JOURNAL OF BACTERIOLOGY, Jan. 1994, p. 189–197
0021-9193/94/$04.00+0

Vol. 176, No. 1

Molecular Characterization, Nucleotide Sequence, and Expression of the *fliO*, *fliP*, *fliQ*, and *fliR* Genes of *Escherichia coli*

JALEH MALAKOOTI,[1]† BERT ELY,[2] AND PHILIP MATSUMURA[1]*

Laboratory of Molecular Biology, Department of Biological Sciences, University of Illinois at Chicago, Chicago, Illinois 60680,[1] and Department of Biological Sciences, University of South Carolina, Columbia, South Carolina 29208[2]

Received 30 August 1993/Accepted 27 October 1993

The *fliL* operon of *Escherichia coli* contains seven genes that are involved in the biosynthesis and functioning of the flagellar organelle. DNA sequences for the first three genes of this operon have been reported previously. A 2.2-kb *Pst*I restriction fragment was shown to complement known mutant alleles of the *fliO*, *fliP*, *fliQ*, and *fliR* genes, the four remaining genes of the *fliL* operon. Four open reading frames were identified by DNA sequence analysis and correlated to their corresponding genes by complementation analysis. These genes were found to encode very hydrophobic polypeptides with molecular masses of 11.1, 26.9, 9.6, and 28.5 kDa for FliO, FliP, FliQ, and FliR, respectively. Analysis of recombinant plasmids in a T7 promoter-polymerase expression system enabled us to identify three of the four gene products. On the basis of DNA sequence analysis and in vivo protein expression, it appears that the *fliP* gene product is synthesized as a precursor protein with an N-terminal signal peptide of 21 amino acids. The FliP protein was homologous to proteins encoded by a DNA sequence upstream of the *flaA* gene of *Rhizobium meliloti*, to a gene involved in pathogenicity in *Xanthomonas campestris* pv. glycines, and to the *spa24* gene of the *Shigella flexneri*. The latter two genes encode proteins that appear to be involved in protein translocation, suggesting that the FliP protein may have a similar function.

The expression of flagellar and chemotaxis genes in *Escherichia coli* and *Salmonella typhimurium* is controlled by a complex regulatory hierarchy (20, 21, 24). The first level of the hierarchy is the *flhDC* master operon, whose protein products are required for expression of all the other genes of the regulon. Transcription of the *flhD* operon itself is controlled by the cyclic AMP (cAMP)-cAMP receptor protein complex (1, 22, 41). The second level of the hierarchy is controlled directly by the *flhDC* gene products and consists of seven operons containing a total of 27 genes. Class II genes code for proteins that are required for early stages of flagellum synthesis and assembly. The third level contains the late genes. These are the structural genes for the flagellar filament, the motility genes, and the chemotaxis genes. The expression of the level III genes is dependent on the expression and assembly of all of the early genes in level II (27). The cascade of gene expression in this regulatory scheme parallels the assembly of the gene products in the same hierarchial manner, from a less complex to a more complex structure (18, 43, 44).

Over 40 genes, clustered at four different regions on the chromosome, are devoted to the processes involved in the biogenesis or functioning of the bacterial flagellum. Electron microscopic examination has revealed that this complex structure is composed of three sections: a filament that extends beyond the cell envelope, a basal body that is embedded in the cell wall, and a hook that acts as a universal joint, connecting the filament to the basal body (2, 8). Additional structures on the cytoplasmic face of the MS ring have been identified in electron micrographs (11). On the basis of genetic analysis, the presence of two more substructures, the switch complex (50, 51) and a flagellum-specific export apparatus (49), has been inferred. Three proteins, FliG, FliM, and FliN, have been shown to be part of the switch complex (50), located in the cytoplasmic face of the inner membrane and interfacing with the basal body MS ring (12, 19). These proteins are involved in the early stages of flagellum assembly, which is prior to MS ring formation in the case of FliG and shortly after in the cases of FliN and FliM (49). Very little is known about the flagellum-specific transport machinery. Vogler et al. (49), using temperature-sensitive flagellar mutants of *S. typhimurium*, have shown that the *flhA*, *fliH*, *fliI*, and *fliN* gene products may be involved in the process of flagellum-specific export. Of the flagellar proteins, the FlgB, FlgC, FlgF, FlgG, FlgE, FlgK, FlgL, FliD, and FliC proteins are believed to be transported to their final destination via the flagellum-specific transport system (14, 25, 28). Thus far, only two flagellar proteins, FlgH and FlgI, have been found to have canonical signal sequences that are transported to their final destinations (the L and P rings, respectively) via the primary signal peptidase-dependent transport pathway (8, 13, 16, 17). There are still about 10 genes in the level II regulatory class whose gene products and roles in the assembly pathway have not yet been identified (28). We have determined the DNA nucleotide sequences for four of these genes, *fliO*, *fliP*, *fliQ*, and *fliR*, and have identified all but the *fliR* gene product. We show evidence that the FliP protein appears to be a transmembrane protein that contains a cleavable signal sequence at its N terminus. Sequence homology of the *fliP* gene product to proteins involved in the translocation of the virulence factors suggests a role for this protein in the flagellum-specific transport system.

* Corresponding author. Present address: Department of Microbiology and Immunology (M/C 790), University of Illinois, 901 S. Wolcott Ave., Chicago, IL 60612. Phone: (312) 996-2286.

† Present address: Department of Biological Sciences, University of South Carolina, Columbia, SC 29208.

TABLE 1. Bacterial strains, plasmids, and phages used

Strain, plasmid, or phage	Relevant genotype or characteristic	Reference or source
E. coli strains		
YK410	F⁻ *araD139 lac (U169) strA thi pyrC46 naDA thyA his*	20
YK4150	*fliP4150*	Parent, YK410
YK4114	*fliQ1083*	Parent, YK410
YK4113	*fliR871*	Parent, YK410
RP4184	*fliO52*	2
JM103	F⁺ (*traD proAB lacI*^{q}) *thi rpsL supE endA sbcB hsdR4*	48
XLI-Blue	F⁺ (*proAB lacI*^{q} M15 Tn*10*) *recA1 endA1 gyrA95 thi hsdR17* (r_K^- m_K^+)	Stratagene
Plasmids		
pJM8	pRL31 derivative containing the entire *fliL* operon	29
pJM9	2.2-kb *Pst*I fragment from pJM8 cloned in M13 mp8; excised with *Hin*dIII and *Bam*HI cloned into pRL22	29
pJM11	0.7-kb *Pst*I-*Hae*III fragment from pJM9 cloned into *Pst*I-*Sma*I-digested pUC19	This study
pJM12	1.0-kb *Hae*III fragment from pJM8 cloned into pUC18 in proper orientation with respect to *lacZ* promoter	This study
pJM13	0.8-kb *Sau*3A fragment was cloned into M13 mp8 at *Bam*HI site, then excised as a 0.6-kb *Taq*I-*Pst*I fragment, and cloned into *Acc*I-*Pst*I-digested pUC18	This study
pJM14	1.2-kb *Dra*I-*Pst*I fragment from pJM9 cloned in *Sma*I-*Pst*I-digested pUC18	This study
pJM15	1.0-kb *Pst*I-*Hin*fI fragment from pJM9 cloned in *Pst*I-*Eco*RI-digested pUC19	This study
pJM91	2.2-kb *Pst*I fragment from pJM9; excised with *Hin*dIII and *Bam*HI cloned into pBluescript SK⁺	This study
pJM111	0.7-kb *Pst*I-*Hae*III fragment from pJM11; excised with *Pst*I-*Sst*I cloned in pBluescript SK⁺	This study
pJM121	1.0-kb *Hae*III fragment from pJM12; excised with *Kpn*I and *Bam*HI cloned in pBluescript SK⁺	This study
pJM131	0.6-kb *Taq*I-*Pst*I fragment from pJM13; excised with *Kpn*I and *Pst*I and cloned into pBluescript SK⁺	This study
pJM141	1.2-kb *Dra*I-*Pst*I fragment from pJM14; excised with *Kpn*I and *Pst*I and cloned into pBluescript SK⁺	This study
pJM171	Exonuclease III-generated deletions of pJM91 from $PstI_{II}$ site; has a truncated *fliP* gene	This study
pUC18	Amp^r cloning vehicle	48
pUC19	Amp^r cloning vehicle	48
pBluescript SK⁺	Amp^r cloning vehicle	Stratagene
pBluescript KS⁺	Amp^r cloning vehicle	Stratagene
pGP1-2	Kan^r plasmid with T7 RNA polymerase gene under control of p_L promoter and c*Its*	45
Phages		
M13mp18	Used for single-strand sequencing	31
M13mp19	Used for single-strand sequencing	31
P1	Used for generalized transduction	31

MATERIALS AND METHODS

Bacterial strains, bacteriophages, plasmids, and media. The bacterial strains, plasmids, and phages used in this study are listed in Table 1. All *E. coli* strains were grown at 37°C in Luria broth (L-broth) medium or minimal medium M9 (30). Difco Bacto Agar was added to L broth at 1.7% for solid media and 0.35% for motility media. Antibiotic (Sigma Chemical Co., St. Louis, Mo.) supplements were used at the following concentrations: penicillin, 100 mg/liter; tetracycline, 30 mg/liter; and kanamycin, 51 mg/liter.

DNA manipulations. Restriction mapping, subcloning, and transformation were carried out by standard procedures (30). Restriction enzymes and other DNA-modifying enzymes were used as recommended by the suppliers.

DNA sequence analysis. Subfragments of the 2.2-kb *Pst*I restriction fragment that were generated by digestion with different restriction enzymes were isolated after electrophoresis and cloned into M13 mp18 or M13 mp19 phage vectors. DNA sequences were determined by the dideoxy-chain termination method (39). Nucleotide sequences were determined for both strands, using overlapping fragments. Sequence data manipulations and analysis of the DNA and protein sequences were carried out with the University of Wisconsin Genetics Computer Group package (9).

Complementation analysis. To eliminate recombination between plasmids and the chromosome of the *fli* mutant strains, a *recA* allele was introduced into mutants by P1 transduction (32). Transformation-competent mutant strains were transformed with plasmids pJM9, pJM11, pJM12, pJM13, pJM14, and pJM15 (Table 1). The transformants were selected on L-penicillin agar plates and screened for restoration of motility and chemotaxis behavior on L-penicillin soft agar plates.

Protein labeling and detection. For T7 protein expression, the method of Tabor and Richardson (45) was used. The flagellar genes were cloned downstream from the T7 promoter in Bluescript plasmids and introduced into XL1-Blue(pGP1-2) by transformation. Plasmid pGP1-2 contains the T7 RNA polymerase gene attached to strong λ p_L promoter and the temperature-sensitive λ repressor gene. To express the cloned gene product, cells were grown at 33°C in L broth to an optical density of 0.5 at 600 nm. Cells from 0.2 ml of this culture were pelleted and then washed with 10 volumes of M9 medium. The cells were collected by centrifugation and resuspended in 1 ml of M9 medium supplemented with 20 μg of thiamine per ml, 0.4% glycerol, and all amino acids minus methionine and cysteine at a final concentration of 20 μg/ml. This cell suspension was incubated at 33°C for 90 min. The expression of genes controlled by the T7 promoter was induced by heat induction at 42°C for 20 min. To inactivate host RNA polymerase, rifampin was added to the cell culture at a concentration of 200 μg/ml, and incubation was continued at 42°C for another 10 min. Next, the cells were incubated at 33°C for 20 min, labeled

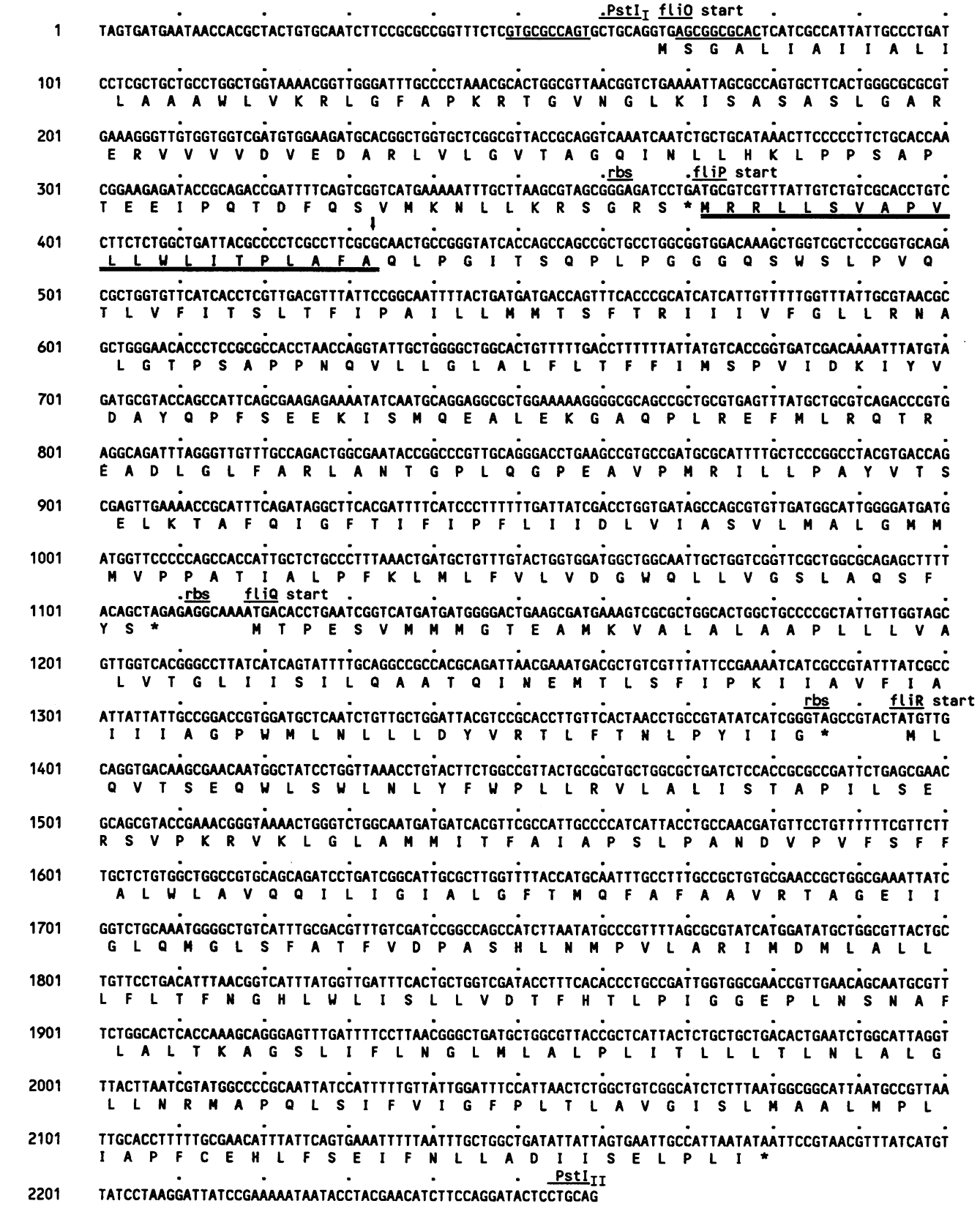

FIG. 1. DNA nucleotide and deduced amino acid sequences of the *fliO*, *fliP*, *fliQ*, and *fliR* genes. Numbers at the left refer to DNA nucleotides. Potential ribosome binding sites (rbs) are overlined. DNA sequences corresponding to the potential secondary structure in the *fliO* gene are underlined. Stop codons are indicated by asterisks. Two *Pst*I restriction sites are shown (proximal [*Pst*I$_I$] and distal [*Pst*I$_{II}$]). The potential signal peptide and its cleavage site are shown by a thick underline and a vertical arrow, respectively. The initiator codon for the *fliO* gene, GTG, which normally codes for a valine residue, is shown by an M for formylmethionine.

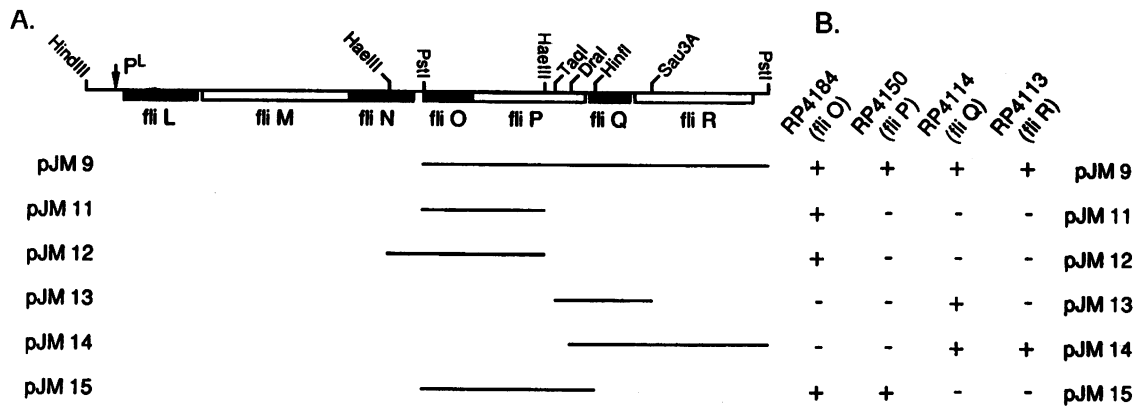

FIG. 2. (A) Physical map of the FliL operon. All seven genes of the operon are indicated. The solid bars represent the DNA fragments cloned in expression vehicles pUC18 and pUC19. All of the flagellar insert DNA in the expression vectors is transcribed through the *lacZ* promoter. (B) Complementation analysis. The ability (+) or inability (−) of the plasmids to complement mutation in the genes which were assigned to this region of DNA by genetic analysis is indicated.

by the addition of 10 μCi of Tran-^{35}S label (ICN Pharmaceuticals, Inc.), and incubated for an additional 5 min at 33°C. After labeling, the cells were pelleted and disrupted by addition of 2× sample buffer. The expression products were fractionated on sodium dodecyl sulfate (SDS)-polyacrylamide gels (4).

RESULTS AND DISCUSSION

DNA sequence analysis. The nucleotide sequence of a 2.2-kb *Pst*I DNA fragment that corresponds to the second half of the *fliL* operon was determined. The DNA and deduced amino acid sequences are shown in Fig. 1. Four open reading frames (ORFs), oriented in the same direction, were found within this region. This region adjoined the previously sequenced portion of the *fliL* operon (23, 29), resulting in a continuous sequenced DNA region of 4,404 bp, from a *Hin*dIII restriction site preceding the promoter to the distal *Pst*I site downstream from ORF4. The DNA sequence shown in Fig. 1 begins at the stop codon of the *fliN* gene (29), the third gene in the operon. Characterization of the ORFs and their corresponding genes is described below.

ORF1, the *fliO* gene. ORF1 starts with a GTG codon at bp 66 and extends to a TGA codon at bp 371. A 62-bp intercistronic DNA region is located between the end of the *fliN* gene and the proposed initiator codon for ORF1. Complementation tests with pJM9, pJM11, and pJM12 (Table 1 and Fig. 2) revealed that the DNA region downstream from the proximal *Pst*I site (*Pst*I$_I$) was capable of complementing a *fliO* mutant to wild-type motility and chemotaxis. Thus, ORF1 was assigned to the *fliO* gene. Initially, the ATG codon at bp 6 was thought to be the start codon for ORF1. However, since pJM11 is deleted of the first start codon at bp 6 and still is able to complement a *fliO* mutation (Fig. 2), it was concluded that the GTG at bp 66 was the most probable start codon for the *fliO* gene. In addition, in pJM11, the DNA sequence downstream from the *Pst*I$_I$ site was fused to the first four amino acid residues of the *lacZ* gene, which resulted in an out-of-phase fusion followed by a stop codon at position 67 (Fig. 1). Therefore, if translation is initiated at the *lacZ* start codon, it would not restore the FliO activity to this plasmid unless there is another start codon beyond the *Pst*I$_I$ restriction site. These data support the conclusion that the GTG codon downstream from the *Pst*I restriction site at position 66 is the initiator codon for the *fliO* gene. There is no obvious sequence homology to the consensus ribosome binding site, TAAGGAGGT (40), upstream from the *fliO* start codon except for an AGGT that overlaps the start codon. The *fliO* gene is 306 bp, encoding for a protein of 101 amino acid residues with a calculated molecular mass of 11.1 kDa. The 62-bp intercistronic region between the *fliN* and *fliO* genes, although not uncommon in bacterial operons, seems peculiar for an operon with an otherwise compact organization. Only a few such cases have been reported for flagellar operons (17). This region also contains a potential stem-loop structure, which if formed would overlap with the first 13 nucleotides of ORF1. A long intergenic region containing a similar structure has also been reported for the *flgG* and *flgH* genes (20). The role of this region and whether it has any regulatory significance are unclear. However, the presence of the secondary structure and a less efficient GTG initiation codon may indicate some role in the differential expression of the members of this operon. Differential expression is prevalent among operons with one or more genes encoding for integral membrane proteins (34, 38).

ORF2, the *fliP* gene. Complementation tests (Fig. 2) demonstrated that plasmid pJM15, which carries both ORF1 and ORF2, was capable of complementing the *fliO* and *fliP* mutants. Since pJM11 and pJM12, containing ORF1, complemented only the *fliO* defect, ORF2 must correspond to the *fliP* gene. Analysis of the DNA sequence (Fig. 1) showed that the initiation codon of the *fliP* gene overlapped with the termination codon of the upstream *fliO* gene, suggesting that the two genes may be coupled translationally (3). This potential coupling represents the second instance within the *fliL* operon, since the *fliN* start codon also overlaps with the stop codon of the *fliM* gene (29). The *fliP* gene is 738 nucleotides in length and is preceded by a DNA region, GGAG, which has homology to the consensus ribosome binding site (Fig. 1). The amino acid sequence deduced from the nucleotide sequence is 245 amino acid residues, with a calculated molecular mass of 26.9 kDa. The N-terminal amino acid sequence of the *fliP* gene product has the characteristics of a signal peptide (33). The putative signal sequence contains a long, predominantly hydrophobic region preceded by a short, positively charged N-terminal sequence. It is followed by a potential recognition site for signal peptidase cleavage. A helix-breaking proline is present four residues before the A-F-A cleavage site (Fig. 1) (35). Therefore, the predicted signal peptide consists of 21 amino acid residues, and the mature FliP protein would have a

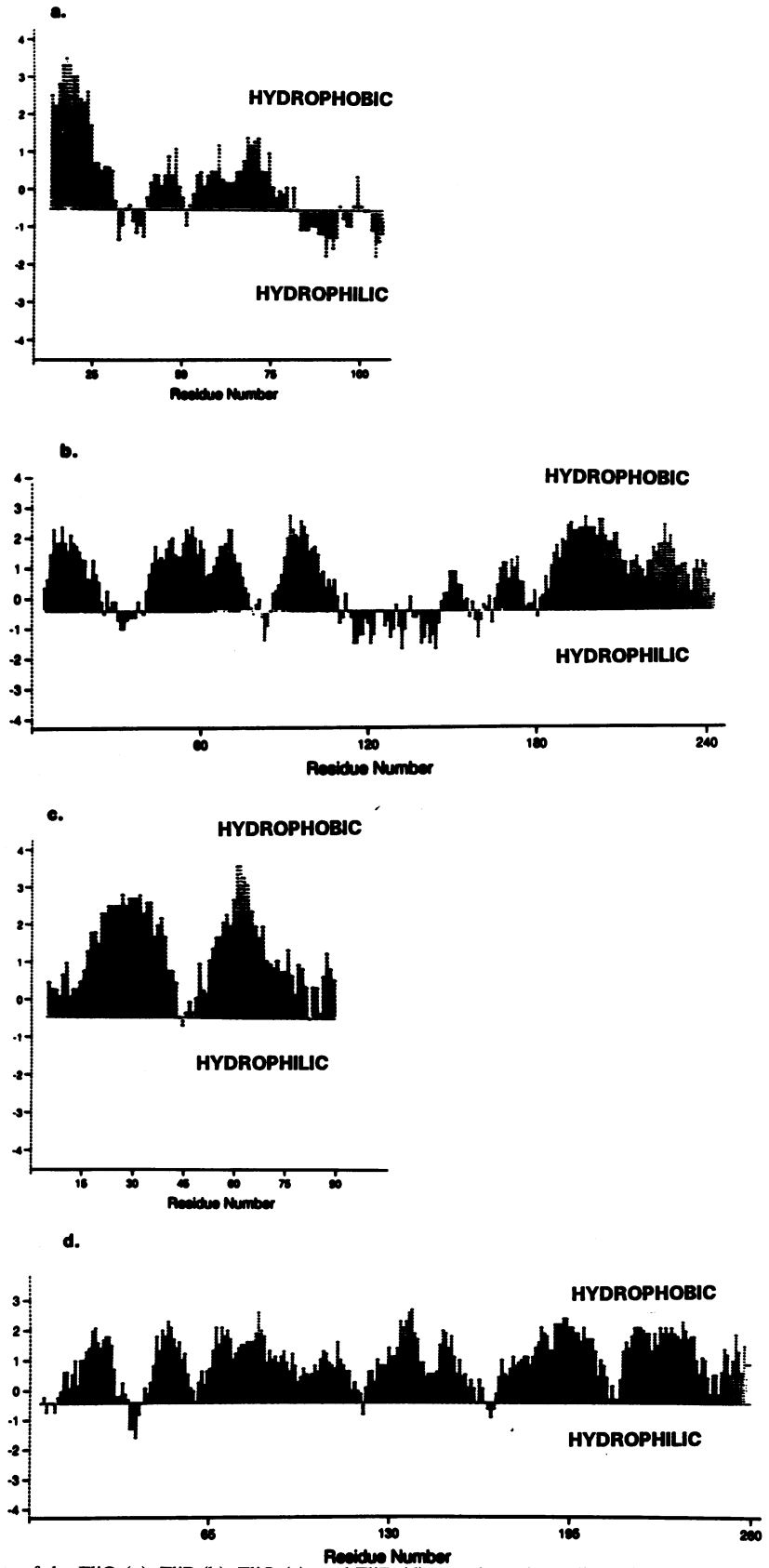

FIG. 3. Hydropathy plots of the FliO (a), FliP (b), FliQ (c), and FliR (d) proteins. The ordinate is the average hydropathic index of a stretch of nine amino acid residues, and the abscissa is the position of each amino acid. The midline at −0.4 corresponds to the mean hydropathic value of a pool of sequenced soluble proteins. The positive and negative values correspond to hydrophobic and hydrophilic segments, respectively.

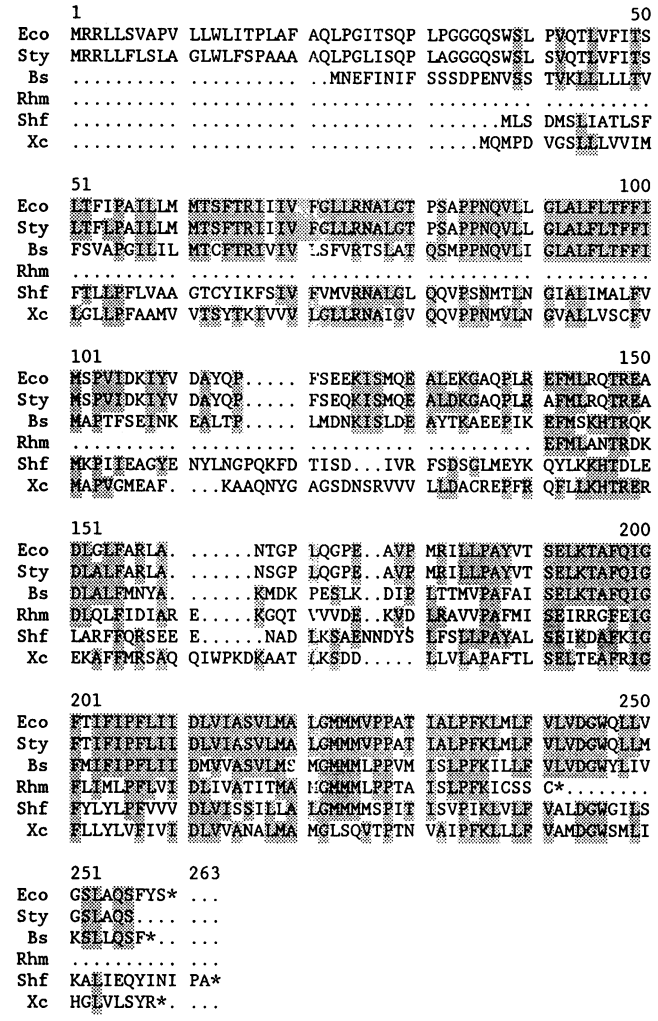

FIG. 4. Amino acid alignment of the FliP protein. The alignment was made by using the Pileup program of the University of Wisconsin Genetics Computer Group package (9). Gaps were introduced to improve the alignments. Residues that are conserved between the *E. coli* (Eco) FliP protein and at least two of the other proteins are shaded. Only identical residues are included in the shaded area. The *S. typhimurium* (Sty) *fliP* gene has been sequenced up to a *Hin*dIII restriction site, which is only 10 bp short of the end of the *fliP* gene. Bs, *B. subtilis*; Rhm, *R. meliloti*; Shf, *S. flexneri*; Xc, *X. campestris*.

calculated molecular mass of 24.1 kDa. The FliP protein contains a very high proportion (82.5%) of nonpolar amino acids. Leucine is the most abundant residue, comprising 17% of the total amino acids. The hydropathy profile of the FliP protein indicates that in addition to the N-terminal signal peptide, this polypeptide contains three hydrophobic segments that have the characteristics of membrane-spanning regions (Fig. 3) (26). Furthermore, the hydrophobic C-terminal segment is long enough to span the membrane at least two times. Proline residues are prevalent in the hydrophobic segments of the *fliP* gene product. Although proline residues are known to break α helices, they have been found in the membrane-located α helices of many proteins that function as transporters (6). The majority of the total charged amino acids (20 of 27) are distributed in a central hydrophilic region among amino acids 105 to 180.

ORF3, the *fliQ* gene. Immediately downstream from the *fliP* gene, a 270-bp ORF (ORF3) was detected (Fig. 1). Complementation analysis (Fig. 2) revealed that pJM13, which contains only the ORF3 DNA region, was capable of restoring motility to a *fliQ* mutant. A strong ribosome binding site, GAGG, was located 4 bp upstream from the start codon (Fig. 1). The predicted amino acid sequence consists of 89 residues, with a calculated molecular mass of 9.6 kDa. Two regions of high hydrophobicity, at residues 16 to 40 and 55 to 75, suggested that the FliQ protein contains two transmembrane segments (Fig. 3) (26).

ORF4, the *fliR* gene. Seven base pairs downstream from the end of the *fliQ* gene, ORF4 begins at bp 1395 and ends with a TAA codon 786 bp away. Complementation tests (Fig. 2) showed that pJM14, which carries both ORF3 and ORF4, was capable of complementing the *fliQ* and *fliR* defects. Since the *fliQ* mutant was complemented by a clone containing ORF3, ORF4 was designated the *fliR* gene. The *fliR* gene is preceded by a weak ribosome binding site 7 bp upstream from the ATG start codon (Fig. 1). The deduced amino acid sequence is extremely hydrophobic, with an unusually high leucine content of 22%. The hydropathy profile of this gene suggests that it contains at least five membrane-spanning regions (Fig. 3). The average hydrophobicity value for these segments was greater than 1.6, suggesting that they are all membrane-spanning regions (26). The *fliR* gene product has a calculated molecular mass of 28.5 kDa. No typical rho-independent transcription termination site (37) was observed in the DNA nucleotides following the translation stop codon for the *fliR* gene. However, the DNA region downstream from the distal *Pst*I ($PstI_{II}$) restriction site includes the *rcsA* gene (42, 46). The DNA sequences from bp 2254 to a *Nru*I site 130 bp downstream were 100% identical to the upstream region of the reported *rcsA* gene (42) (data not shown). Since there is no space for additional genes between the termination codon of the *fliR* gene and the start of the *rcsA* gene, we conclude that the *fliR* gene is the last gene in the *fliL* operon.

Comparison of the amino acid sequences of the FliO, FliP, FliQ, and FliR proteins of *E. coli* with sequences of other proteins. The *fliO*, *fliP*, *fliQ*, and *fliR* gene and deduced amino acid sequences were compared with DNA and protein sequences available in data banks. A partial DNA sequence of an unidentified gene located upstream of the *flaA* gene of *Rhizobium meliloti* (36) showed 48 and 64% identities to the *fliP* amino acid and DNA sequences, respectively. Inspection of the *R. meliloti* DNA sequence suggested that it was missing one nucleotide. The addition of one nucleotide at position 271 of the *R. meliloti* sequence, corresponding to position 1040 of the *E. coli* DNA sequence (Fig. 1), would eliminate the early termination of the *R. meliloti* polypeptide and extend the region of homology to the end of the *fliP* gene (Fig. 4). Therefore, we suggest that the corresponding gene is the *fliP* counterpart in *R. meliloti*. This search also revealed very strong homologies between FliP and two other proteins that belong to virulence systems in plant and mammalian pathogens (15, 47). In the first case, 39% identity over a 228-amino-acid overlap was observed between FliP and the ORF2 gene product, which is involved in pathogenicity of *Xanthomonas campestris* pv. glycines 8ra (15) (Fig. 4). It is suggested that the pathogenicity genes in *X. campestris* pv. campestris may be involved in the secretion of proteins from these microorganisms by controlling translocation across the outer membrane (10). Hwang et al. (15) have indicated that the pathogenicity gene, ORF2, of *X. campestris* pv. glycines 8ra was similar to the gamma subunit of oxaloacetate decarboxylase from *Klebsiella pneumoniae*, which is involved in sodium ion transport. However, we could not find

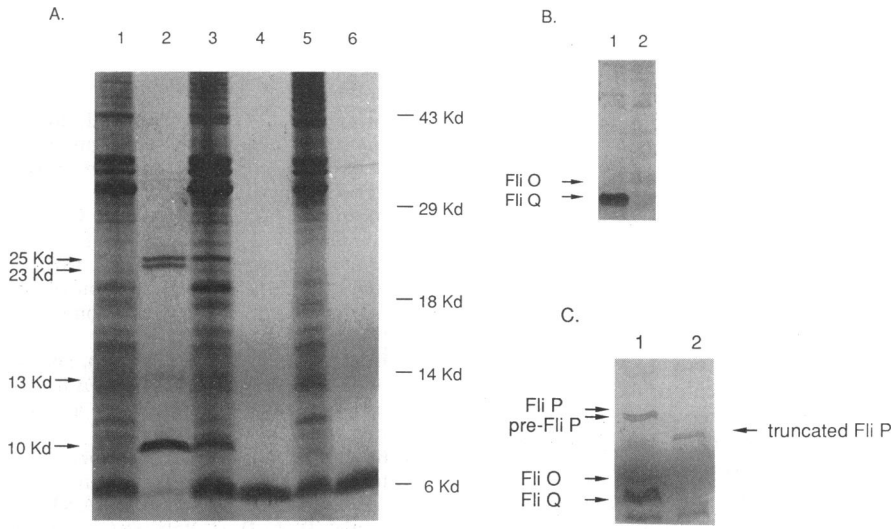

FIG. 5. Autoradiograph of polypeptides labeled with [^{35}S]methionine and resolved by SDS-polyacrylamide gel electrophoresis. Samples were prepared from *E. coli* XL1-Blue containing pGP1-2 plus the indicated plasmids as described in Materials and Methods. (A) Lanes: 1 to 3, samples containing pJM91; 4 and 5, control sample containing pBluescript SK$^+$; 6, control sample containing only pPG1-2. In lanes 1 and 5, cells were labeled without heat induction or rifampin treatment; in lane 3, cells were labeled after heat induction without rifampin treatment; in lanes 2, 4, and 6, cells were heat induced and treated with rifampin prior to labeling. Positions of molecular size markers are given on the right. (B) Expression of pJM131 (lane 1) and pJM111 (lane 2). (C) Expression of pJM91 (lane 1) and pJM171 (lane 2). The insert-specific proteins are indicated.

any significant homology between the oxaloacetate decarboxylase and FliP amino acid sequences. In the second case, 39% identity was shown between the FliP protein and the *Shigella flexneri* Spa24 protein (47). Spa24 is a member of a series of proteins involved in surface presentation of the invasion plasmid antigens. These comparisons suggest that *fliP* may play some role in the transport of flagellar proteins.

Recently, sequences homologous to those of the *E. coli fliP* and *fliQ* genes have been identified in *Bacillus subtilis* (5), and sequences homologous to those of the *fliQ* and *fliR* genes have been identified in *Caulobacter crescentus* (52). Interestingly, a conventional cleavable signal peptide is not present at the N terminus of the *B. subtilis* FliP protein, suggesting a location in the inner membrane for this polypeptide. Lastly, FliO and FliP of *E. coli* and *S. typhimurium* (26a) show 67 and 91% amino acid identity, respectively.

Identification of proteins encoded by *fliO*, *fliP*, *fliQ*, and *fliR*. To determine gene-polypeptide correlation, the 2.2-kb *Pst*I fragment containing the *fliOPQR* genes was cloned in pBluescript SK$^+$ in the orientation whereby the genes would be transcribed via the strong T7 phage promoter. The expression of this plasmid (pJM91) resulted in the synthesis of four insert-specific proteins of 25, 23, 13, and 10 kDa (Fig. 5). To determine which gene corresponded to each of these polypeptides, a number of pBluescript derivative plasmids carrying different portions of the 2.2-kb *Pst*I fragment were constructed. When plasmid pJM131, containing only the *fliQ* gene, was expressed, only one protein band of 10 kDa was observed (Fig. 5). Also, this polypeptide was no longer expressed by a *fliQ*-deleted pJM91 derivative, pJM171. Therefore, the *fliQ* gene encodes the 10-kDa protein. A weakly expressed 13-kDa protein of the predicted molecular weight of the FliO protein was present in pJM111 (Fig. 5B, lane 2) and in pJM121 (data not shown). Plasmid pJM121 contained the 3′ end of the neighboring *fliN* gene and the entire *fliO* gene, and it also had both possible translation initiation codons for the *fliO* gene (see above). In contrast, in pJM111, the region upstream from the *Pst*I$_I$ site containing the *fliN-fliO* intercistronic region and the ATG codon at bp 6 was missing (Table 1). Since the two plasmids encoded polypeptides of the same size, we feel confident that the GTG at position 66 bp is the functional start codon and the 13 kDa protein is the gene product of the *fliO* gene (Fig. 5).

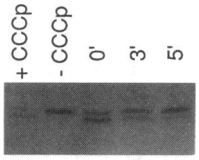

FIG. 6. Processing of the signal sequence of the FliP protein. pJM91-containing cells were labeled in the T7 RNA polymerase expression system as described in Materials and Methods. Cells were labeled for 1 min with [^{35}S]methionine. At this point, 100-μl aliquots were removed and incubation was continued for 30 min with (+CCCP) or without (−CCCP) addition of CCCP (15 mmol/ml) and 2 volumes of L broth containing excess unlabeled methionine. A pulse-chase experiment was performed by addition of 100 μl of the 1-min-labeled sample to 200 μl of L broth containing excess cold methionine and incubation at 35°C for 0 min (0′), 3 min (3′), and 5 min (5′). Samples were centrifuged briefly, suspended in loading buffer, and boiled for 5 min prior to gel electrophoresis.

Plasmid pJM171 is a 3′-deletion product of pJM91 in which the DNA region beyond bp 971 (Fig. 1) is deleted, and the remaining DNA fragment is ligated with the vector DNA. As a result of this ligation, six in-frame amino acid codons are added to the end of the truncated FliP before a stop codon is encountered. The expression of pJM171 resulted in synthesis of the 13 kDa FliO protein and two proteins of 20 and 18 kDa in place of the 25- and 23-kDa proteins synthesized by pJM91. If the 25- and 23-kDa protein bands in pJM91 expression (Fig. 5) were two different forms of the FliP polypeptide (i.e., precursor and mature forms as suggested by DNA sequence analysis), deletion of the 3′ end of the *fliP* gene would be

expected to result in the replacement of both protein bands by smaller proteins. Thus, these results indicated that both the 25- and 23-kDa protein bands were encoded by the *fliP* gene. Although multiple methionine residues were present in the predicted amino acid sequence of FliR, attempts to visualize the *fliR* gene product in minicells or the T7 expression system by using Tran-^{35}S methionine for protein labeling were unsuccessful.

Processing of the FliP protein. It has been shown that processing of outer membrane proteins or periplasmic proteins requires a membrane potential for proper processing of the signal peptides (7, 35). To determine whether the two closely migrating bands at 25 and 23 kDa were products of the signal peptide processing of the *fliP* gene product, we treated the labeled cells with carboxycyanide *m*-chlorophenylhydrazone (CCCP), which inhibits signal sequence processing (7). As shown in Fig. 6, in the presence of CCCP, processing of the precursor protein was prevented. However, instead of the expected accumulation of more slowly migrating unprocessed protein, increased levels of a faster-migrating protein were seen. Similarly, a pulse-chase experiment done in the absence of CCCP resulted in accumulation of increased levels of the more slowly migrating protein over the time course, which was concomitant with loss of the fast-migrating protein (Fig. 6). These data suggested that the fast-migrating protein is the precursor form and the slowly migrating polypeptide is the mature form of the FliP protein.

In addition to performing the CCCP treatment, we treated cells carrying the pGP1-2 and pJM91 with sodium azide prior to pulse-labeling. Sodium azide is an inhibitor of SecA-dependent protein translocation and causes the accumulation of the precursor form of the transported proteins (34). This experiment revealed that FliP precursor processing was inhibited in the presence of sodium azide and that the precursor form of the FliP protein is, in fact, the fast-migrating band under the conditions used for gel electrophoresis (data not shown). The reason for this abnormal migration of the processed and unprocessed forms of the FliP protein remains to be determined.

The existence of an N-terminal signal sequence at the beginning of the FliP protein is the third incidence of a signal sequence in the flagellar system. The only flagellar proteins known previously to have canonical signal sequences were FlgH and FlgI, the structural protein components of the L and P rings, respectively (13, 16). Jones and Macnab (18) have shown that *fliP* and *fliQ* are expressed prior to the expression of any of the basal body or hook proteins. They have also shown that in partial basal body structures containing only the MS ring, two additional FlhD/C-dependent proteins of 23 and 26 kDa are present. It is tempting to suggest that the FliP protein may be a structural component of the basal body, because *fliP* expression is also regulated by the *flhD* and *flhC* gene products, and *fliP* encodes proteins with molecular weights similar to those seen on partial basal body preparations. However, strong homologies between the FliP protein and proteins involved in translocation of the virulence factors of different systems may be an indication of its direct or indirect involvement in the flagellum-specific transport apparatus.

ACKNOWLEDGMENTS

We thank Robert Macnab for sharing DNA sequences prior to publication and Tom Silhavy for helpful discussion concerning the use of sodium azide for inhibition of protein translocation.

This study was supported by PHS grants GM34765 to Bert Ely and AI18985 to Philip Matsumura.

REFERENCES

1. **Adler, J., and B. Templeton.** 1967. The effect of environmental conditions on the motility of *Escherichia coli*. J. Gen. Microbiol. **46**:175–184.
2. **Aizawa, S. I., G. E. Dean, C. J. Jones, R. M. Macnab, and S. Yamaguchi.** 1985. Purification and characterization of the flagellar hook-basal body complex of *Salmonella typhimurium*. J. Bacteriol. **161**:836–849.
3. **Aksoy, S., C. L. Squires, and C. Squires.** 1984. Translational coupling of the *trpB* and *trpA* genes in the *Escherichia coli* tryptophan operon. J. Bacteriol. **157**:363–367.
4. **Ames, G. F.-L.** 1973. Resolution of bacterial proteins by polyacrylamide gel electrophoresis on slabs. J. Biol. Chem. **249**:634–644.
5. **Bischoff, D. S., M. D. Weinreich, and G. W. Ordal.** 1992. Nucleotide sequences of *Bacillus subtilis* flagellar biosynthetic genes *fliP* and *fliQ* and identification of a novel flagellar gene, *fliZ*. J. Bacteriol. **174**:4017–4025.
6. **Brandl, C. J., and C. M. Deber.** 1986. Hypothesis about the function of membrane-buried proline residues in transport proteins. Proc. Natl. Acad. Sci. USA **83**:917–921.
7. **Daniel, D. J., D. G. Bole, S. C. Quay, and D. L. Oxender.** 1981. Role for membrane potential in the secretion of protein into periplasm of *Escherichia coli*. Proc. Natl. Acad. Sci. USA **78**:5396–5400.
8. **DePamphilis, M. L., and J. Adler.** 1971. Fine structure and isolation of hook-basal complex of flagella from *Escherichia coli* and *Bacillus*. J. Bacteriol. **105**:384–395.
9. **Devereux, J., P. Haeberli, and O. Smithies.** 1984. A comprehensive set of sequence analysis programs for the VAX. Nucleic Acids Res. **12**:387–395.
10. **Dow, J. M., G. Scofield, K. Trafford, P. C. Turner, and M. J. Daniels.** 1987. A gene cluster in *Xanthomonas campestris* pv. campestris required for pathogenicity controls the excretion of polygalacturonate lyase and other enzymes. Physiol. Mol. Plant Pathol. **31**:261–271.
11. **Driks, A., and D. J. DeRosier.** 1990. Additional structures associated with bacterial flagellar basal body. J. Mol. Biol. **211**:669–672.
12. **Francis, N. R., V. M. Irikura, S. Yamaguchi, D. J. DeRosier, and R. M. Macnab.** 1992. Localization of the salmonella typhimurium flagellar switch protein FliG to the cytoplasmic M-ring face of the basal body. Proc. Natl. Acad. Sci. USA **89**:6304–6308.
13. **Homma, M., Y. Komeda, T. Iino, and R. M. Macnab.** 1987. The *flaFIX* gene product of *Salmonella typhimurium* is a flagellar basal body component with a signal peptide for export. J. Bacteriol. **169**:1493–1498.
14. **Homma, M., K. Kutsukake, M. Hasebe, T. Iino, and R. M. Macnab.** 1990. FlgB, FlgC, FlgF, and FlgG. A family of structurally related proteins in the flagellar basal body of *Salmonella typhimurium*. J. Mol. Biol. **211**:465–477.
15. **Hwang, I. S., M. Lim, and P. Shaw.** 1992. Cloning and characterization of pathogenicity genes from *Xanthomonas campestris* pv. glycines. J. Bacteriol. **174**:1923–1931.
16. **Jones, C. J., M. Homma, and R. M. Macnab.** 1987. Identification of proteins of the outer (L and P) rings of the flagellar basal body of *Escherichia coli*. J. Bacteriol. **169**:1489–1492.
17. **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 sequence and deduced protein sequences. J. Bacteriol. **171**:3890–3900.
18. **Jones, C. J., and R. M. Macnab.** 1990. Flagellar assembly in *Salmonella typhimurium*: analysis with temperature-sensitive mutants. J. Bacteriol. **172**:1327–1339.
19. **Khan, I. H., T. S. Reese, and S. Khan.** 1992. The cytoplasmic component of the bacterial flagellar motor. Proc. Natl. Acad. Sci. USA **89**:5956–5960.
20. **Komeda, Y.** 1982. Fusion of operons to lactose genes on a Mu *lac* bacteriophage. J. Bacteriol. **150**:16–26.
21. **Komeda, Y.** 1986. Transcriptional control of flagellar genes in *Escherichia coli* K-12. J. Bacteriol. **168**:1315–1318.
22. **Komeda, Y., H. Suzuki, J. Ishidsu, and T. Iino.** 1975. The role of cAMP in flagellation of *Salmonella typhimurium*. Mol. Gen. Genet. **142**:289–298.

23. **Kuo, S. C., and D. E. Koshland, Jr.** 1986. Sequence of the *flaA* (*cheC*) locus of *Escherichia coli* and discovery of a new gene. J. Bacteriol. **166:**1007–1012.
24. **Kutsukake, K., Y. Ohya, and T. Iino.** 1991. Transcriptional analysis of the flagellar regulon of *Salmonella typhimurium*. J. Bacteriol. **172:**741–747.
25. **Kuwajima, G., I. Kawagishi, M. Homma, J.-I. Asaka, E. Kondo, and R. M. Macnab.** 1989. Export of an N-terminal fragment of *Escherichia coli* flagellin by a flagellum-specific pathway. Proc. Natl. Acad. Sci. USA **86:**4953–4957.
26. **Kyte, J., and R. F. Doolittle.** 1982. A simple method for displaying the hydropathic character of a protein. J. Mol. Biol. **157:**105–132.

26a. **Macnab, R. M.** Personal communication.

27. **Macnab, R. M.** 1992. Genetics and biogenesis of bacterial flagella. Annu. Rev. Genet. **26:**131–158.
28. **Macnab, R. M., and J. S. Parkinson.** 1991. Genetic analysis of bacterial flagellum. Trends Genet. **7:**196–200.
29. **Malakooti, J., Y. Komeda, and P. Matsumura.** 1989. DNA sequence analysis, gene product identification, and localization of flagellar motor components of *Escherichia coli*. J. Bacteriol. **171:** 2728–2734.
30. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual. Cold Spring Laboratory, Cold Spring Harbor, N.Y.
31. **Messing, J.** 1983. New M13 vectors for cloning. Methods Enzymol. **101:**20–78.
32. **Miller, J. H.** 1972. Experiments in molecular genetics. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
33. **Oliver, D.** 1985. Protein secretion in *Escherichia coli*. Annu. Rev. Microbiol. **39:**615–648.
34. **Oliver, D. B., R. J. Cabelli, K. M. Dolan, and G. P. Jarosik.** 1990. Azide-resistant mutants of *Escherichia coli* alter the SecA protein, an azide-sensitive component of the protein export machinery. Proc. Natl. Acad. Sci. USA **87:**8227–8231.
35. **Oxender, D. L., R. Landick, P. Nazos, and B. R. Copeland.** 1984. Role of membrane potential in protein folding and secretion in *Escherichia coli*, p. 4–7. *In* L. Leive and D. Schlessinger (ed.), Microbiology—1984. American Society for Microbiology, Washington, D.C.
36. **Pleier, E., and R. Schmitt.** 1989. Identification and sequence analysis of two related flagellin genes in *Rhizobium meliloti*. J. Bacteriol. **171:**1467–1475.
37. **Rosenberg, M., and D. Court.** 1979. Regulatory sequence involved in the promotion and termination of RNA transcription. Annu. Rev. Genet. **13:**319–353.
38. **San Francisco, M. J. D., L. S. Tisa, and B. P. Rosen.** 1989. Identification of the membrane component of the anion pump encoded by arsenical resistance operon of R-factor R773. Mol. Microbiol. **3:**15–21.
39. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74:**5463–5467.
40. **Shine, J., and L. Dalgarno.** 1974. The 3′ terminal sequence of *Escherichia coli* 16 S ribosomal RNA: complementarity to nonsense triplets and ribosome binding sites. Proc. Natl. Acad. Sci. USA **71:**1342–1346.
41. **Silverman, M., and M. Simon.** 1974. Characterization of *Escherichia coli* flagellar mutants that are insensitive to catabolite repression. J. Bacteriol. **120:**1196–1203.
42. **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.
43. **Suzuki, T., T. Iino, T. Horiguchi, and S. Yamaguchi.** 1978. Incomplete structures in nonflagellate mutants of *Salmonella typhimurium*. J. Bacteriol. **133:**904–915.
44. **Suzuki, T., and Y. Komeda.** 1981. Incomplete flagellar structures in *Escherichia coli* mutants. J. Bacteriol. **145:**1036–1041.
45. **Tabor, S., and C. C. Richardson.** 1985. A bacteriophage T7 RNA polymerase/promoter system for controlled exclusive expression of specific genes. Proc. Natl. Acad. Sci. USA **82:**1074–1078.
46. **Torres-Cabassa, A. S., and S. Gottesman.** 1987. Capsule synthesis in *Escherichia coli* K-12 is regulated by proteolysis. J. Bacteriol. **169:**981–989.
47. **Venkatesan, M. M., J. M. Buysse, and E. V. Oaks.** 1992. Surface presentation of *Shigella flexneri* invasion plasmid antigens requires the products of the *spa* locus. J. Bacteriol. **174:**1990–2001.
48. **Vieira, J., and J. Messing.** 1982. The pUC plasmids, and M13mp7-derived system of insertion mutagenesis and sequencing with synthetic universal primers. Gene **19:**259–268.
49. **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.
50. **Yamaguchi, S., S.-I. Aizawa, M. Kihara, M. Isomura, C. J. Jones, and R. M. Macnab.** 1986. Genetic evidence for a switching and energy-transducing complex in the flagellar motor of *Salmonella typhimurium*. J. Bacteriol. **168:**1172–1179.
51. **Yamaguchi, S., H. Fujita, A. Ishihara, S.-I. Aizawa, and R. M. Macnab.** 1986. Subdivision of multifunctional flagellar genes of *Salmonella typhimurium* into regions responsible for assembly, rotation, and switching. J. Bacteriol. **166:**187–193.
52. **Zhuang, W. Y., and L. Shapiro.** Personal communication.