JOURNAL OF BACTERIOLOGY, Nov. 1984, p. 577–585
0021-9193/84/110577-09$02.00/0

Vol. 160, No. 2

Identification of *Escherichia coli* Region III Flagellar Gene Products and Description of Two New Flagellar Genes

DOUGLAS H. BARTLETT AND PHILIP MATSUMURA*

Department of Biological Sciences, University of Illinois at Chicago, Chicago, Illinois 60680

Received 26 April 1984/Accepted 8 August 1984

Region III flagellar genes in *Escherichia coli* are involved with the assembly and rotation of the flagella, as well as taxis. We subcloned the *flaB* operon from a λ*fla* transducing phage onto plasmid pMK2004. Two additional genes were found at the *flaB* locus, and we subdivided the *flaB* gene into *flaB1*, *flaB11*, and *flaB111*. The *cheY* suppressor mutations which have previously been mapped to *flaB* were further localized to *flaB11* (Parkinson et al., J. Bacteriol. 155:265–274, 1983). Until now, gene product identification has not been possible for these genes because of their low levels of gene expression. Overexpression of the flagellar genes was accomplished by placing the *flaB* operon under the control of the *lacUV5* or *tac* promoters. Plasmid-encoded proteins were examined in a minicell expression system. By correlating various deletions and insertions in the *flaB* operon with the ability to complement specific flagellar mutants and code for polypeptides, we made the following gene product assignments: *flaB1*, 60 kilodaltons; *flaB11*, 38 kilodaltons; *flaB111*, 28 kilodaltons; *flaC*, 56 kilodaltons; *fla0*, 16 kilodaltons; and *flaE*, 54 kilodaltons.

The study of bacterial flagella has been useful in understanding such divergent areas as flagellar organelle assembly, the energy transduction involved in conversion of proton motive force to flagellar rotation, and the sensory processes directing taxes to a variety of stimuli, such as chemicals, oxygen, light, and pH (5, 27, 34, 40). The flagellum itself has been divided into three morphologically distinct units. These are the basal body components which anchor the flagellum to the cell, the hook which provides a flexible coupling to the cell, and the flagellar filament itself. The flagellar filament protein (flagellin), hook protein, and some of the proteins found in hook basal body preparations have been correlated with specific flagellar genes (23, 31, 41). In *Escherichia coli*, there are more than 30 genes that have been found to be required for producing the flagella (designated *fla* genes), 3 genes required for flagellar rotation (designated *mot* genes), and 11 genes required for chemotaxis (designated *che* genes) (27). These genes are positioned on the *E. coli* chromosome at three main regions (region I located at 23 min, region II at 43 min, and region III at 43 min). To characterize biochemically the *fla*, *mot*, and *che* systems, gene product identification has been performed. To date, of the 41 genes known for *fla*, *mot*, and *che* functions, product identifications of 16 have been published (17, 23, 25, 30, 31, 41, 42, 45). These have all represented either region I or II genes.

We initiated an analysis of region III. Region III contains genes with *fla*, *mot*, and *che* phenotypes. The *flaA* and *flaB* gene products are believed to be at the interface of the chemotactic machinery and the flagellum structure (8, 44). The mutations of revertants of some chemotactic mutants have been mapped to these loci (35, 36). Nonchemotactic mutations in *flaA* have been isolated and designated *cheC* (44). By analogy with studies in *Salmonella typhimurium*, the *E. coli flaB* gene may acquire Che^- or Mot^- phenotypes (8, 10). A Mot^- mutant suggests a close relationship between the *flaB* gene product and flagellar rotation. Just downstream from the *flaA* gene, another motility gene, *motD*, is present (21). The *flaE* gene is believed to regulate the assembly of the hook structure and the polymerization of flagellin onto the hook (46). The function of the other region III genes is largely unknown, except that they are important in the assembly of the basal body structure.

Hybrid λ phage carrying region I and II genes have been used to program gene product expression in UV-irradiated λ lysogens (41, 42). Although hybrid λ phage carrying region III flagellar genes have been available for some time, no gene production identifications have been made. This is because of the low level of expression of these genes. We subcloned the *flaB* operon from λ*fla* phage onto plasmids. These genes were overexpressed by using high-level promoters, and their gene products were examined in a plasmid-minicell expression system (31). Gene product identifications were made by using in vitro and in vivo recombinant DNA technology. Plasmids with deletions in the *flaB* locus were used in complementation tests, and two new genes were identified.

MATERIALS AND METHODS

Bacteria, plasmids, and phage. The bacteria, plasmids, and phage used in the manipulation of flagellar DNA are listed in Table 1. The *scyB* and flagellar mutants used for complementation analysis are shown in Table 2. Bacteria were grown up in L-broth (9). Solid media contained 1.7% agar, except motility medium, which contained 0.38% agar (Difco Laboratories, Detroit, Mich.). Antibiotics were purchased from Sigma Chemical Co., St. Louis, Mo. The antibiotic concentrations used for the selection of resistant transformant colonies were as follows: penicillin, 100 μg/ml; kanamycin, 50 μg/ml; and chloramphenicol, 25 μg/ml. Incubations were at 37°C except for the complementation of *fla* and *scy* mutants, which were performed at 30°C (43).

The cloning, transformation, and plasmid purification procedures used have been described previously (31). Restriction endonucleases were purchased from Amersham Corp., Arlington Heights, Ill.), or Bethesda Research Laboratories, Gaithersburg, Md. *Pst*I and T4 DNA ligase were purified by the method of Pirrotta and Bickle (37). Endonuclease III, S1 nuclease, and both unphosphorylated and phosphorylated

* Corresponding author.

TABLE 1. *E. coli* K-12 strains, phages, and plasmids

Strain or plasmid	Relevant genotype and/or comments[a]	Source (reference)
E. coli K-12		
χ1488	*minA minB met*	R. Meagher et al. (32)
C600	*lacY*	Appleyard (2)
NK5012	*sulI*	N. Kleckner (20)
Phages		
λ*fla*36Δ26	*flaN flaB flaC flaO flaE flaA* transducing phage	Y. Komeda et al. (22) J. S. Parkinson et al. (36)
λ577	λ b_2 b_{522} cI *O*am::Tn*9*	N. Kleckner (unpublished data)
Plasmids		
pBR322	Cloning vehicle	F. Bolivar et al. (4)
pMK2004	Cloning vehicle	M. Kahn et al. (19)
pRL31	p*lacUV5* expression vehicle	This study
pKK223-3	p*tac* expression vehicle	J. Brosius via D. Figurski (1, 11)

[a] The complete genotypes are given in the references cited.

*Eco*RI linkers were purchased from Bethesda Research Laboratories. In the case of the unphosphorylated linkers, phosphorylation was performed by the method of Maniatis et al. (28). End labeling of the unphosphorylated linkers was accomplished by using polynucleotide kinase obtained from P-L Biochemicals, Milwaukee, Wis., and γ-^{32}P-labeled ATP obtained from New England Nuclear Corp., Boston, Mass.

Plasmid pRL31 was constructed by cloning the 0.6-kilobase (kb) *Eco*RI-to-*Sal*I fragment from the p*lacUV5* vehicle pOP203-3 (15) into pMK2004. The *Bam*HI-to-*Sal*I sites within the tetracycline resistance gene were deleted, and the M13mp8 polylinker was cloned downstream from p*lacUV5* in the *Eco*RI-to-*Pst*I sites (R. Linzmeier, M.S. thesis, University of Illinois at Chicago, Chicago, Ill., 1982).

Phage DNA was extracted by using a sodium dodecyl sulfate extraction procedure (39).

Endonuclease III-S1 nuclease deletions. Our procedure was similar to that used by Roberts et al. (38) and Guo and Wu (16). Plasmid DNA free of single-strand nicks was essential, presumably because of the normal endonuclease activity of S1 nuclease. *Exo*III buffer contained 66 mM Tris (pH 8), 5 mM $MgCl_2$, and 1 mM β-mercaptoethanol. For deletions larger than 1 kb, no salt was added to the buffer. For deletions between 250 base pairs 1 kb long, 50 mM NaCl was added to the buffer, and for deletions less than 250 base pairs long 80 mM NaCl was added. *Exo*III was added to a final concentration of 20 to 30 U/pmol of plasmid DNA. Digestions were carried out at 22 and 37°C. At 22°C in high-salt preparations 0 to 30 min of digestion resulted in 0- to 250-base pair deletions, in medium-salt preparations 30 to 90 min of digestion resulted in 250- to 1,000-base pair deletions, and at 37°C in no-salt preparations 10 to 25 min of digestion resulted in 1- to 2-kb deletions. *Exo*III was inactivated by heating at 65°C for 10 min. DNA was precipitated, suspended in S1 buffer (100 mM sodium acetate, pH 4.6, 300 mM NaCl, 2.0 mM $ZnSO_4$) containing 10 U of S1 nuclease per pmol, and incubated at 18°C for 30 min. S1 nuclease was inactivated by raising the pH to 8.0 and adding EDTA to a concentration of 10 mM.

Tn*9* mutagenesis. Our transposon mutagenesis procedure was similar to that of Fouts and Barbour (14). A 2-ml portion of midlog pDB4/C600 was infected with λ::Tn*9* phage at a multiplicity of infection of 0.1 in λ yeast extract-maltose broth (10 g of tryptone per liter, 2.5 g of NaCl per liter, 2 g of maltose per liter, 0.1 g of yeast extract per liter). After 30 min at 30°C, the cells were diluted to 10 ml in L-broth and incubated with shaking at 30°C for 1 h. The mixture was diluted to 100 ml in L-broth, and selection was maintained for both the plasmid (kanamycin resistance) and the transposon (chloramphenicol resistance). After growth at 37°C for 17 h, plasmid DNA was extracted, transformed into strain C600, and selected for both plasmid and transposon markers. Individual colonies were screened for plasmids bearing Tn*9* insertions (3).

RESULTS

Cloning of region III flagellar genes. The source of the region III flagellar genes was the specialized transducing phage λ*fla*36Δ26. λ*fla*36 was first isolated by Komeda et al. (22) and subsequently was deleted of DNA downstream from *flaA* by Parkinson (33). In agreement with the data of Parkinson (33), we also found this phage capable of complementing *flaN*, *flaB*, *flaC*, *flaO*, *flaE*, and *flaA* mutants to motility and near wild-type behavior on motility agar plates.

λ*fla* phage were grown up to a high titer, and their DNA was extracted. Various restriction fragments from λ*fla*36Δ26 were subcloned onto plasmid vectors and screened for their ability to complement region III flagellar mutants. A 9.8-kb *Sma*I subclone was found to carry the entire *flaB* operon as well as *flaA*. This *Sma*I construction is diagrammed in Fig. 1. The 9.8-kb *Sma*I restriction fragment from λ*fla*36Δ26 was subcloned into the *Sma*I site of the pBR322 derivative pMK2004. The resultant plasmid was designated pDB3. Since λ*fla*36Δ26 has only one additional *Sma*I site compared with wild-type λ phage DNA, some of the *Sma*I fragment contained within pDB3 must contain λ DNA. In fact, a DNA sequence analysis has shown that 1.3 kb of λ DNA is present downstream from the *flaA* locus (J. Malakooti and P. Matsumura, unpublished data). pDB3 served as the parental plasmid for all later constructions.

Figure 1 shows a partial restriction map of pDB3. For Fig. 2, the orientation and positioning of the *flaB* operon was determined by using restriction enzyme-generated deletions. The flagellar genes were localized by correlating specific regions with genetic complementation groups. In addition to the *flaB* operon, the *flaA* gene was also localized by determining that *flaA* complementation activity is contained on the 1.6-kb *Bam*HI fragment.

TABLE 2. *scy* and *fla* mutations

Mutant allele(s)	Parental strain	Source	Reference
scy mutations			
scyB4495, *4500*, *4501*, *4503*, *4504*, *4514*, and *4520*	RP437	J. S. Parkinson et al.	36
fla mutations			
flaN 1860	MS1350	M. Silverman and M. Simon	43
flaB72, *111*, *131*, *264*, *394*, *722*, *915*, *945*, *957*, *1015*, *6208*, *6209*, and *6210*	MS1350		43
flaO1862	MS1350		43
flaE234	MS1350		43
flaN4145	YK410	Y. Komeda et al.	21
flaB4167	YK410		21
flaC4178	YK410		21
flaO4112	YK410		21
flaE4105	YK410		21
flaA4166	YK410		21
flaA4160	YK410		21

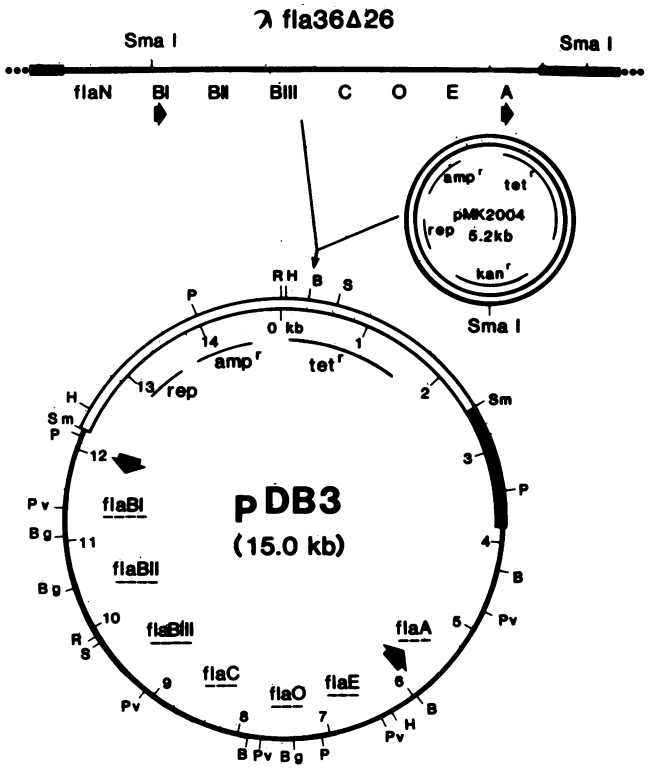

FIG. 1. Construction of pDB3. λ*fla*36Δ26 contains region III flagellar DNA. A *Sma*I construction from λ*fla*36Δ26 into pMK2004 which contained region III DNA was designated pDB3. The thin solid line indicates flagellar DNA, the thick solid line indicates λ DNA, and the thick open line indicates the plasmid vehicle. Dots represent remaining flanking λ DNA sequences. Arrows show the starts and polarities of the *flaB* and *flaA* operons. The numbers shown on pDB3 are size markers (in kilobases). R, *Eco*RI; B, *Bam*HI; S, *Sal*I; Sm, *Sma*I; P, *Pst*I; Pv, *Pvu*II; H, *Hin*dIII; Bg, *Bgl*II.

Two new flagellar genes. The *flaB* locus was subdivided into three complementation groups. By comparing the complementation activity of a variety of restriction enzyme- and nuclease-generated deletions with a collection of *flaB* mutants, it was possible to establish three distinct complementation groups (Fig. 3). Plasmid pDB3 complemented all *flaB* mutants, whereas another construction, pBF1, which contained the entire *flaB* region except for 70 bases upstream from the *Pst*I site, failed to complement any *flaB* mutant. Therefore, either transcriptional signals which are necessary for expression of the *flaB* operon were deleted from pBF1, or a polar mutation in the first gene was made. This placed the start of the *flaB* operon over 2-kb upstream from the *Eco*RI site.

To delineate and define the genes in the *flaB* operon, various deletions within the *flaB* operon was generated. In some cases, deletions were generated in the 5′ coding region of the *flaB* locus. Presumably, such deletions removed the endogenous promoter for the *flaB* operon. However, in these cases transcription was ensured by cloning flagellar DNA next to the *lacUV5* promoter. The *lacUV5* promoter contains an up-promoter mutation in the *E. coli* lactose promoter (15). The vehicle harboring this promoter, pRL31, carries just downstream from the *lacUV5* promoter multiple restriction sites derived from the M13 polylinker present in mp8 (Linzmeier, M.S. thesis). The 9.9-kb-*Eco*RI-to-7.1-kb-

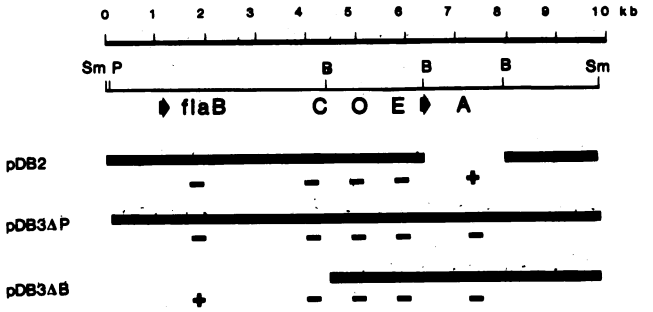

FIG. 2. Localization of flagellar genes. *fla* complementation ability was correlated to the physical map of pDB3 in order to position the genes. The insertion from pDB3 is arranged in a linear fashion with the size markings indicated. pDB2 contains a 1.6-kb *Bam*HI subclone which has *flaA* activity. pDB3ΔP is a *Pst*I deletion of pDB3 which has lost the ability to complement any region III flagellar mutants. pDB3ΔB is a *Bam*HI deletion of pDB3 which retains only *flaB* complementation ability. pDB3ΔB is also capable of recombination in our *flaC* mutants to restore motility. The thick solid lines indicate the extents of the deletions. Sm, *Sma*I; P, *Pst*I; B, *Bam*HI.

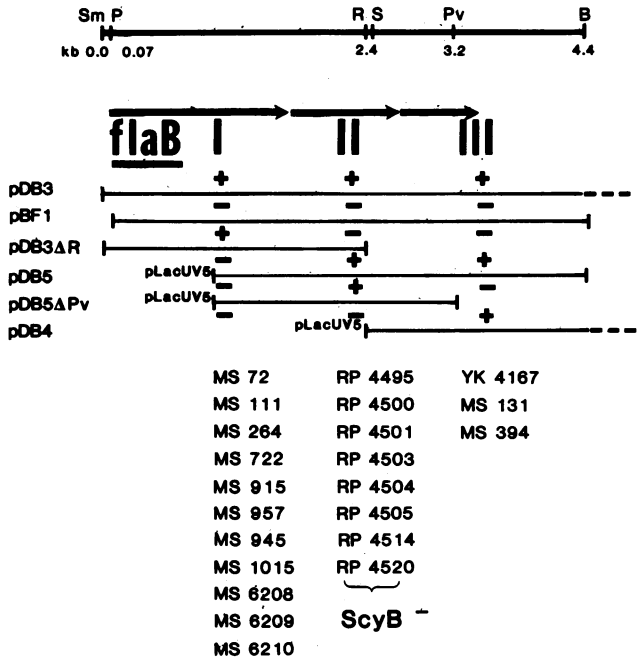

FIG. 3. Identification of the three *flaB* genes. The six plasmids shown on the left were found to complement different combinations of *flaB* mutants. pDB3, the parental plasmid, complements all *flaB* mutants. pBF1 is a 12.2-kb *Pst*I-to-7.8-kb-*Bam*HI subclone from pDB3. Although having removed less than 100 bases from the 5′ end of the *flaB* operon, this plasmid does not complement any *flaB* mutant. pDB3ΔR is an *Eco*RI-generated deletion of pDB3 which complements one set of *flaB* mutants. pDB4 contains *flaB* operon DNA from 9.9 to 7.1 kb on pDB3. pDB4 complements all *flaB* mutants not complemented by pDB3ΔR. pDB5 contains from 11.1 to 7.8 kb from pDB3, and its transcription is mediated by the *lacUV5* promoter. pDB5ΔPv is a *Pvu*II deletion of pDB5 in which 1.2 kb has been removed from the 3′ end of the *flaB* operon insertion. The *Pvu*II deletion of pDB5 only complements *scyB* mutants. Complementation is indicated by a plus sign; no complementation is indicated by a minus sign. Brackets are used to represent the amount of flagellar DNA which each plasmid contains. Sm, *Sma*I; P, *Pst*I; R, *Eco*RI; S, *Sal*I; Pv, *Pvu*II, B, *Bam*HI.

*Pst*I fragment from pDB3 (map units according to Fig. 1) was cloned into pRL31. The resultant plasmid, designated pDB4, is shown in Fig. 3. pDB4 contained all of the *flaB* locus downstream from the *Eco*RI site and was capable of complementing all *flaB* mutants not complemented by an *Eco*RI deletion of pDB3 (for example, YK4167 in Fig. 3). These data suggested that there was at least one additional flagellar gene at the *flaB* locus.

A second additional gene was found when complementation was examined in *scyB* mutants. *scyB* mutants are a class of suppressors of *cheY* mutants (36). The mutations of these pseudorevertants have been mapped to both the *flaB* (*scyB*) and *flaA* (*scyA*) loci (35, 36). *scy* mutants are classified as motile but generally nonchemotactic. Thus, *scyB* mutants are thought to be analogous to the *cheV* mutants which have been isolated from *S. typhimurium* (8). Although pDB3 restored normal chemotactic behavior in all *scyB* mutants, neither the *Eco*RI deletion of pDB3 (pDB3ΔR) nor pDB4 had this ability. Another plasmid, pDB5, which localized the *scyB* complementation function, was constructed. The 12.3-kb-*Hin*dIII-to-9.9-kb-*Eco*RI region from pDB3 was cloned into pBR322, generating pDB3-1. Then an exonuclease III-S1 nuclease procedure was used to produce a range of 5′ deletions within the *flaB* operon (38). pDB3-1 was cleaved at its unique *Hin*dIII site. Then, to produce different deletion sizes, plasmid DNA was subjected to exonuclease III treatment for varying time periods, followed by treatment with S1 nuclease, which generated blunt ends. In the final step, *Eco*RI linkers were cloned into the deletion endpoints. This procedure resulted in deleted plasmids which contained 5′ deletions that ranged in size from 0.6 to 1.1 kb, as well as an additional *Eco*RI site at the deletion endpoints. *flaB* DNA from the plasmid with the largest deletion (1.1 kb) was placed under transcriptional control of the *lacUV5* promoter, and the remaining 3′ portion of the *flaB* loci was rejoined. This plasmid was designated pDB5. pDB5 was capable of complementing *scyB* mutants, as well as that set of *flaB* mutants complemented by pDB4. In addition, a *Pvu*II deletion of pDB5 (pDB5ΔPv) which lost 1.2 kb of flagellar DNA from its 3′ end was only capable of complementing *scyB* mutants. Plasmids pDB3ΔR, pDB5ΔPv, and pDB4 represent three different and overlapping subclones which complemented a particular class of *flaB* or *scyB* mutants. This divides the *flaB* locus into three genes, which we designated *flaB1*, *flaB11*, and *flaB111*.

Overexpression and identification of the *flaB1* and *flaB11* gene products. Further evidence for the existence of three genes at the *flaB* locus was obtained. All genes within the *flaB* operon were overexpressed, and their gene products were identified. Plasmid-encoded gene products were expressed in a plasmid-minicell system and labeled with [^{35}S]methionine. The plasmid-minicell system only allows de novo protein synthesis from the plasmids contained within the minicells. pDB3, which has the *flaB* operon under transcriptional control of its endogenous promoter, was unable to program the synthesis of even radiolabeled amounts of flagellar gene products.

The problem of low gene expression was overcome by using plasmids which expressed flagellar genes from high-level transcriptional fusions. Some of the *lacUV5* constructions described previously were used to enhance expression products. pDB4 expressed *flaB111*, and pDB5 expressed both *flaB11* and *flaB111*. To obtain a plasmid capable of overexpressing *flaB1*, the *Sma*I-to-*Eco*RI region of pDB3 which contained *flaB1* complementation activity was cloned next to the *lacUV5* promoter. This plasmid was designated pDB6. pDB6 contains the endogenous promoter of the *flaB* operon and the *flaB1* gene downstream from the *lacUV5* promoter. pDB6- and pDB5-encoded gene products are shown in Fig. 4. The overexpression vector, pRL31, produced a 30-kilodalton (kd) protein band and a low-molecular-weight protein band (Fig. 5B, lane 2). The 30-kb band corresponds in size to the kanamycin resistance protein of the vector. In addition to these proteins, pDB6 produced a

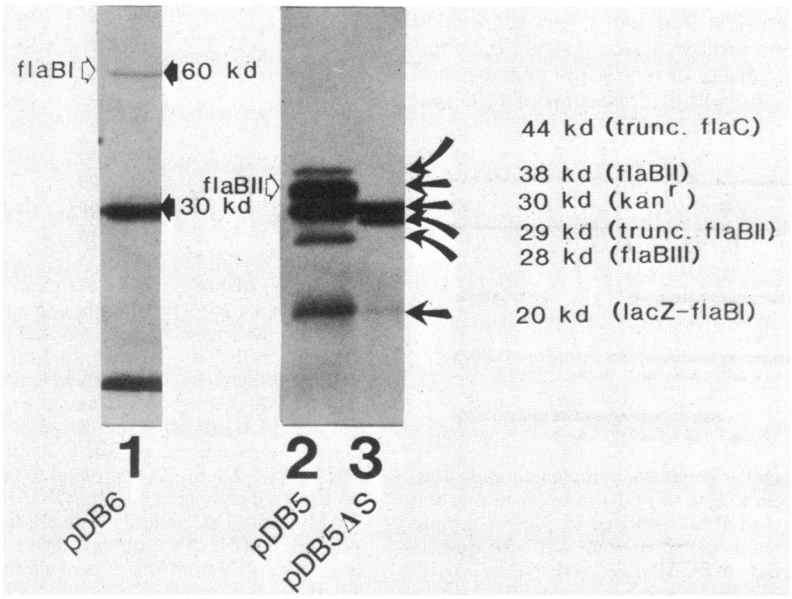

FIG. 4. Identification of the *flaB1* and *flaB11* gene products. Protein expression was monitored in a minicell expression system. pDB6-encoded proteins are shown in lane 1 of the autorad. In addition to the 30-kd and low-molecular-weight vector proteins (see Fig. 5, lane 2), pDB6 also produces a 60-kd protein which represents the *flaB1* gene product. Lane 2 shows pDB5-encoded proteins. These are 44, 38, 30, 28, and 20 kd in size. Lane 3 shows a *Sal*I deletion of pDB5 which removes both *flaB11* and *flaB111* complementation. trunc., Truncated.

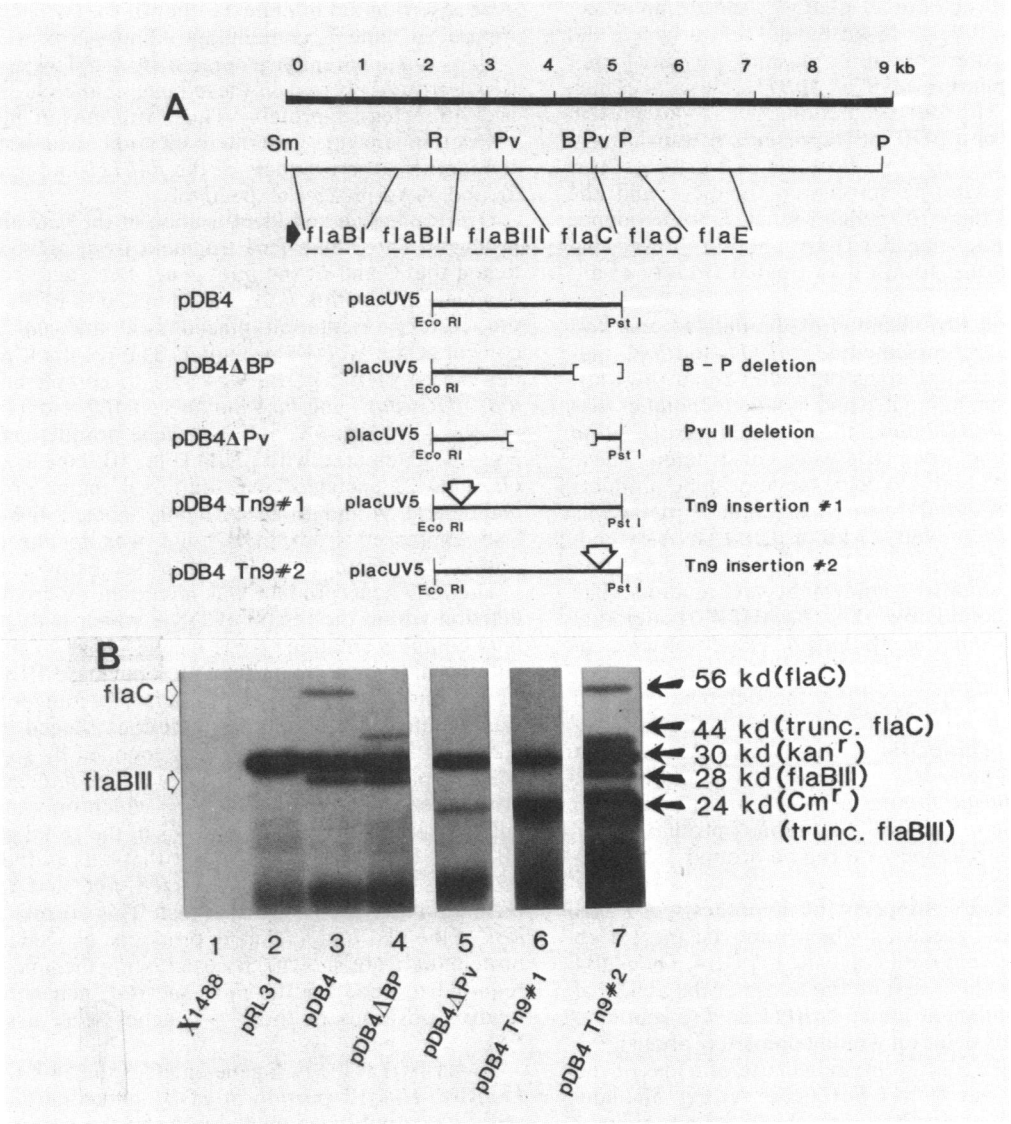

FIG. 5. Identification of the *flaB111* and *flaC* gene products. (A) Plasmid pDB4 and its derivatives which were used for identifying the *flaB111* and *flaC* proteins. The amount of flagellar DNA contained in each plasmid is indicated by the solid line within the brackets. ΔBP signifies a *Bam*HI-to-*Pst*I deletion. The open arrows indicate the positions of the Tn*9* insertions. The lines coming up from the genes to the restriction map better delineate the gene positions. (B) Protein expression of pDB4 and its derivatives. Lane 1 shows the expression of minicell strain λ1488 without any plasmid as a control. Lane 2 shows the vehicle pRL31. The 30-kd band represents the kanamycin resistance protein. Lane 3, pDB4, which produces 56- and 28-kd bands; lane 4, *Bam*HI-to-*Pst*I deletion of pDB4, which produces a truncated (trunc.) *flaC* gene product at 44 kd; lane 5, pDB4 *Pvu*II deletion, which produces a truncated *flaB111* gene product at 24 kd; lane 6, pDB4 Tn*9* insertion 1, which prevents the expression of both *flaB111* and *flaC*; lane 7, pDB4 Tn9 insertion 2, which does not interfere with *flaB111* or *flaC* expression.

60-kd protein (Fig. 4, lane 1). Since pDB6 contains DNA required for *flaB1* complementation activity, we believe that this represents the *flaB1* gene product. DNA sequence analysis has demonstrated an open reading frame coding for a 60-kd protein in this region of DNA (B. Frantz et al., unpublished data). Assuming that 1-kb of DNA contains on the average 37 kd of protein-encoding capacity, then 60 kd of protein corresponds to 1.6 kb of DNA. Since the beginning of the operon has been positioned near the 12.2-kb *Pst*I site by genetic and DNA sequence data, the 3′ end of the *flaB1* gene maps approximately 700 bases upstream from the *Eco*RI site which is in *flaB11*. Thus, the minimum size of the *flaB11* gene is 700 bases or 26 kd of protein-encoding capacity.

Figure 4, lane 2, shows the gene products encoded by pDB5. pDB5 complemented both *scyB* and *flaB111* mutants. In addition to the vector bands, four new proteins appeared; they were 44, 38, 28, and 20 kd in size. The 44- and 28-kd polypeptides represent a truncated *flaC* gene product and the *flaB111* gene product, respectively (see below). Because the 38-kd band was the only protein larger than the 26-kd minimum size estimate of the *flaB11* gene product, we predicted that it represents the *flaB11* gene product. Support for this prediction was obtained by analysis of a deletion in which flagellar DNA 3′ to the *Sal*I site in pDB5 was removed. If the *flaB11* gene product was 38 kd, then this deletion would be expected to produce a truncated *flaB11* protein in the range of 28 kd. Figure 4, lane 3, shows this to

be the case. The *Sal*I deletion of pDB5 (pDB5ΔS) no longer synthesized the 44-, 38-, or 28-kd bands, and a new 29-kd band appeared. Furthermore, this plasmid was no longer capable of complementing *scyB* or *flaB111* mutants. Thus, the results agreed with our prediction. The 20-kd protein expressed in pDB5 and pDB5ΔS represents a translational fusion between the first eight amino acids of β-galactosidase encoded within the *lacUV5* promoter fragment and the carboxy terminus of the *flaB1* gene product. DNA sequence analysis of the promoter-flagellar DNA junction showed that an in-phase translational fusion was created (Frantz et al., unpublished data).

Overexpression and identification of the *flaB111* and *flaC* gene products. pDB4 complemented *flaB111* and *flaC* mutants and expressed two proteins of 56 and 28 kd (Fig. 5B, lane 3). A combination of in vitro and in vivo techniques was used to identify the *flaB111* and *flaC* gene products. Deletion and insertion derivatives of pDB4 were constructed. These are shown in Fig. 5A, with the corresponding gene products shown in Fig. 5B. A *Bam*HI-to-*Pst*I deletion of pDB4 was made. This deletion removed 0.7 kb of pDB4 DNA extending 2.1 to 2.8 kb beyond the *Eco*RI site and resulted in a loss of ability of the plasmid to complement *flaC* mutants. The gene products encoded by this *Bam*HI-*Pst*I deletion (pDB5ΔBP) are shown in Fig. 5B, lane 4. There was a loss of the 56-kd band, which was replaced by a 44-kd protein. As a working hypothesis, we assumed that this deletion extended into the 3′ coding region of the *flaC* gene, removing 12 kd of its coding sequence. This pDB4 deletion then removed about 325 bases of the *flaC* gene downstream from the *Bam*HI site. Since upstream from the *Bam*HI site there must be enough DNA to code for the remaining 44 kd of *flaC* protein, the 5′ end of the gene was mapped to a region around 1.2 kb up from the *Bam*HI site.

Another pDB4 deletion supports this hypothesis. A *Pvu*II deletion of pDB4 was prepared which removed the 1.4 kb from 0.8 to 2.2 kb beyond the *Eco*RI site in pDB4. The pDB4 *Pvu*II deletion (pDB4ΔPv) resulted in a loss of the ability of the plasmid to complement either *flaB111* or *flaC* mutants. We expected that this deletion would completely abolish the synthesis of the larger 56-kd protein because of a loss of the translation initiation site for the *flaC* gene. As Fig. 5B, lane 5, shows, this was true. Furthermore, the 28-kd protein was replaced by a 24-kd one. Apparently, this *Pvu*II deletion extended into the 3′ region of the *flaB111* gene, removing 4 kd of protein-encoding capacity or approximately 100 bases of the *flaB111* coding region. These data suggest that the location of the *flaB111* gene extends from 0.2 to 0.95 kb down from the *Eco*RI site. This agrees well with the positions previously assigned to both the *flaB11* and *flaC* genes.

The location and gene product identification of the *flaB111* and *flaC* genes were verified through the use of in vivo-generated transposon insertions and the chloramphenicol-resistant transposon Tn*9* (6). Two Tn*9* insertions were mapped. One Tn*9* insertion mapped at 300 bases downstream from the *Eco*RI site, and a second Tn*9* insertion was located 2.3 kb downstream from the *Eco*RI site. The first insertion inactivated both *flaB111* and *flaC* activity, whereas the second insertion did not affect either gene. These data were expected since transposon insertions are generally polar on downstream gene expression. Both insertions were consistent with the polarity of the operon going from *flaB111* to *flaC*. The proteins encoded by these two insertion plasmids are shown in Fig. 5B, lanes 6 and 7. Note that the upstream insertion prevented the expression of both the *flaB111* and *flaC* gene products, whereas the downstream gene insertion did not affect either of the two gene products. Figure 5B, lane 7, contains an additional protein at 32 kd, which was apparently generated by the Tn*9* insertion. The 24-kd protein present in these lanes is the chloramphenicol acetyltransferase protein. Thus, transposon insertion and restriction enzyme-generated deletion data both illustrate the fact that the products of the *flaB111* and *flaC* genes are 28 and 56 kd in size, respectively.

Overexpression and identification of the *flaO* and *flaE* gene products. The 3.8-kb *Pst*I fragment from pDB3 which contained the 3′ end of the *flaO* gene, *flaE*, and from the next downstream operon, *flaA*, was cloned into pDB4. A clone in the correct orientation placed *flaO* and *flaE* under the control of the *lacUV5* promoter. The resultant plasmid was designated pDB7. pDB7 was able to complement *flaB111*, *flaC*, *flaO*, *flaE*, and *flaA* mutants. pDB7 and its derivatives are shown in Fig. 6A, and their gene products are shown in Fig. 6B. Compared with pDB4 (Fig. 5B, lane 3), pDB7 (Fig. 6B, lane 2) encoded two additional major protein bands mapping at 54 and 16 kd. A lightly labeled 47-kd band was also expressed from pDB7 and was localized to DNA downstream from the *flaB* operon (see below).

The *flaO* gene product was identified by creating a small deletion within the region of DNA which is responsible for *flaO* complementation. *Bgl*II cleaves once in pDB7. *Bgl*II-cut plasmids were treated with exonuclease III followed by S1 nuclease and religated. As before, a range of deletions was obtained. This time, the deletions ranged in size from 0.2 to 2.4 kb. The gene products encoded by some of these deletion plasmids were investigated. The *flaO* gene was the first to be interrupted. A deletion which only affected *flaO* complementation resulted in a loss of the 16-kd protein (Fig. 6B, lane 1). Thus, we conclude that this is the *flaO* gene product. As more DNA was deleted, *flaC* and *flaE* complementation activities were also lost. This correlated with the loss of the *flaC* and *flaO* gene products, as well as the 54-kd protein (data not shown). By calculating the amount of DNA required to code for the *flaO* and *flaE* gene products, the relative positions of these two genes were assigned (Fig. 5A).

To verify that the *flaE* gene encodes the 54-kd protein, the *Eco*RI-to-*Hin*dIII portion of pDB6 encoding *flaB111*, *flaC*, and *flaO* activity was cloned into the *tac* promoter plasmid pKK223-3. The *tac* promoter contains the −35 region of the *E. coli trp* promoter and the −10 region of the *E. coli lacUV5* promoter, as well as the *lac* operator region (1, 11). This *Eco*RI-to-*Hin*dIII construction was designated pDB8 (Fig. 6A). The minicell products of pDB8 are shown in Fig. 6B, lane 3). The previously designated 28-kd *flaB111* product, the 56-kd *flaC* product, and the 16-kd *flaO* product were all visible. However, the 54-kd product encoded on pDB6 was replaced by a 47-kd band. This was in complete agreement with the expected molecular weight of a truncated *flaE* gene product, as deduced from previous mapping. Thus, the *flaO* and *flaE* gene products were found to be 16 and 54 kd, respectively.

DISCUSSION

In this study, the molecular weights of the *E. coli flaB* operon gene products were determined. Our results are summarized in Fig. 7. Each of the six gene products was identified in the following way. First, the ability to complement different Fla$^-$ mutants was correlated with specific DNA fragments generated by nucleases and restriction enzymes. Second, the molecular weights of overexpressed *flaB* operon proteins and their truncated derivatives were corre-

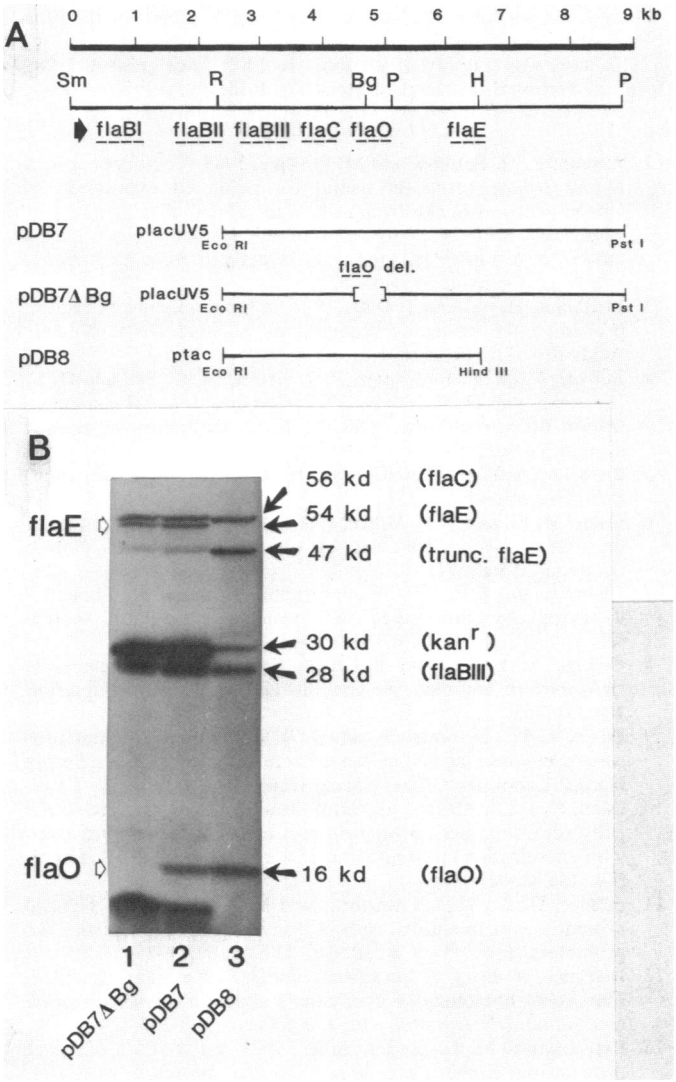

FIG. 6. Identification of the *flaO* and *flaE* gene products. (A) Plasmids used to identify *flaO* and *flaE*. (B) Autorad of the plasmid-encoded proteins. Lane 1 shows that the *flaO* deletion within pDB7 (pDB7ΔBg) results in the loss of the 16-kd protein. Lane 2, pDB7; lane 3, pDB8. Note the truncated (trunc.) *flaE* protein at 47 kd.

lated to the DNA fragments which genetically complemented the different genes in the *flaB* operon. From this analysis, the entire coding capacity from the start of *flaB1* to the start of *flaA* was determined, and we conclude that there is no room for any additional genes besides those mentioned here. Additional evidence for the gene product assignments for the three *flaB* genes will be presented in a subsequent paper by Frantz et al.; there, the complete nucleotide sequence of the *flaB* locus will be presented. The molecular weight of the first gene in the adjacent operon, *flaA*, was previously determined to be 38,000 (D. Clegg and D. E. Koshland, Am. Soc. Biol. Chem. Abstr. 2215, 1983). In our hands, an identical weight for the *flaA* gene product has been found (data not shown).

The parental plasmid pDB3 contains 2.4 kb of insertion DNA downstream from *flaE*. However, only *flaA* complementation activity was found in this region. There was a 47-kd protein which was localized to the region downstream from *flaA*. This was based on the fact that *Exo*III-S1 deletions which remove this DNA in pDB6 also prevent the expression of the 47-kd protein. Komeda has isolated region III DNA from λ*fla*36 phage containing all of region III, as well as a region III hybrid ColE1 plasmid from the Clark-Carbon plasmid bank, and discovered differences in the restriction map from ours downstream from *flaA* (Y. Komeda, personal communication). It is quite possible that in our case some form of genetic rearrangement occurred in this region. In any case, the 47-kd protein downstream from *flaA* in pDB3 was not correlated with any region III flagellar gene product.

The principle overexpression vector used, pRL31, contains the *lacUV5* promoter (15). In addition to making transcriptional fusions, this vehicle is also capable of producing translational fusions between the first eight amino acids of β-galactosidase encoded by the *lacUV5* promoter fragment and any downstream gene in the correct translational phase. The generation of fusion proteins could have complicated gene product analysis; however, the DNA sequence of these constructions indicated that only one of the flagellar overexpression plasmids is in the correct phase to produce such fusions (Frantz et al., unpublished data). This plasmid, pDB5, produced a *lacZ-flaB1* fusion protein of the predicted molecular weight (20,000). With pDB5 and pDB5ΔS, the 20-kd *lacZ-flaB1* fusion protein was not confused with *fla* gene products. Therefore, none of the gene product assignments reflects an artifact of the overexpression system which was used.

In a previous study, *E. coli* region III flagellar gene products were expressed from λ*fla* phase in UV-irradiated λ lysogens labeled with ^{14}C-labeled amino acids (24). A transducing phage carrying the *flaN*, *flaB*, and *flaA* operons was reported to program the synthesis of six proteins that were 48, 40, 38, 36, 30, and 27 kd in size. However, no data were shown. In our minicell translation system without overexpressing the *flaB* operon genes, no proteins significantly expressed above background were evident.

Two new gene assignments have been made, and the *flaB* locus has been divided into the *flaB1*, *flaB11*, and *flaB111* genes. A similar situation exists in the analogous locus in the closely related bacterium *S. typhimurium*. There seems to be virtually complete functional homology between the *fla*, *mot*, and *che* genes of *E. coli* and *S. typhimurium* (12, 26). The analogous locus in *S. typhimurium* is *flaA11* (10). Yamaguchi has recently divided the *flaA11* gene into *flaA11.1* and *flaA11.2*, with the *flaA11.2* gene possessing *che* and *mot* as well as *fla* phenotypes (10; S. Yamaguchi,

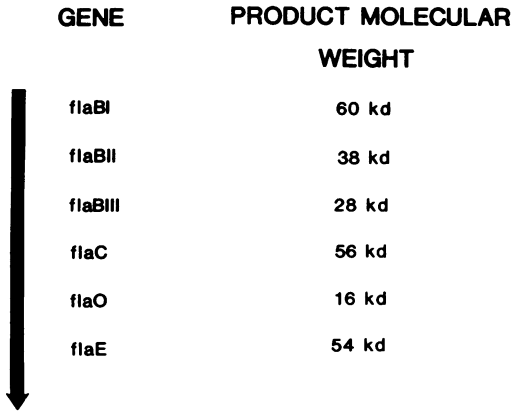

FIG. 7. *E. coli flaB* operon genes and gene products.

unpublished data). Based on our results, we expect a third gene in *S. typhimurium* corresponding to the *flaB111* gene in *E. coli*

Silverman et al. have previously placed the position of an *Eco*RI site to the right of *flaB* (45). This *Eco*RI site is actually within *flaB11*. Our results indicate that 11 of 14 tested *flaB* mutations lie to the left of the *Eco*RI site, but 3 of 14 lie to the right of this site. The *flaB1* gene is twice the size of the *flaB111* gene, yet it contains four times as many mutations as *flaB111*. Because our sample size was so small, no conclusions can be made regarding mutation frequencies between these two genes. However, it has been noted that the frequency of obtaining *flaB* mutants relative to other flagellar mutants is quite high (M. Silverman and J. S. Parkinson, personal communication). This can now be explained by the large total target size of the three *flaB* genes (3.3 kb).

The mutations of the nonchemotactic *scyB* mutants described by Parkinson et al. all map to *flaB11* (36). This suggests that the *flaB11* gene product interacts with the *cheY* gene product in the sensory or adaptation process of chemotaxis or both. A similar type of interaction has also been proposed between the *flaA* and *cheY* gene products (35). It is interesting to note that the *flaB11* and *flaA* proteins have also been suggested to interact with the product of the *cheZ* gene and are similar in size and the fact that both genes may acquire Che^- phenotypes with either counterclockwise or clockwise rotational biases. Even though the mutations of only *scyB* mutants have mapped within the *flaB11* gene, we have given this gene a *fla* designation rather than a *che* designation because of the previous *flaB* assignment to this locus. A null phenotype in *flaB11* of Fla^- has yet to be demonstrated.

flaE mutants produce polyhook structures with no flagellar filaments (18, 46). A current model of flagellar assembly predicts that the *flaE* protein resides in the hook, where it is believed to regulate hook length and initiate flagellin polymerization (46). We have found that radiolabeled *flaE* protein synthesized in minicells is localized totally in the cytoplasmic fraction (unpublished data). It is possible that the *flaE* gene product is site limited and present on the hook structure in very low quantities and that our localization into the cytoplasmic fraction is an artifact of abnormal overexpression in minicells. Alternatively, the *flaE* protein may normally be cytoplasmic and function to regulate gene expression or the secretion of other flagellar proteins.

The identification and overexpression of region III flagellar gene products allow several biochemical questions to be asked. Can any protein-protein interactions be demonstrated? Purified *cheY*, hook, and flagellin proteins are now available (13, 29, 41). The *flaB11* protein may bind purified *cheY*, and *flaE* protein may interact with both hook and flagellin proteins. On the other hand, *flaE* protein may be found to be a DNA-binding protein. These region III products may be examined for specific modifications, reversible or irreversible, and processing. Also, these gene products may be localized to various subcellular fractions. This localization may also be investigated in different flagellar and motility mutants. In this way a hierarchy of assembly order for these proteins within the hook basal body structure may be ascertained. Some of these experiments are now in progress.

ACKNOWLEDGMENTS

We are grateful to J. S. Parkinson, Y. Komeda, M. Silverman, and M. Simon for certain bacterial strains, to J. S. Parkinson and Nancy Kleckner for phage strains, and to D. Figurski for the *ptac* expression vehicle pKK223-3.

This work was supported by Public Health Service grant A118985 from the National Institutes of Health to P.M.

LITERATURE CITED

1. **Amann, E., J. Brosius, and M. Ptashne.** 1983. Vectors bearing a hybrid *trp-lac* promoter useful for regulated expression of cloned genes in *Escherichia coli*. Gene **25**:167–178.
2. **Appleyard, R.** 1954. Segregation of new lysogenic types during growth of a doubly lysogenic strain derived from *Escherichia coli* K-12. Genetics **39**:440–452.
3. **Birnboim, H. C., and J. Doly.** 1979. A rapid alkaline extraction procedure for screening recombinant plasmid DNA. Nucleic Acids Res. **7**:1513–1518.
4. **Bolivar, F., R. L. Rodriguez, P. J. Green, M. C. Betlach, H. L. Heyneker, and H. W. Boyer.** 1977. Construction and characterization of new cloning vehicles. II. A multipurpose cloning system. Gene **2**:95–113.
5. **Boyd, A., and M. Simon.** 1982. Bacterial chemotaxis. Annu. Rev. Physiol. **44**:501–517.
6. **Calos, M. P., and J. H. Miller.** 1980. Molecular consequences of deletion formation mediated by the transposon Tn9. Nature (London) **285**:38–41.
7. **Chamberlain, J. P.** 1979. Fluorographic detection of radioactivity in polyacrylamide gels with water soluble fluor, sodium salyclate. Anal. Biochem. **98**:132–135.
8. **Collins, A. L. T., and B. A. D. Stocker.** 1976. *Salmonella typhimurium* mutants in general chemotaxis. J. Bacteriol. **128**:754–761.
9. **Davis, R. W., D. Bottstein, and J. R. Roth.** 1980. A manual for genetic engineering and advance bacterial genetics. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
10. **Dean, G. E., S. Aizawa, and R. B. Macnab.** 1983. *flaA11* (*motC*, *cheV*) of *Salmonella typhimurium* is a structural gene involved with energization and switching of the flagellar motor. J. Bacteriol. **154**:84–91.
11. **deBoer, H. A., L. J. Comstock, and M. Vasser.** 1983. The *tac* promoter is a functional hybrid derived from the *trp* and *lac* promoters. Proc. Natl. Acad. Sci. U.S.A. **80**:21–25.
12. **Defranco, A. L., J. S. Parkinson, and D. E. Koshland, Jr.** 1979. Functional homology of chemotaxis genes in *Escherichia coli* and *Salmonella typhimurium*. J. Bacteriol. **139**:107–114.
13. **Depamphilis, M. L., and J. Adler.** 1971. Purification of intact flagella from *Escherichia coli* and *Bacillus subtilis*. J. Bacteriol. **105**:376–383.
14. **Fouts, K. E., and S. D. Barbour.** 1981. Transductional mapping of *ksgB* and a new Tn5-induced kasugamycin resistance gene, *ksgD*, in *Escherichia coli*. J. Bacteriol. **145**:914–919.
15. **Fuller, F.** 1982. A family of cloning vectors containing the *lacUV5* promoter. Gene **19**:43–54.
16. **Guo, L., and R. Wu.** 1982. New rapid methods for DNA sequencing based on exoIII digestion followed by repair synthesis. Nucleic Acids Res. **10**:2065–2069.
17. **Hilmen, M., and M. Simon.** 1976. Motility and the structure of bacterial flagella, p. 35–45. *In* R. Goldman, T. Pollard, and J. Rosenbaum (ed.), Cell motility. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
18. **Homma, M., K. Kutsukake, T. Iino, and S. Yamaguchi.** 1984. Hook-associated proteins essential for flagellar filament formation in *Salmonella typhimurium*. J. Bacteriol. **157**:100–108.
19. **Kahn, M., R. Kolter, C. Thomas, D. Figurski, R. Meyer, E. Remaut, and D. R. Helinski.** 1979. Plasmid cloning vehicles derived from plasmids, colE1, F, R6K, and RK2. Methods Enzymol. **68**:268–280.
20. **Kleckner, N.** 1978. DNA sequence analysis of Tn10 insertions: origin and role of 9bp flanking repetitions during Tn10 translocation. Cell **16**:711–720.
21. **Komeda, Y., K. Kutsukake, and T. Iino.** 1980. Definition of additional flagellar genes (*fla*) of *Escherichia coli* K-12. Genetics **94**:277–290.
22. **Komeda, Y., K. Shimada, and T. Iino.** 1977. Isolation of a specialized transducing bacteriophage for flagellar genes (*fla*) of

Escherichia coli K-12. J. Virol. **22**:654–661.
23. **Komeda, Y., M. Silverman, P. Matsumura, and M. Simon.** 1978. Genes for the hook-basal body proteins of the flagellar apparatus in *Escherichia coli*. J. Bacteriol. **134**:655–667.
24. **Kondoh, H.** 1977. Isolation and characterization of nondefective transducing lambda bacteriophage carrying *fla* genes of *Escherichia coli* K-12. J. Bacteriol. **130**:736–745.
25. **Kondoh, H., B. R. Paul, and M. M. Howe.** 1980. Use of λpMu bacteriophages to isolate λ specialized transducing bacteriophages carrying genes for bacterial chemotaxis. J. Virol. **35**:619–628.
26. **Kutsukake, K., T. Iino, Y. Komeda, and S. Yamaguchi.** 1980. Functional homology of *fla* genes between *Salmonella typhimurium* and *Escherichia coli*. Mol. Genet. Gene. **178**:59–67.
27. **Macnab, R., and S. Aizawa.** 1984. Bacterial motility and the bacterial flagellar motor. Annu. Rev. Biophys. Bioenerg. **13**:51–83.
28. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning, a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
29. **Matsumura, P., J. J. Rydell, R. Lindzmeier, and D. Vacante.** 1984. Overexpression and sequence of the *Escherichia coli cheY* gene and biochemical activities of the *cheY* protein. J. Bacteriol. **160**:36–41.
30. **Matsumura, P., M. Silverman, and M. Simon.** 1977. Synthesis of *mot* and *che* gene products of *Escherichia coli* programmed by hybrid colE1 plasmids in minicells. J. Bacteriol. **132**:996–1002.
31. **Matsumura, P., M. Silverman, and M. Simon.** 1977. Cloning and expression of the flagellar hook gene on hybrid plasmids in minicells. Nature (London) **265**:758–760.
32. **Meagher, R. B., R. C. Tait, M. Betlach, and H. W. Boyer.** 1977. Protein expression in minicells by recombinant plasmids. Cell **10**:521–536.
33. **Parkinson, J. S.** 1981. Genetics of bacterial chemotaxis, p. 265–290. *In* S. W. Glover and D. A. Hopwood (ed.), Genetics as a tool in microbiology. Cambridge University Press, London.
34. **Parkinson, J. S., and G. L. Hazelbauer.** 1983. Bacterial chemotaxis: molecular genetics of sensory transduction and chemotactic gene expression, p. 293–318. *In* J. Beckwith, J. Davies, and J. Gallant (ed.), Gene function in prokaryotes. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
35. **Parkinson, J. S., and S. R. Parker.** 1979. Interaction of the *cheC* and *cheZ* gene products is required for chemotactic behavior in *Escherichia coli*. Proc. Natl. Acad. Sci. U.S.A. **76**:2390–2394.
36. **Parkinson, J. S., S. R. Parker, P. B. Talbert, and S. E. Houts.** 1983. Interactions between chemotaxis genes and flagellar genes in *Escherichia coli*. J. Bacteriol. **155**:265–274.
37. **Pirotta, V., and T. A. Bickle.** 1980. General purification schemes for restriction endonucleases. Methods Enzymol. **65**:89–95.
38. **Roberts, T. M., R. Kacich, and M. Ptashne.** 1979. A general method for maximizing gene expression of a cloned gene. Proc. Natl. Acad. Sci. U.S.A. **76**:760–764.
39. **Schleif, R. F., and P. C. Wensink.** 1981. Practical methods in molecular biology. Springer-Verlag, New York.
40. **Silverman, M.** 1980. Building bacterial flagella. Q. Rev. Biol. **55**:395–408.
41. **Silverman, M., P. Matsumura, R. Draper, S. Edwards, and M. Simon.** 1977. Expression of flagellar genes carried by bacteriophage lambda. Nature (London) **261**:248–250.
42. **Silverman, M., P. Matsumura, and M. Simon.** 1976. The identification of the *mot* gene product with *Escherichia coli*-lambda hybrids. Proc. Natl. Acad. Sci. U.S.A. **73**:3126–3130.
43. **Silverman, M., and M. Simon.** 1972. Flagellar assembly mutants in *Escherichia coli*. J. Bacteriol. **112**:986–993.
44. **Silverman, M., and M. Simon.** 1973. Genetic analysis of bacteriophage Mu-induced flagellar mutants in *Escherichia coli*. J. Bacteriol. **116**:114–122.
45. **Silverman, M., and M. Simon.** 1976. Identification of polypeptides necessary for chemotaxis in *Escherichia coli*. J. Bacteriol. **130**:1317–1325.
46. **Suzuki, T., and T. Iino.** 1981. Role of the *flaB* gene in flagellar hook formation in *Salmonella* spp. J. Bacteriol. **148**:973–979.