar Assembly

KOUHEI OHNISHI,† FAN FAN,‡ GARY J. SCHOENHALS,§ MAY KIHARA, AND ROBERT M. MACNAB*

Department of Molecular Biophysics and Biochemistry, Yale University, New Haven, Connecticut 06520-8114

Received 27 May 1997/Accepted 18 July 1997

The flagellar genes *fliO*, *fliP*, *fliQ*, and *fliR* of *Salmonella typhimurium* are contiguous within the *fliLMNOPQR* operon. They are needed for flagellation but do not encode any known structural or regulatory components. They may be involved in flagellar protein export, which proceeds by a type III export pathway. The genes have been cloned and sequenced. The sequences predict proteins with molecular masses of 13,068, 26,755, 9,592, and 28,933 Da, respectively. All four gene products were identified experimentally; consistent with their high hydrophobic residue content, they segregated with the membrane fraction. From N-terminal amino acid sequence analysis, we conclude that *fliO* starts immediately after *fliN* rather than at a previously proposed site downstream. FliP existed in two forms, a 25-kDa form and a 23-kDa form. N-terminal amino acid analysis of the 23-kDa form demonstrated that it had undergone cleavage of a signal peptide—a rare process for prokaryotic cytoplasmic membrane proteins. Site-directed mutation at the cleavage site resulted in impaired processing, which reduced, but did not eliminate, complementation of a *fliP* mutant in swarm plate assays. A cloned fragment encoding the mature form of the protein could also complement the *fliP* mutant but did so even more poorly. Finally, when the first transmembrane span of MotA (a cytoplasmic membrane protein that does not undergo signal peptide cleavage) was fused to the mature form of FliP, the fusion protein complemented very weakly. Higher levels of synthesis of the mutant proteins greatly improved function. We conclude that, for insertion of FliP into the membrane, cleavage is important kinetically but not absolutely required.

In previous work, we have characterized most of flagellar region IIIb of the chromosome of *Salmonella typhimurium*, from *fliE* (encoding a basal-body protein) through *fliF* (basal-body MS ring), *fliG* (switch), *fliH* (export), *fliI* (export and ATPase), *fliJ* (unknown function), *fliK* (hook-length control), *fliL* (unknown function), *fliM* (switch), and *fliN* (switch) (4, 7, 10, 12–14, 19, 22, 30). Only *fliO*, *fliP*, *fliQ*, and *fliR* were left uncharacterized.

The *fliO*, *fliP*, *fliQ*, and *fliR* genes are contiguous within the *fliLMNOPQR* operon. Their products have not been found in the flagellar structure (11), and there is no evidence that they are regulatory genes. They may therefore play a role in flagellar assembly rather than in flagellar function, perhaps being involved in the export of external components such as the rod proteins, hook protein, and flagellin. The *fliP*, *fliQ*, and *fliR* genes (as well as several other flagellar genes) have homologs in the gene systems for export of virulence factors in a variety of pathogenic bacteria; for example, they correspond to the *spa* (surface presentation of antigen) genes *spaP*, *spaQ*, and *spaR* in *S. typhimurium* and to *spa24*, *spa9*, and *spa29* in *Shigella* spp. (6). We describe here the characterization of the *S. typhimurium fliO*, *fliP*, *fliQ*, and *fliR* genes and their products.

Of particular interest was the finding that FliP, a cytoplasmic membrane protein, undergoes signal peptide cleavage. This event is common in eukaryotes but rare in prokaryotes, in which signal peptide cleavage is normally restricted to proteins that are destined for the periplasm, the outer membrane, or the cell exterior (1).

* Corresponding author. Mailing address: Department of Molecular Biophysics and Biochemistry 0734, Yale University, P.O. Box 208114, 266 Whitney Ave., New Haven, CT 06520-8114. Phone: (203) 432-5590. Fax: (203) 432-9782. E-mail: robert_macnab@qm.yale.edu.

† Present address: Marine Biotechnology Institute, Kamaishi, Iwate 026, Japan.

‡ Present address: Microbiology Department, SmithKline Beecham Pharmaceuticals, UP1345, Collegeville, PA 19426.

§ Present address: R. W. Johnson, PRI, San Diego, CA 92121.

MATERIALS AND METHODS

Strains and plasmids. The following commercial strains and vectors were used: BL21(DE3) (26); pBR322, pACYC184, and pUC19 (New England Biolabs, Beverly, Mass.); pUC118 (Takaro Shuzo, Shiga, Japan); HMS174(DE3), pET3a, pET11a, pET19b, and pET22b (Novagen, Madison, Wis.); pTrc99A (Pharmacia, Piscataway, N.J.); and pBluescript II KS(+) and pBluescript II SK(+) (Stratagene, La Jolla, Calif.). pTrc99A-FF4 is a modified form of pTrc99A in which the *Nde*I site within the vector has been removed and the *Nco*I site in the multiple cloning region has been converted to an *Nde*I site by PCR mutagenesis (3a).

S. typhimurium strain SJW1103 is wild type for motility (33). *Escherichia coli* strains used for complementation were YK4458 (*fliO*), YK4150 (*fliP*), YK4120 (*fliQ*), and YK4144 (*fliR*) (28). UH869 is an *E. coli* minicell-producing strain (8).

pKK1673 (a gift from K. Kutsukake, described in reference 14) is a pBR322-based plasmid containing the *S. typhimurium* region IIIb flagellar genes from *fliF* to beyond *fliR*. pKK1673Δ*Kpn*I is a derivative of pKK1673 with a deletion of the *Kpn*I fragment extending from *fliG* to the middle of *fliM* (14). Plasmids constructed during this study and carrying various *fliO*, *fliP*, *fliQ*, and *fliR* inserts are listed in Table 1.

^{35}S labeling of proteins synthesized under control of the T7 promoter or in minicells. *E. coli* strain BL21(DE3)pLysS was transformed with plasmids based on the pET series (Novagen) or pBluescript II KS/SK(+) (Stratagene) and containing an insert, with genes of interest expressed under control of the T7 promoter. A fresh transformant colony was used to inoculate 10 ml of Luria broth containing 50 μg of ampicillin per ml. Cells were grown at 37°C to an optical density at 600 nm of 0.5. Cells from 0.5 ml of this culture were pelleted and washed with 10 volumes of M9 medium. The cells were collected by centrifugation and resuspended in 1 ml of M9 medium supplemented with 0.5% glycerol, 10 μg of thiamine per ml, and all amino acids except methionine at a final concentration of 10 μg/ml. The expression of genes controlled by the T7 polymerase was induced by the addition of isopropyl-β-D-thiogalactopyranoside (IPTG) to 1 mM and subsequent incubation for 1 h at 37°C. Rifampin was added to a final concentration of 0.2 mg/ml to inactivate the host RNA polymerase, and the cells were incubated for another 30 min. Proteins were then labeled by the addition of 15 μCi of [^{35}S]methionine (Amersham, Arlington Heights, Ill.) and further incubation of the cells for 5 min. After labeling, the cells were pelleted and resuspended in protein sample buffer containing 1 or 2% sodium dodecyl sulfate (SDS). In the case of FliR, we found that most of the protein remained

TABLE 1. Plasmids constructed during this study

Plasmid	Vector[a]	Contents of insert
pMKM2	pBR322	10-kb *Eco*RI fragment from pKK1673Δ*Kpn*I
pAWU1	pBluescript II KS(+)	1.5-kb *Eco*RI-*Sal*I fragment from pMKM2
pAWU2	pUC118	2.5-kb *Hin*dIII-*Pst*I fragment from pMKM2
pAWU1/pAWU2	NA[b]	Combined information from pAWU1 and pAWU2
pAWU10	pACYC184	*fliO*, *fliP*, *fliQ*, and *fliR*
pAWU20	pBluescript II KS(+)	Same as pAWU10
pAWU11	pACYC184	Part of *fliN*, *fliO*, *fliP*, *fliQ*, and part of *fliR*
pAWU21	pBluescript II KS(+)	Same as pAWU11
pAWU13	pACYC184	*fliQ* (wild type)
pAWU23	pBluescript II KS(+)	Same as pAWU13
pAWU4	pBluescript II KS(+)	*fliN*, *fliO*, and most of *fliP*
pAWU14	pACYC184	*fliO* and most of *fliP*
pAWU24	pBluescript II KS(+)	Same as pAWU14
pAWU7	pBluescript II KS(+)	*fliQ* and *fliR*
pAWU8	pBluescript II SK(+)	*fliQ*, *fliR*, and part of *rcsA*
pGS17	pUC19	*fliO* (wild type)
pGS34	pUC19	C-terminally FLAG-tagged *fliO*
pGS45	pET11a	C-terminally His-tagged *fliO*
pIP31	pET22b	*fliP* (wild type)
pPFF3	pTrc99A-FF4[c]	Same as pIP31
pIP32	pET22b	*fliP* with altered signal sequence
pPFF4	pTrc99A-FF4	Same as pIP32
pIP40	pET22b	C-terminally His-tagged *fliP*
pIP41	pET22b	*fliP* C-terminal truncation with His-tag
pGS37	pUC19	C-terminally FLAG-tagged *fliP*
pPFF1	pET22b	*fliP* minus signal peptide
pPFF2	pTrc99A-FF4	Same as pPFF1
pPFF5	pET22b	*motA-fliP* fusion
pPFF6	pTrc99A-FF4	Same as pPFF5
pGS66	pUC19	*fliQ* (wild type)
pGS36	pUC19	Internally FLAG-tagged *fliQ*
pRFF1000	pET11a	*fliR* (wild type)
pRFF1300	pET19b	N-terminally His-tagged *fliR*
pRFF1301	pET19b	N-terminally truncated and His-tagged *fliR*
pRFF1302	pET19b	N-terminally His-tagged and C-terminally truncated *fliR*
pRFF1304	pET19b	N-terminally His-tagged and internally deleted *fliR*

[a] (i) pBR322 and pACYC184 are low-copy-number, general purpose plasmids. (ii) pUC19 and pUC118 are high-copy-number, general purpose plasmids. (iii) pBluescript II KS and SK(+) are high-copy-number, high-expression phagemids carrying the T7 promoter. (iv) The pET series are low-copy-number, high-expression plasmids carrying either the T7 or the T7*lac* promoter and permit optimization of both transcription and translation levels; in complementation tests, in which the host lacks the T7 polymerase gene, pET-based plasmids have low levels of expression. (v) pTrc99A is a medium-copy-number, high-expression plasmid carrying the *trc* promoter.

[b] NA, not applicable.

[c] See Materials and Methods.

in the stacking gel under these conditions and that the use of 4% SDS in the sample buffer greatly improved its solubilization.

For pulse-chase labeling, BL21(DE3) cells with the *fliP* gene under control of the T7 promoter were grown in the presence of rifampin. Cells were induced by IPTG and labeled with [^{35}S]methionine for 1 min at 37°C. At this point, unlabeled methionine was added to a final concentration of 20 mM. Samples were withdrawn at various times for SDS-polyacrylamide gel electrophoresis (SDS-PAGE) separation and autoradiography. Minicell labeling was carried out as described with strain UH869 (8).

Swarm tests. Cells were tested for their swarming ability on semisolid agar (1% Bacto-Tryptone, 0.35% Bacto-Agar [Difco, Detroit, Mich.], 0.7% NaCl).

Purification of C-terminally His-tagged FliO. C-terminally His-tagged FliO was overexpressed from plasmid pGS45 in *E. coli* strain HMS174(DE3). Following manufacturer's instructions (except that 1% [wt/vol] SDS was used in the loading buffer, the early wash buffers, and the elution buffer and 0.5% [wt/vol] SDS was used in the dialysis buffer), His-tagged FliO was bound to Ni-charged nitrilotriacetic acid resin (Novagen), washed, eluted, dialyzed to remove salts, and concentrated. The partially purified FliO-His was then subjected to SDS-PAGE (15% acrylamide), transferred to polyvinylidene difluoride membrane (70 V for 2 h), and stained with Coomassie blue (Bio-Rad, Hercules, Calif.). A major band corresponding to the long form of FliO was seen at about 17 kDa and was excised from the membrane and subjected to N-terminal amino acid sequencing.

N-terminal amino acid sequence analyses. Analyses were carried out by the Keck Foundation Biotechnology Resource Laboratory, Yale University.

Nucleotide sequence accession number. The DNA sequences of the *S. typhimurium fliO*, *fliP*, *fliQ*, and *fliR* genes have been deposited in GenBank under accession no. L49021.

RESULTS

Cloning of the *fliOPQR* portion of *S. typhimurium* flagellar region IIIb. Plasmid pKK1673 contains a large *Bam*HI fragment of *S. typhimurium* DNA encompassing all of flagellar regions IIIa and IIIb except for a deletion that extends from *fliC* in region IIIa through the intervening nonflagellar region and into *fliE* in flagellar region IIIb. A *Kpn*I deletion derivative, pKK1673Δ*Kpn*I, lacks the DNA from the start of *fliG* through the middle of *fliM*. A 10-kb *Eco*RI fragment of this insert was subcloned into pBR322 to give plasmid pMKM2, which complemented *fliO*, *fliP*, *fliQ*, and *fliR* mutants of *E. coli* (Fig. 1). Plasmids containing different portions of the *Eco*RI fragment were constructed and used for complementation analysis by using *E. coli* strains with mutations in the various genes as hosts. The results confirmed the gene order that was obtained by analysis of polarity with Tn*10* insertions (15).

DNA sequence of the *fliO*, *fliP*, *fliQ*, and *fliR* genes. The *fliL* operon of *S. typhimurium* consists of seven genes, *fliL–fliR*. The DNA sequence from *fliL* through *fliN* has been published previously (14). The sequence of *fliO* through *fliR* was obtained in this study (Fig. 2).

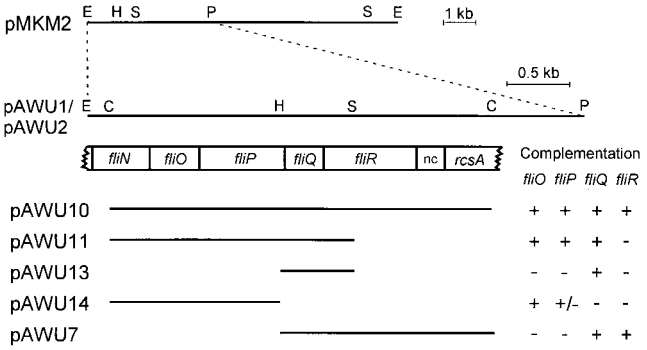

FIG. 1. Physical map of the part of flagellar region IIIb containing the *fliO*, *fliP*, *fliQ*, and *fliR* genes and complementation analysis of various plasmids. Lines represent the insert DNA in each plasmid. Restriction sites are abbreviated as follows: C, *Cla*I; E, *Eco*RI; H, *Hin*dIII; P, *Pst*I; S, *Sal*I. pAWU1/pAWU2 represents the overlapping information from plasmids pAWU1 and pAWU2; the actual plasmid was not constructed. *E. coli* mutants with defects in the genes shown above each column were used as hosts in complementation tests; the ability of transformed cells to swarm on semisolid agar plates is indicated. +, close to wild-type swarming; +/−, marginal swarming; −, no swarming.

The stop codon of *fliN* (TAA) overlaps with an ATG start codon (TAATG), which is preceded by a reasonable ribosome-binding consensus. We tentatively identified this as the start codon of *fliO* (but see below). The *fliO* and *fliP* genes have a TGATG overlap, whereas *fliP* and *fliQ* and *fliQ* and *fliR* are separated by 9 and 8 bp, respectively. The *fliP*, *fliQ*, and *fliR* genes all have moderate to strong well-placed ribosome-binding consensus sequences.

Our identification of the *fliO*, *fliP*, *fliQ*, and *fliR* genes is based on the following reasons. (i) All are predicted to be authentic coding regions by the program TESTCODE (5), (ii) the gene sizes are consistent with the observed sizes of the protein products (see below), and (iii) the levels of identity between the *S. typhimurium* and *E. coli* sequences are much higher at the amino acid level (see below) than at the nucleotide level (see discussion in reference 21).

The deduced protein sequences of FliO, FliP, FliQ, and FliR correspond to molecular masses of 13,068, 26,755, 9,592, and 28,933 Da, respectively. Comparison of the deduced sequences for FliP, FliQ, and FliR with those for *E. coli* (17) revealed a high degree of identity (91, 94, and 87%, respectively). The sequence of FliO was less well conserved between the two species (68% identity), with the N- and C-terminal regions being the most divergent.

After the terminus of *fliR*, there is a gap of 282 bp that contains no obvious open reading frames and is predicted to be noncoding. This is followed by an open reading frame that corresponds to the 5′ end of the *rcsA* gene (25), which encodes a positive regulator of capsular polysaccharide synthesis. Thus, as in *E. coli* (17), *fliR* is the last gene in the *S. typhimurium* flagellar region IIIb.

Identification of products encoded by the *fliOPQR* region. We transferred the inserts from the four pACYC184-based

```
    ..FliN                                                                                                  FliN end
106   I  A  Q  G  E  V  V  V  V  A  D  K  Y  G  V  R  I  T  D  I  I  T  P  S  E  R  M  R  R  L  S  R  *
  1 TGATTGCCCA GGGTGAAGTG GTCGTCGTGG CGGATAAATA CGGCGTACGT ATCACCGATA TTATTACGCC ATCCGAGCGT ATGCGTCGTT TGAGTCGTTA
                                                                                                          S-D
    FliO start                                                                  (FliO start 2?)
  1 M  M  K  T   E  A  T   V  S  Q   P  T  A  P   A  G  S   P  L  M   Q  V  S  G   A  L  I   G  I  I   A  L  I  L
101 ATGATGAAGA CAGAAGCCAC GGTTTCTCAG CCCACCGCGC CAGCCGGATC GCCGCTCATG CAGGTGAGCG GCGCATTAAT TGGCATTATT GCGCTCATTC

.........
                                                                                  FliO end  FliP start          signal
101 T  E  A  P   V  A  P   P  A  D   F  Q  N  M   M  K  S   L  L  K   R  S  G  R   S  * M   R  R  L  L   F  L  S
401 ACGGAAGCAC CCGTTGCCCC ACCGGCGGAT TTCCAGAACA TGATGAAGAG CTTACTCAAG CGTTCCGGGA GATCCTGATG CGCCGTTTGT TATTCCTTTC
                                                                                  S-D
    sequence
  9   L  A  G   L  W  L  F   S  P  A   A  A  A ↓ Q  L  P  G   L  I  S   Q  P  L   A  G  G  G   Q  S  W   S  L  S
501 TCTGGCGGGT CTGTGGCTTT TTAGTCCCGC CGCCGCTGCG CAACTGCCGG GGCTTATCAG CCAACCATTA GCGGGCGGAG GCCAAAGCTG GTCATTATCC

.........
          FliP end                FliQ start
242 S  F  Y  S   *                  M  T   P  E  S  V   M  M  M   G  T  E   A  M  K  V   A  L  A   L  A  A   P  L  L  L
1201 AGCTTTTACA GTTAGAGGCG CAAAATGACT CCTGAATCCG TCATGATGAT GGGCACCGAG GCGATGAAAG TCGCTCTCGC CCTGGCTGCG CCGCTGTTAC
     HindIII      S-D
.........
                                                                                                        FliQ end
  60  I  A  I   I  V  A  G   P  W  M   L  N  L   L  L  D  Y   V  R  T   L  F  S   N  L  P  Y   I  I  G   *
1401 CATCGCAATT ATCGTTGCCG GCCCGTGGAT GCTCAACCTG CTGCTGGATT ACGTGCGCAC TCTGTTCAGC AATTTACCTT ATATCATCGG ATAAACAGAA
                                                                                                      S-D
       FliR start
  1    M  I  Q   V  T  S   E  Q  W  L   Y  W  L   H  L  Y   F  W  P  L   L  R  V   R  A  L   I  S  T  A   P  I  L
1501 CGATGATCCA GGTTACAAGC GAGCAATGGC TATACTGGCT GCATCTCTAT TTCTGGCCTT TGCTGCGCGT ACGGGCGCTG ATTTCAACGG CGCCCATTCT

.........
                                                                                                      FliR end
234  P  L  I   A  P  F  C   E  H  L   F  S  E   I  F  N  L   L  A  D   I  V  S   E  M  P  I   N  N  N   P  *
2201 GCCATTAATT GCGCCGTTTT GCGAACATTT ATTCAGCGAA ATTTTCAATT TGCTTGCTGA TATTGTTAGC GAGATGCCGA TAAATAATAA CCCATAATTT

.........
                                                                                                RcsA start
  1                                                                                             M  S  T  I   I  M  D
2501 ACTATTCGTG TTCCGTCTGA TTATGGTGAG TTATTCAGGT AAGGGGAATT ATCGTTACGC ATTGAGTGAG GGGTATGCCA TGTCAACGAT TATTATGGAT
                                                                                                  S-D
```

FIG. 2. DNA sequence and amino acid sequence of the *fliO*, *fliP*, *fliQ*, and *fliR* genes and gene products of *S. typhimurium*. The flanking regions (the 3′ end of *fliN* and the 5′ end of *rcsA*, a gene unrelated to flagellar function) are also shown. The internal portion of each gene sequence (.....) has been omitted for brevity; the complete sequence is available in GenBank under accession number L49021. *, termination codon; S-D, Shine-Dalgarno or ribosome-binding consensus. For the significance of FliO start 2, see the text. The cleaved signal sequence for FliP is shown by overlining, with an arrow (↓) at the cleavage site.

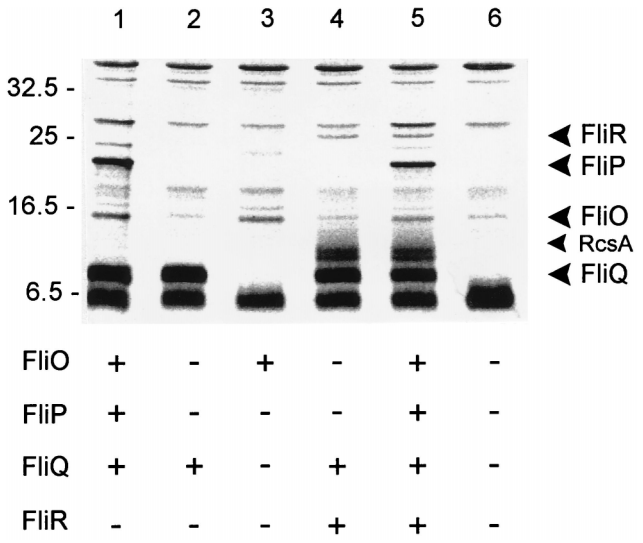

FIG. 3. Autoradiograph of flagellar proteins synthesized from various pBluescript II KS(+)-based plasmids under control of the T7 promoter. Lane 1, pAWU21; lane 2, pAWU23; lane 3, pAWU24; lane 4, pAWU7; lane 5, pAWU20; lane 6, pBluescript II KS(+). The flagellar genes present on the various plasmid inserts are shown below (cf. Fig. 1), and the positions of bands identified with the corresponding flagellar proteins are shown to the right. Numbers on the left indicate the positions of prestained molecular weight markers (in thousands).

plasmids shown in Fig. 1 (pAWU10, pAWU11, pAWU13, and pAWU14) into pBluescript II KS(+) in order to express the genes under control of the T7 promoter. The corresponding plasmids were named pAWU20, pAWU21, pAWU23, and pAWU24 [pAWU7 was already a pBluescript II KS(+) derivative]. Radiolabeled cells were harvested and analyzed by SDS-PAGE and autoradiography (Fig. 3). Plasmid pAWU20 (lane 5), which contains all four genes as judged by complementation, showed bands at about the predicted positions for FliO, FliP, FliQ, and FliR. It was not clear whether FliO had one or two bands, and the band corresponding to FliR was not seen in all experiments (see below).

The deletion derivatives of pAWU20 produced patterns (Fig. 3) that were consistent with the deletion boundaries and complementation patterns shown in Fig. 1. The inserts in pAWU14 and pAWU24 end at a *Hin*dIII site shortly before the end of the *fliP* gene, and plasmid pAWU14 complemented a *fliP* mutant very poorly (Fig. 1). The band predicted to correspond to truncated FliP was not seen with pAWU24 (Fig. 3, lane 3), presumably because the truncated protein was unstable. Other lines of evidence suggest that the C terminus of FliP is important for stability and function; pIP40 (with a C-terminal His tag) and pGS37 (with a C-terminal FLAG tag) (9) both complemented poorly.

Confirmation of the identity of the gene products. The identification of proteins by the means described above can be subject to some uncertainty, especially if the protein appears to have multiple forms or if detection is not reliable. In order to confirm the identity of FliO, FliP, FliQ, and FliR, we subcloned the genes and modified them in various ways. The results are described below and shown in Fig. 4.

(i) FliO. We subcloned *fliO* into pUC19 (in the reverse orientation to avoid problems with overexpression) starting with the ATG codon that overlaps the stop codon of *fliN*, yielding plasmid pGS17. Next, we modified pGS17 so that it encoded the same sequence fused to a C-terminal FLAG tag, yielding plasmid pGS34. In minicell expression experiments, pGS17 gave two bands at about the position expected for wild-type FliO (Fig. 4a, lane 1); with pGS34, both bands were shifted upward (lane 2). This result confirmed the identity of FliO and established that both bands correspond to forms of FliO. The specific identity of the lower band remains unclear.

(ii) FliP. Using pBluescript KS II(+) as the vector, we usually observed only one protein band associated with the presence of *fliP* on plasmids (Fig. 3). However, when the *fliP* gene was cloned into a pET vector to give plasmid pIP31, two bands were seen, one at about 25 kDa and the other at about 23 kDa, with the lower band being more intense (Fig. 4b, lane 1). To establish whether both bands were related to FliP (and also for future protein purification), we constructed two pET-based plasmids encoding forms of FliP with C-terminal His tags. One plasmid, pIP40, encoded the full-length FliP sequence, and the other, pIP41, encoded a C-terminally truncated FliP. pIP40 and pIP41 still produced two protein bands but shifted in the manner predicted relative to the two forms produced by pIP31 (Fig. 4b, lanes 2 and 3). We conclude that both bands are forms of FliP.

(iii) FliQ. We subcloned *fliQ* into pUC19, yielding plasmid pGS66. Next, we modified pGS66 so that the *fliQ* gene was disrupted by an internal FLAG tag, yielding plasmid pGS36. Minicell expression with pGS66 revealed a band at about the position expected for wild-type FliQ (Fig. 4c, lane 1); with pGS36, this band had shifted upward slightly (lane 2), confirming the identity of FliQ.

(iv) FliR. Previous attempts to detect the *fliR* gene product in *E. coli* (17) and *Caulobacter crescentus* (34) were unsuccessful. We also experienced considerable difficulty in detecting FliR. When the gene was expressed from the T7 promoter on pBluescript II, a protein with an apparent molecular mass of about 26 kDa was usually detected (e.g., Fig. 3, lanes 4 and 5). Because the intensity of the band varied considerably, confirmation of its correspondence to FliR was especially important.

We therefore cloned *fliR* and various deletion alleles into pET vectors, such that the gene was optimally positioned for translation initiation as well as for transcription from the T7 promoter. Some of the constructs employed an N-terminal His tag for use in future purification of FliR.

In T7 expression experiments (Fig. 4d), the plasmid encoding wild-type FliR produced a protein band with an apparent molecular mass of 26 kDa (pRFF1000 [Fig. 4d, lane 1]). With the plasmid encoding an N-terminally His-tagged full-length FliR (pRFF1300, [Fig. 4d, lane 2]) the band had shifted to an apparent molecular mass that was about 2 kDa larger. The His-tagged deletion versions of FliR gave about the expected mobility shifts (Fig. 4d, lanes 3 to 5).

The translational start for *fliO*. The location of the DNA sequence in the vicinity of the *fliN* termination codon suggested that the overlapping ATG codon represented the start of the *fliO* gene. However, it has been reported that there is a large intergenic gap of about 60 bp between *fliN* and *fliO* in *E. coli* (17). We therefore wished to establish the N-terminal sequence of the FliO protein to resolve this issue directly.

A C-terminally His-tagged version of FliO was overproduced from plasmid pGS45, whose insert starts with the ATG codon that overlaps the *fliN* stop codon. The product was purified by Ni affinity chromatography, subjected to SDS-PAGE, and electrotransferred to polyvinylidene difluoride membrane. The major (upper) band, with an apparent molecular mass of about 17.5 kDa, was excised and subjected to N-terminal amino acid sequence analysis. This yielded a perfect match (MMKTEATVSQTPA) to the first 13 amino acids

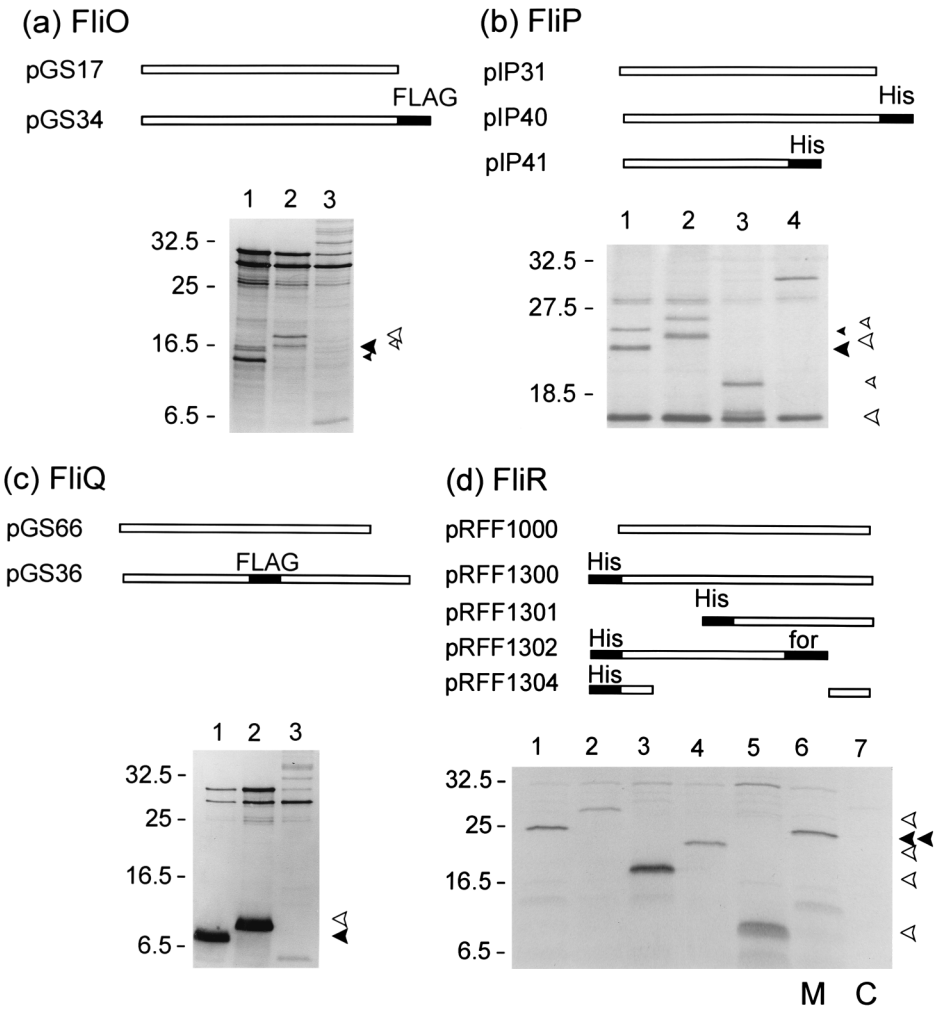

FIG. 4. Confirmation of the identification of the products of the *fliO*, *fliP*, *fliQ*, and *fliR* genes by band mobility shifts in SDS-PAGE. For *fliO* and *fliQ*, a minicell expression system was used; for *fliP* and *fliR*, genes were expressed under control of the T7 promoter. For each gene, the inserts on the plasmids used for the analysis are shown (not to scale) above the corresponding autoradiogram. Histidine and FLAG epitope tags are shown in black. Positions of the relevant proteins are shown by arrowheads, with black arrowheads indicating wild-type protein(s) and white arrowheads indicating proteins carrying tags, and in some cases, other alterations such as truncations or deletions. The small arrowheads in Fig. 4a indicate the second form of FliO; in Fig. 4b, they indicate the precursor forms of FliP. Numbers to the left of the autoradiograms indicate the positions of prestained molecular-weight markers. (a) Lane 1, wild-type FliO (pGS17); lane 2, C-terminally FLAG-tagged FliO (pGS34); lane 3, vector control (pUC19). Note the presence of two forms of FliO in lanes 1 and 2. (b) Lane 1, wild-type FliP (pIP31); lane 2, C-terminally His-tagged FliP (pIP40); lane 3, C-terminally truncated and His-tagged FliP (pIP41); lane 4, vector control (pET22b). Note the presence of two forms of two forms of FliP in lanes 1 to 3. (c) Lane 1, wild-type FliQ (pGS66); lane 2, internally FLAG-tagged FliQ (pGS36); lane 3, vector control (pUC19). (d) Lane 1, wild-type FliR (pRFF1000); lane 2, N-terminally His-tagged FliR (pRFF1300); lane 3, N-terminally truncated and His-tagged FliR (pRFF1301); lane 4, C-terminally truncated and N-terminally His-tagged FliR (pRFF1302), with the out-of-frame truncation generating foreign sequence (for) from the vector; lane 5, internally deleted and N-terminally His-tagged FliR (pRFF1304). Lanes 6 and 7 contain the membrane fraction (M) and cytosolic fraction (C), respectively, of sonicated cells transformed with pRFF1000 (wild-type FliR); FliR was found exclusively in the membrane fraction.

generated by translation from the ATG codon that overlaps the *fliN* stop codon (Fig. 2).

FliO, FliP, FliQ, and FliR are integral membrane proteins. Based on hydrophobicity analysis (3), all four proteins are strongly predicted to be located in the cytoplasmic membrane (data not shown). Cells were lysed either by French press or by sonication, and the high-speed pellet (corresponding roughly to the membrane fraction) was isolated from the high-speed supernatant (corresponding to the cytoplasmic fraction). All four proteins were found exclusively or predominantly in the membrane fraction, as is illustrated for FliR in Fig. 4d (lanes 6 and 7).

The predicted topologies of the proteins are shown schematically in Fig. 5. FliO is predicted to have a small N-terminal cytoplasmic domain, a single transmembrane helix, and a substantial periplasmic domain. FliP, after cleavage of its signal peptide (see below), is predicted to have a small N-terminal periplasmic domain, a transmembrane helix, a short cytoplasmic loop, a second transmembrane helix, and a large periplasmic domain; the last 80 or so amino acids are mostly hydrophobic and could constitute either a third transmembrane helix and a C-terminal sequence in the cytoplasm or a third transmembrane helix, a short cytoplasmic loop, a fourth transmembrane helix, and a short C-terminal sequence in the periplasm. FliQ has a predominantly hydrophobic sequence throughout and is predicted to consist of a short N-terminal sequence in the cytoplasm, a transmembrane helix, a fairly small periplasmic domain, a second transmembrane helix, and a short C-terminal sequence in the cytoplasm. FliR is predicted to have a substantial N-terminal cytoplasmic domain; the remainder of

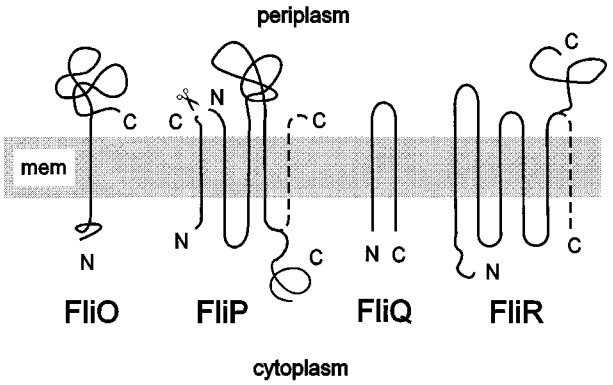

FIG. 5. Schematic illustration of the predicted topologies of FliO, FliP, FliQ, and FliR (see text). Dashed lines represent alternative possibilities. FliP is shown immediately after cleavage of its signal peptide. mem, cytoplasmic membrane.

its sequence is almost exclusively hydrophobic and appears to consist of a series of transmembrane helices connected by short loops. The number of helices is difficult to specify, but there are probably either five or six, meaning the C terminus would be located in the periplasm or cytoplasm, respectively.

FliP is a cytoplasmic membrane protein with a cleaved signal sequence. FliP exists in two forms differing in apparent molecular mass by about 2 kDa (Fig. 4b and reference 17). Pulse-chase experiments established that the upper band is the precursor of the lower band (Fig. 6, pIP31). We wished to establish the process responsible for generating the lower band. One possibility was signal peptide cleavage. From an inspection of the sequence (Fig. 2), we judged that the most likely cleavage site was between Ala21 and Gln22. N-terminal sequence analysis of mature FliP confirmed that cleavage had occurred and that the cleavage site was as predicted (data not shown). The deduced molecular mass for the mature form of FliP is 24,471 Da, i.e., about 2.3 kDa smaller than the precursor form. This difference is similar to the observed difference in apparent molecular mass.

Role of signal peptide cleavage in FliP function. Is signal peptide cleavage important for FliP function? To address this question, we attempted to inactivate the cleavage site by making the mutations A21D and Q22P, to give plasmid pIP32. However, pulse-chase experiments established that processing, though greatly reduced, was not eliminated (Fig. 6, pIP32). The processing that occurred was presumably at an adjacent position, since proline is not observed at the +1 position of cleaved proteins (20).

Next, we made two plasmids with extreme modifications to FliP. In one, pPFF1, only the mature form of FliP was encoded

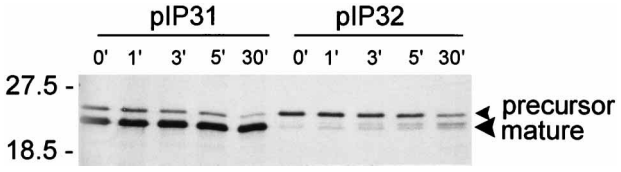

FIG. 6. Pulse-chase analysis of FliP synthesis and processing. Plasmids contained either the wild-type *fliP* gene (plasmid pIP31) or a version encoding a protein with the mutations A21D and Q22P at the signal sequence cleavage site (plasmid pIP32). Cells were labeled with [^{35}S]methionine for 1 min, at which point excess unlabeled methionine was added and samples were withdrawn for analysis at the times shown. Molecular weight markers (in thousands) are indicated on the left.

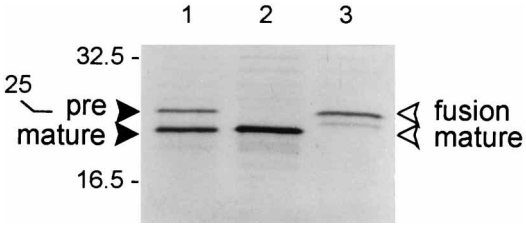

FIG. 7. Autoradiogram of different versions of FliP expressed under the control of the T7 promoter. Lane 1 (pIP31), wild-type FliP, showing both the precursor form (pre) and the mature form following signal peptide cleavage; lane 2 (pPFF1), mature FliP, synthesized without the signal sequence; lane 3 (pPFF5), MotA-FliP fusion, with the FliP signal sequence replaced by the first transmembrane segment of MotA. Molecular weight markers (in thousands) are indicated on the left.

by the plasmid insert. In the other, pPFF5, the signal sequence of FliP was replaced by the first transmembrane span of a cytoplasmic membrane protein (MotA) that does not have a cleavable signal sequence (32).

We then examined the products synthesized by these plasmids (Fig. 7). As expected, the mature form of FliP synthesized directly by plasmid pPFF1 had the same electrophoretic mobility as the mature form produced by signal peptide cleavage of the precursor form synthesized by plasmid pIP31. The MotA-FliP fusion protein synthesized by plasmid pPFF5 had approximately the same electrophoretic mobility as the precursor form of wild-type FliP. These constructions used pET-based vectors and correspond to low levels of expression in hosts lacking T7 polymerase. We also constructed equivalent pTrc-based plasmids (pPFF3, pPFF4, pPFF2, and pPFF6) in order to examine complementation at higher expression levels.

For pET-based plasmids (Fig. 8a), a progressive deterioration in complementation ability was seen as FliP was changed from the wild-type form to (i) the form with an altered signal sequence, (ii) the mature form lacking the signal peptide, and (iii) the MotA-FliP fusion protein. For pTrc-based plasmids (Fig. 8a), the form with the altered signal sequence complemented as effectively as wild-type FliP, while the mature form complemented almost as well. The MotA-FliP fusion protein, while complementing significantly better than when it was expressed from the pET-based plasmid, still complemented

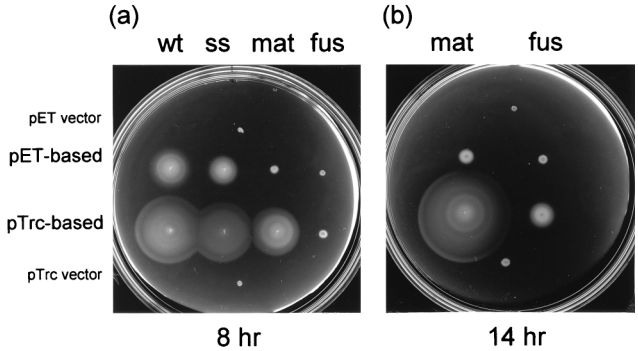

FIG. 8. Semisolid agar swarm plates of *fliP* mutant YK4150 transformed with various plasmids. (a) Top row, vector pET22b; second row, pET-based plasmids encoding wild-type FliP (wt), FliP with altered signal sequence (ss), mature FliP (mat), and MotA-FliP fusion (fus); third row, as in second row except that plasmids are pTrc-based; bottom row, vector pTrc99A-FF4. Incubation time was 8 h at 30°. (b) Top row, vector pET22b; second row, pET-based plasmids encoding mature FliP and MotA-FliP fusion; third row, as in second row except that plasmids are pTrc-based; bottom row, vector pTrc99A-FF4. Incubation time was 14 h at 30° C.

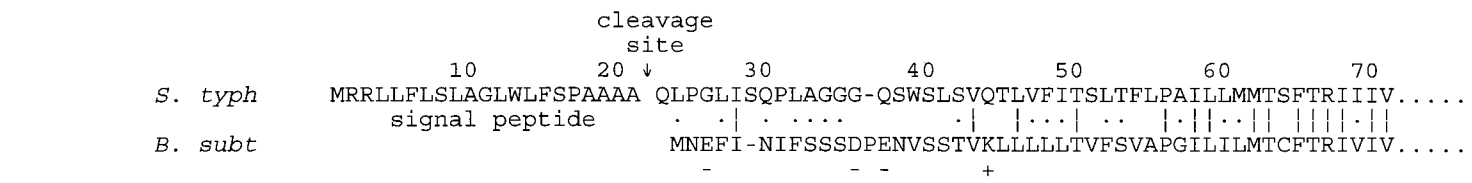

FIG. 9. Alignment of the N-terminal regions of the FliP proteins of *S. typhimurium* (this study) and *B. subtilis* (2). Identities are indicated by vertical bars, and similarities (I-L-V-M-F, F-Y, A-G, D-E, H-K-R, N-Q, and S-T) are indicated by dots. Note the N-terminal overhang of the *S. typhimurium* sequence, which corresponds to a cleavable signal peptide that is lacking in *B. subtilis*. Note also the prevalence of negatively charged residues (−) in the N-terminal region of the *B. subtilis* sequence (and the absence of such residues in the corresponding region of the mature *S. typhimurium* sequence); this may facilitate transfer of the N terminus of the *B. subtilis* protein across the membrane.

much more poorly than the others; prolonged incubation (Fig. 8b), however, made it clearer that some complementation was occurring.

DISCUSSION

In this study we have examined the properties of the FliO, FliP, FliQ, and FliR flagellar proteins of *S. typhimurium*. A study of these proteins in *E. coli* has been published by Malakooti et al. (17). We have extended the scope of their study in several regards and have clarified some aspects that seemed puzzling.

There is no large gap between the *fliN* and *fliO* genes. Large noncoding regions between genes in an operon are unusual in prokaryotes, which are normally quite economical in this regard. In the flagellar gene system, for example, adjacent genes are usually separated by no more than a few bases, and in many instances they actually overlap (16).

In the case of *fliO*, there is an ATG codon in TAATG overlap with the terminus of *fliN* (Fig. 2), and the sequence then continues as a long open reading frame which ends in TGATG overlap with the start of *fliP*. However, a plasmid carrying an insert that starts about 60 bases downstream of the terminus of *E. coli fliN* was reported to complement a *fliO* mutant (17). Translation was presumed to start from a GTG codon that is near the beginning of the insert and (in the full sequence) in the same frame as the ATG codon mentioned above. This suggested that the lost information is either noncoding or coding but nonessential.

We have now shown by N-terminal amino acid analysis that translation from the first ATG codon does occur. Admittedly, the insert we cloned was placed in an optimal situation with respect to both transcription and translation, so it might be argued that we artificially forced translation to initiate from that site. However, in plasmid pGS17, in which transcription is presumably from a gratuitous promoter, there is a major protein band at about the expected position for translation from the first start and no major band at the expected position for translation from the second start.

We conclude that, in the wild-type cell, translation of *fliO* occurs primarily from the first site. A more detailed analysis of the issue of the first and second start sites of *fliO* and of a stem-loop sequence that surrounds the putative second start will be presented elsewhere.

FliP is a cytoplasmic membrane protein with a cleavable signal peptide. We have established that, in the process of insertion into the cytoplasmic membrane, FliP undergoes signal peptide cleavage, with the precursor 25-kDa form giving way to the 23-kDa mature form. Taken in isolation, this fact suggests that FliP is either a periplasmic protein, an outer membrane protein, or an exported protein, since these are (with few exceptions) the only types of prokaryotic proteins that undergo signal peptide cleavage (1).

However, FliP is far too hydrophobic to be a soluble periplasmic protein, as cell fractionation experiments confirmed. Also, its hydrophobicity profile argues that FliP is not an outer membrane protein (outer membrane proteins usually have a different profile, with shorter hydrophobic regions, corresponding to transmembrane segments made up of β strands rather than the α helices that make up the transmembrane segments of cytoplasmic membrane proteins [2a]). Direct evidence that FliP is a cytoplasmic membrane protein is the fact that the cloned mature form is functional (Fig. 8), even though it has never undergone signal peptide cleavage and so could not have the outer membrane as its destination.

In prokaryotes, signal peptide cleavage of cytoplasmic membrane proteins is unusual (1). Bacteriophage M13 (fd) procoat protein (27) and subunit II of cytochrome *c* oxidase of *Paracoccus denitrificans* (24) are among the few examples in which it is known to occur. A consequence of the cleavage is that the N terminus of the mature protein is located in the periplasm. This location can be and sometimes is attained without signal peptide cleavage, as in the case of signal peptidase I itself (18, 31). In the case of FliP, there are two interesting examples of direct insertion without signal peptide cleavage. The first is the artificial example we generated in this study, namely, the cloned mature form of *S. typhimurium* FliP. The second is a natural example, namely, FliP of *Bacillus subtilis* (2), which aligns closely to FliP of *S. typhimurium* throughout most of its sequence, except that it has 22 fewer amino acids at its N terminus than the precursor form of *S. typhimurium* FliP (Fig. 9). Thus, the final topologies of FliP in the two species are presumably the same (with the N terminus in the periplasm), but they are achieved by different means—simple insertion in the case of *B. subtilis* and insertion followed by peptide cleavage in the case of *S. typhimurium.*

As expected, the MotA-FliP fusion protein did not undergo any processing, yet it retained some degree of function. As well as being of specific interest in the context of flagellar protein export and assembly, our data regarding the functionality of these versions of FliP may be of general interest in the context of insertion of prokaryotic proteins into the cytoplasmic membrane, a process which is still not well understood.

The roles of FliO, FliP, FliQ, and FliR. In this study we established that all four of these proteins are membrane bound. FliP, FliQ, and FliR clearly have homologs in the virulence systems of pathogenic bacteria. Both flagellar assembly and the presentation of virulence factors to the host require protein export. In both cases this export occurs, without signal peptide cleavage, by pathways that have been categorized as Class III (6, 23, 29). The simplest explanation for the observed homologies is that they reflect similarities in the export appa-

ratus itself; in other words, FliP, FliQ, and FliR (and possibly FliO, although there is no obvious homolog in this case) are likely to be components of the flagellar export apparatus and may be physically associated with the flagellum itself.

ACKNOWLEDGMENTS

We thank Albert Wu for assistance in constructing and testing the plasmids shown in Fig. 1. Manidipa Raha participated in early attempts to identify the *fliR* gene product. We are grateful to Deborah Kendall (University of Connecticut) for help in identifying published examples of prokaryotic cytoplasmic membrane proteins with a cleaved signal peptide.

This work was supported by USPHS grants AI12202 and GM40335.

REFERENCES

1. **Broome-Smith, J. K., S. Gnaneshan, L. A. Hunt, F. Mehraeln-Ghomi, L. Hashemzadeh-Bonehi, M. Tadayyon, and E. S. Hennessey.** 1994. Cleavable signal peptides are rarely found in bacterial cytoplasmic membrane proteins. Mol. Membr. Biol. **11:**3–8.
2. **Carpenter, P. B., A. R. Zuberi, and G. W. Ordal.** 1993. *Bacillus subtilis* flagellar proteins FliP, FliQ, FliR and FlhB are related to *Shigella flexneri* virulence factors. Gene **137:**243–245.

2a. **Engelman, D. M.** Personal communication.

3. **Engelman, D. M., T. A. Steitz, and A. Goldman.** 1986. Identifying nonpolar transbilayer helices in amino acid sequences of membrane proteins. Annu. Rev. Biophys. Biophys. Chem. **15:**321–353.

3a. **Fan, F.** Unpublished data.

4. **Fan, F., and R. M. Macnab.** 1996. Enzymatic characterization of FliI: an ATPase involved in flagellar assembly in *Salmonella typhimurium*. J. Biol. Chem. **271:**31981–31988.
5. **Fickett, J. W.** 1982. Recognition of protein coding regions in DNA sequences. Nucleic Acids Res. **10:**5303–5318.
6. **Galán, J. E.** 1996. Molecular genetic bases of *Salmonella* entry into host cells. Mol. Microbiol. **20:**263–271.
7. **Homma, M., T. Iino, and R. M. Macnab.** 1988. Identification and characterization of the products of six region III flagellar genes (*flaAII.3* through *flaQII*) of *Salmonella typhimurium*. J. Bacteriol. **170:**2221–2228.
8. **Homma, M., K. Kutsukake, and T. Iino.** 1985. Structural genes for flagellar hook-associated proteins in *Salmonella typhimurium*. J. Bacteriol. **163:**464–471.
9. **Hopp, T. P., K. S. Prickett, V. L. Price, R. T. Libby, C. J. March, D. P. Cerretti, D. L. Urdal, and P. J. Conlon.** 1988. A short polypeptide marker sequence useful for recombinant protein identification and purification. Bio/Technology **6:**1204–1210.
10. **Jones, C. J., M. Homma, and R. M. Macnab.** 1989. L-, P-, and M-ring proteins of the flagellar basal body of *Salmonella typhimurium*: gene sequences and deduced protein sequences. J. Bacteriol. **171:**3890–3900.
11. **Jones, C. J., R. M. Macnab, H. Okino, and S.-I. Aizawa.** 1990. Stoichiometric analysis of the flagellar hook-(basal-body) complex of *Salmonella typhimurium*. J. Mol. Biol. **212:**377–387.
12. **Kawagishi, I., M. Homma, A. W. Williams, and R. M. Macnab.** 1996. Characterization of the flagellar hook length control protein FliK of *Salmonella typhimurium* and *Escherichia coli*. J. Bacteriol. **178:**2954–2959.
13. **Kawagishi, I., V. Müller, A. W. Williams, V. M. Irikura, and R. M. Macnab.** 1992. Subdivision of flagellar region III of the *Escherichia coli* and *Salmonella typhimurium* chromosomes and identification of two additional flagellar genes. J. Gen. Microbiol. **138:**1051–1065.
14. **Kihara, M., M. Homma, K. Kutsukake, and R. M. Macnab.** 1989. Flagellar switch of *Salmonella typhimurium*: gene sequences and deduced protein sequences. J. Bacteriol. **171:**3247–3257.
15. **Kutsukake, K., Y. Ohya, S. Yamaguchi, and T. Iino.** 1988. Operon structure of flagellar genes in *Salmonella typhimurium*. Mol. Gen. Genet. **214:**11–15.
16. **Macnab, R. M.** 1988. The end of the line in bacterial sensing: the flagellar motor. Cold Spring Harbor Symp. Quant. Biol. **53:**67–75.
17. **Malakooti, J., B. Ely, and P. Matsumura.** 1994. Molecular characterization, nucleotide sequence, and expression of the *fliO*, *fliP*, *fliQ* and *fliR* genes of *Escherichia coli*. J. Bacteriol. **176:**189–197.
18. **Moore, K. E., and S. Miura.** 1987. A small hydrophobic domain anchors leader peptidase to the cytoplasmic membrane of *Escherichia coli*. J. Biol. Chem. **262:**8806–8813.
19. **Müller, V., C. J. Jones, I. Kawagishi, S.-I. Aizawa, and R. M. Macnab.** 1992. Characterization of the *fliE* genes of *Escherichia coli* and *Salmonella typhimurium* and identification of the FliE protein as a component of the flagellar hook-basal body complex. J. Bacteriol. **174:**2298–2304.
20. **Pugsley, A. P.** 1993. The complete general secretory pathway in gram-negative bacteria. Microbiol. Rev. **57:**50–108.
21. **Raha, M., M. Kihara, I. Kawagishi, and R. M. Macnab.** 1993. Organization of the *Escherichia coli* and *Salmonella typhimurium* chromosomes between flagellar regions IIIa and IIIb, including a large non-coding region. J. Gen. Microbiol. **139:**1401–1407.
22. **Raha, M., H. Sockett, and R. M. Macnab.** 1994. Characterization of the *fliL* gene in the flagellar regulon of *Escherichia coli* and *Salmonella typhimurium*. J. Bacteriol. **176:**2308–2311.
23. **Salmond, G. P. C., and P. J. Reeves.** 1993. Membrane traffic wardens and protein secretion in Gram-negative bacteria. Trends Biochem. Sci. **18:**7–12.
24. **Steinrücke, P., G. C. M. Steffens, G. Panskus, G. Buse, and B. Ludwig.** 1987. Subunit II of cytochrome *c* oxidase from *Paracoccus denitrificans*. DNA sequence, gene expression and the protein. Eur. J. Biochem. **167:**431–439.
25. **Stout, V., A. Torres-Cabassa, M. R. Maurizi, D. Gutnick, and S. Gottesman.** 1991. RcsA, an unstable positive regulator of capsular polysaccharide synthesis. J. Bacteriol. **173:**1738–1747.
26. **Studier, F. W., and B. A. Moffatt.** 1986. Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. J. Mol. Biol. **189:**113–130.
27. **Sugimoto, K., H. Sugisaki, T. Okamoto, and M. Takanami.** 1977. Studies on bacteriophage fd DNA. IV. The sequence of messenger RNA for the major coat protein gene. J. Mol. Biol. **110:**487–507.
28. **Suzuki, T., and Y. Komeda.** 1981. Incomplete flagellar structures in *Escherichia coli* mutants. J. Bacteriol. **145:**1036–1041.
29. **Van Gijsegem, F., C. Gough, C. Zischek, E. Niqueux, M. Arlat, S. Genin, P. Barberis, S. German, P. Castello, and C. Boucher.** 1995. The *hrp* gene locus of *Pseudomonas solanacearum*, which controls the production of a type III secretion system, encodes eight proteins related to components of the bacterial flagellar biogenesis complex. Mol. Microbiol. **15:**1095–1114.
30. **Vogler, A. P., M. Homma, V. M. Irikura, and R. M. Macnab.** 1991. *Salmonella typhimurium* mutants defective in flagellar filament regrowth and sequence similarity of FliI to F_0F_1, vacuolar, and archaebacterial ATPase subunits. J. Bacteriol. **173:**3564–3572.
31. **von Heijne, G.** 1989. Control of topology and mode of assembly of a polytopic membrane protein by positively charged residues. Nature (London) **341:**456–458.
32. **Wilson, M. L., and R. M. Macnab.** 1988. Overproduction of the MotA protein of *Escherichia coli* and estimation of its wild-type level. J. Bacteriol. **170:**588–597.
33. **Yamaguchi, S., H. Fujita, K. Sugata, T. Taira, and T. Iino.** 1984. Genetic analysis of *H2*, the structural gene for phase-2 flagellin in *Salmonella*. J. Gen. Microbiol. **130:**255–265.
34. **Zhuang, W. Y.** Personal communication.