JOURNAL OF BACTERIOLOGY, June 1984, p. 897–904
0021-9193/84/060897-08$02.00/0

Vol. 158, No. 3

Physical Mapping and Complementation Analysis of Transposon Tn*5* Mutations in *Caulobacter crescentus*: Organization of Transcriptional Units in the Hook Gene Cluster

NORIKO OHTA,[1] EVRYLL SWANSON,[1] BERT ELY,[2] AND AUSTIN NEWTON[1]*

Department of Biology, Princeton University, Princeton, New Jersey 08544,[1] and Department of Biology, University of South Carolina, Columbia, South Carolina 29208[2]

Received 24 January 1984/Accepted 22 March 1984

Using the cloned DNA from the hook protein gene region of *Caulobacter crescentus* (Ohta et al., Proc. Natl. Acad. Sci. U.S.A. **79**:4863–4867, 1982), we have identified and physically mapped 19 Tn*5*-induced and 2 spontaneous insertion mutations to this region of the chromosome. These nonmotile mutants define a major cluster of *fla* genes that covers approximately 17 kilobases on the chromosome (hook gene cluster). Complementation analysis of the mutants using DNA fragments from the region subcloned in the broad host range plasmid pRK290 has shown that these *fla* genes are organized into at least five transcriptional units (I to V). Transcriptional unit II contains at least one gene in addition to the hook protein gene, which makes this the first operon described in *C. crescentus*. Expression of the hook protein gene and the genetically unlinked flagellin A and B genes by this set of mutants also furnishes additional insights into the hierarchial regulation of flagellar genes. We have found that the spontaneous insertion mutant (SC511) of the hook protein gene (*flaK*) makes no flagellin A or B and that genes downstream from the hook protein gene are required in *trans* for expression of the hook protein operon and the flagellin A and B genes. Recombination and complementation results thus place *flaK*, *flaJ*, *flaN*, and *flaO* (R. C. Johnson and B. Ely, J. Bacteriol. **137**:627–634, 1979) in the hook gene cluster, identify at least three new genes (*flbD*, *flbG*, and *flbF*), and suggest that this cluster may contain several additional, as yet unidentified, *fla* genes.

Recently, Johnson and Ely (12) have isolated and classified a set of spontaneous motility mutants in the gram-negative, aquatic bacterium *Caulobacter crescentus* that define at least 24 genes involved in flagellum biosynthesis (13). The organization, expression, and mapping of the flagellar genes in *C. crescentus* are of particular interest because of the unusual pattern of temporal and spatial control of flagellum formation during development of these cells. Cell division in *C. crescentus* is asymmetric, with a nonmotile stalked cell dividing repeatedly to produce a motile swarmer cell with a single polar flagellum. The flagellum is subsequently shed, and a membraneous stalk is formed at the point of flagellum attachment (20a, 26).

Flagellum formation occurs at a discrete time in the cell cycle at the stalk-distal cell pole, and its biosynthesis is regulated in part at the level of gene expression (18, 20, 28). The major flagellar proteins, flagellin A (25 kilodaltons [kd]), flagellin B (27 kd), and the hook protein (70 kd) are synthesized periodically in the cell cycle, and their induction, like the assembly of the DNA bacteriophage receptor sites at the cell pole (9), is coupled to DNA synthesis (28). One goal of the research described here is to understand the coordination between gene expression and the localized assembly of flagellar subunits at the cell pole.

Important tools in the study of flagellar gene organization and expression in *C. crescentus* are the cloned sequences for the hook protein (21), flagellin (19), and *flaE* genes (23) that have been recently isolated. In the work reported here we used the cloned hook protein gene as a probe to identify and physically map 19 Tn*5* insertion mutations located in this region of the *C. crescentus* chromosome. These nonmotile mutants define a major cluster of *fla* genes (hook gene cluster) within a 17-kilobase (kb) stretch of DNA that contains the hook protein gene. DNA fragments from this region of the chromosome were subcloned in the broad host range vector pRK290 (4) and used in a complementation analysis of the insertion mutations. This analysis has shown that the hook gene cluster is organized into at least five transcription units, three of which act in *trans* to regulate expression of the hook protein gene and the genetically unlinked flagellin genes. A preliminary report of these studies appeared earlier (N. Ohta, E. C. Swanson, B. Ely, and A. Newton, Fed. Proc. **42**:2037, 1983).

MATERIALS AND METHODS

Strains and culture conditions. Motility mutants (Table 1) were isolated by Tn*5* mutagenesis (5) or UV irradiation (27) or as spontaneous mutants (12). For the preparation of DNA, strain CB15 and its derivatives were grown in a peptone-yeast extract medium (22). *Escherichia coli* HB101 was used as a host in transformation. Wide host range cloning vector pRK290 and its helper pRK2013 have been described by Ditta et al. (4) and were supplied by D. Helinski. The plasmids were maintained in *E. coli* HB101. Triparental crosses to obtain a *C. crescentus* strain carrying hybrid pRK290 were similar to those described by Ditta et al. (4) for transfer to *Rhizobium* sp. *E. coli* strains were grown in ML medium (1% tryptone, 0.5% yeast extract, 0.5% NaCl, 0.02% $MgSO_47H_2O$).

Isolation of *Caulobacter* DNA and plasmid DNA. Genomic DNAs from CB15 and motility mutants for Southern analysis (31) were prepared from exponentially growing cells. Cells were washed once in 100 mM Tris–10 mM EDTA, pH 8,

* Corresponding author.

TABLE 1. Bacterial strains used in this study

Strain[a]	Genotype	Source or reference
C. crescentus		
CB15	Wild type	ATCC 19089
SC259	*flaW120*	Spontaneous (12)
SC266	*flaK124*[b]	Spontaneous (12)
SC269	*flaJ125*	Spontaneous (12)
SC270	*flaI126*	Spontaneous (12)
SC272	*flaN128*	Spontaneous (12)
SC284	*flaH135*	Spontaneous (12)
SC290	*flaO138*	Spontaneous (12)
SC293	*flaV140*	Spontaneous (12)
SC297	*flaW143*	Spontaneous (12)
SC298	*fla-144*[c]	Spontaneous (12)
SC300	*flaK146*[d]	Spontaneous (12)
SC307	*flaQ150*	Spontaneous (12)
SC511	*flaK155*	Spontaneous (12)
SC1028	*fla-188*::Tn*5 str-152*	This study
SC1032	*fla-198*::Tn*5 str-152*	This study
SC1038	*fla-166*::Tn*5 proA103 str-140*	This study
SC1042	*fla-601*::Tn*5 proA103 str-140*	This study
SC1043	*fla-616*::Tn*5 proA103 str-140*	This study
SC1045	*fla-169*::Tn*5 proA103 str-140*	This study
SC1049	*fla-612*::Tn*5 proA103 str-140*	This study
SC1052	*fla-172*::Tn*5 proA103 str-140*	This study
SC1054	*fla-173*::Tn*5 proA103 str-140*	This study
SC1060	*fla-176*::Tn*5 proA103 str-140*	This study
SC1061	*fla-177*::Tn*5 proA103 str-140*	This study
SC1062	*flaE178*::Tn*5 proA102 str-140*	23
SC1114	*fla-607*::Tn*5 str-152*	This study
SC1132	*fla-608*::Tn*5 str-152*	This study
SC1134	*fla-603*::Tn*5 str-152*	This study
SC1135	*fla-602*::Tn*5 str-152*	This study
SC1166	*fla-609*::Tn*5 str-152*	This study
SC1582	*hunA106*::Tn*5 str-152*	This study
SC1585	*hunB109*::Tn*5 str-152*	This study
SC1721	*cysB102 aar-127 str-196*	Spontaneous (6)
PC5001	*fla-311*::Tn*5*	This study
PC5005	*fla-315*::Tn*5*	This study
PC5007	*fla-317*::Tn*5*	This study
PCM103	*flaJ303*	UV induced (28)
E. coli		
HB101	*hsdR hsdM pro leu thr lacY endI recA str*	
HB101(pRK290)	HB101 carrying pRK290, Tet^r	
HB101(pRK2013)	HB101 carrying pRK2013, Kan^r	

[a] SC strains were isolated in the laboratory of B. Ely; PC strains were isolated in the laboratory of A. Newton.
[b] Formerly *flaM124* (12).
[c] Formerly *flaK144* (12).
[d] Formerly *flaL146* (12).

resuspended in the buffer, and lysed by the addition of lysozyme (final concentration, 400 μg/ml) and Sarkosyl (0.2%). Lysates were then digested with pronase (1 mg/ml) for 2 h at 37°C, followed by chloroform-phenol (1:1) extractions, and DNA was precipitated twice with ethanol. Plasmid DNA was prepared by the cleared lysate technique (15) with 0.1% (wt/vol) Triton X-100 (final concentration) or by an alkaline lysis procedure (17).

Construction of pRK290 hybrids. DNA from λL47.1 Cc8861 or Cc10112 (21; unpublished data) was partially digested with *Bam*HI, inserted into pRK290 at the *Bgl*II site, and ligated with T4 ligase. λL47.1 Cc8855 DNA was partially restricted with *Eco*RI and ligated into the *Eco*RI-digested pRK290 by the same procedure. Hybrid DNA was used to transform *E. coli* HB101 to tetracycline resistance, and clones containing hook region DNA were identified by colony hybridization (8) using appropriate nick-translated fragments of hybrid clone DNA as probes. The identity of hybrid plasmids was confirmed by DNA restriction analysis.

Enzymes. Restriction nucleases and T4 ligase were purchased from Bethesda Research Laboratory, whose instructions were followed for incubation conditions. *E. coli* polymerase I used in nick translation (24) was obtained from New England Biolabs.

Radioimmunoassays and quantitation. Radioimmunoassays for the rates of flagellin and hook protein synthesis were carried out as described previously (27). Cultures were grown in M2 (11) minimal salts medium plus 0.2% glucose to early log phase and labeled for 9 min with [^{35}S]methionine (30 μCi/ml; 1,050 to 1,500 Ci/mmol; Amersham Corp.) without added carrier. Autoradiograms were traced with a Joyce-Loebl densitometer, and the area corresponding to

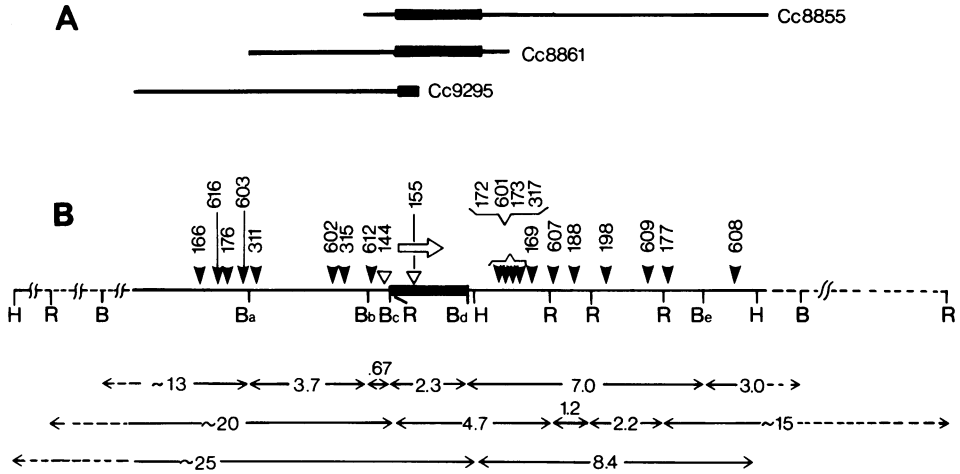

FIG. 1. Physical mapping of nonmotile Tn5 and spontaneous insertion mutations. (A) Overlapping hybrid λL47.1 clones used as probes in Southern hybridization (see text). The solid line represents the cloned insert DNA from *C. crescentus*, and the heavy solid line shows the position of the hook protein structural gene (21). (B) Restriction map and the location of insertion mutants of the hook gene cluster. The broken line represents genomic DNA not present on hybrid clones, the solid line represents cloned segments of the hook cluster, and the heavy solid line represents the hook protein gene; the direction of transcription is indicated by the arrow. Positions of Tn5 insertions are shown by solid arrows, and the positions of spontaneous insertion mutations are shown by open arrows. Distances shown are in kb. The orientation of the restriction map is reversed from that published previously in reference 21. B, *Bam*HI; H, *Hin*dIII; R, *Eco*RI.

each protein band was determined with a peak integrator (MOPIII; Zeiss).

RESULTS

Identification of Tn5 insertion mutations in the hook gene cluster. The structural gene for the hook protein of *C. crescentus* is contained on a 2.3-kb *Bam*HI fragment (Fig. 1), and using overlapping λL47.1 hybrid clones containing this fragment it is possible to probe approximately 50 kb of genomic DNA (21). We have used this approach to identify and physically map Tn5-induced *fla* mutants with insertions in the hook gene cluster. Genomic DNA from 54 mutants was restricted with *Eco*RI, and the fragments were separated by electrophoresis on 0.8% agarose gels, transferred to nitrocellulose filters (31), and probed with nick-translated DNA from hook protein gene clones Cc8855 and Cc8861 (Fig. 1). Five *Eco*RI fragments of ca. 20, 4.7, 1.2, 2.2, and 15 kb were identified in genomic DNA of wild-type cells, and 19 of the 54 mutants screened showed an altered mobility in one of these *Eco*RI fragments. Since Tn5 contains no *Eco*RI sites, the increased mobilities can all be explained by insertion of the 5.7-kb Tn5 element (14). Typical Southern blots of genomic DNA showing the positions of Tn5 insertions in *fla-608*::Tn5, *fla-602*::Tn5, and *fla-173*::Tn5 are shown in Fig. 2A.

The positions of the Tn5 insertions within *Eco*RI fragments were determined by digesting the genomic DNA with *Hin*dIII and *Bam*HI and hybridizing with the same probes. Tn5 has one *Hin*dIII site in each of its inverted repeats (14), and its insertion in a genomic *Hin*dIII fragment generates two new fragments with an additional 1.3 kb of Tn5 DNA. At the same time, there is only one *Bam*HI site which divides the transposon into nearly equal halves. Taking this information into account, the Tn5 insertions were mapped as shown in Fig. 1.

The Tn5 insertions in the hook gene cluster identify genes that code for flagellum formation, since no flagellar structures were detected by electron microscopy of whole cells (data not shown). Although the insertions are scattered throughout approximately 17 kb of DNA, their distribution is not entirely random. One cluster of four Tn5 elements, including *fla-172*::Tn5, occurs within one 1.0-kb region, whereas no Tn5 insertions were found in the 2.3-kb *Bam*HI fragment that codes for the hook protein gene (Fig. 1).

Genomic DNAs from two Tn5 insertion mutants of the *hunA* and *hunB* (histidine utilization; D. M. Ferber and B. Ely, manuscript in preparation) genes were also screened, since both genes are linked genetically to the hook gene cluster (see Fig. 4). However, no insertions were detected within the 50-kb stretch of DNA.

Genomic screen of spontaneous *fla* mutants. We have also examined the possibility that some of the previously mapped *fla* mutants (12) arise by deletion, insertion, or rearrangement within the hook gene cluster that can be detected by Southern analysis of genomic DNA. In the experiments described here DNA from several of these spontaneous *fla* gene mutants (Table 1) was restricted with *Eco*RI, *Hin*dIII, or *Bam*HI. The fragments were separated, blotted as described in the previous section, and then probed with nick-translated Cc8855 DNA and Cc9295 DNA, whose CB15 DNA insert extends farther to the right than that in Cc8861 (Fig. 1). DNA rearrangements were detected in the two mutants previously designated *flaK*, SC298 (*fla-144*), and SC511 (*flaK155*). *fla-144* is an 0.8-kb insertion in the 0.67-kb *Bam*HI fragment, and *flaK155* is a 1.2-kb insertion in the 4.7-kb *Eco*RI fragment (Fig. 2B). Additional restriction analysis of *flaK155* DNA showed a reduction in size of the 2.3-kb *Bam*HI fragment; consequently, we conclude that the *flaK155* element is inserted in the hook protein gene and that it contains at least one *Bam*HI site. This result identifies the hook gene cluster with the *flaK* linkage group of genes previously mapped on the *C. crescentus* chromosome (12).

Effects of *fla*::Tn5 and spontaneous mutations on flagellar gene expression. To examine flagellar gene expression in mutants of the hook gene cluster, a set of isogenic strains was constructed by transducing strain CB15 with bacteriophage φCr30 (11) grown on the *fla*::Tn5 mutants listed in Table 1. The kanamycin resistant marker of Tn5 was used

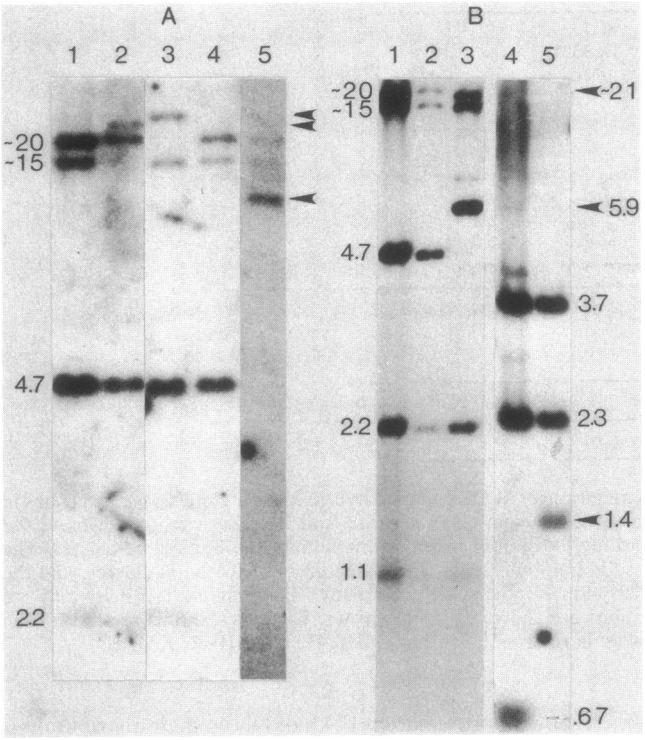

FIG. 2. Identification of insertion mutations in the hook gene cluster of *C. crescentus*. (A) *Eco*RI fragments of genomic DNA from strain CB15 and Tn*5*-induced *fla* mutants were separated on 0.8% agarose gels, and Southern blots were prepared (31) and probed with a mixture of nick-translated λCc8855 and λCc8861 DNA (Fig. 1). Lane 1, CB15; lane 2, SC1132 (*fla-608*::Tn*5*); lane 3, SC1135 (*fla-602*::Tn*5*); lane 4, SC1062 (*flaE178*::Tn*5*, unlinked to hook gene cluster, [23]); lane 5, SC1054 (*fla-173*::Tn*5*). (B) Genomic DNA from strain CB15 and spontaneous *fla* mutants was digested with either *Eco*RI (lanes 1 to 3) or *Bam*HI (lanes 4 and 5) and analyzed as described above by probing with nick-translated DNA from λCc8855 and λCc9295. Lanes 1 and 4, strain CB15; lanes 2 and 5, SC298 (*fla-144*); lane 3, SC511 (*flaK155*). DNA fragment sizes are in kb (Fig. 1); arrows indicate positions of new DNA fragments present in *fla* mutants. Minor bands not indicated by size are products of incomplete digestion.

for selection of recombinants (Table 2), all of which had the expected Fla$^-$ phenotype. Rates of hook protein and flagellin synthesis in representative *fla*::Tn*5* recombinants were determined by labeling exponential cultures with [^{35}S]methionine, dividing the sample into equal parts, and assaying with antiflagellin or antihook protein serum (27). Flagellin antiserum immunoprecipitates flagellins A and B and a 29-kd flagellin, which is coded by a separate gene with partial homology to the other flagellin genes (7), but whose function is unknown. The hook protein antiserum precipitates the 70-kd protein which is assembled into the hook structure and a 68-kd protein which may be a precursor of the 70-kd species (13). The 68-kd protein is synthesized at about one-half the rate of the 70-kd protein (Table 2).

The effect of the insertions on flagellar gene expression depends on their location within the hook cluster (Table 2). Tn*5* insertions downstream (3′ to the hook protein gene; Fig. 1; 21) have a phenotype not previously described for *fla* mutants of *C. crescentus*. These strains do not produce detectable levels of the 70-kd hook protein, and generally we were unable to measure significant levels of the 68-kd hook protein. However, the presence of contaminating bands near the 68-kd position makes it difficult to rule out the possibility that some of the mutants make very low levels of this protein. The 3′ Tn*5* insertion mutants also make extremely low or undetectable levels of flagellins A and B, which are coded for by genes unlinked to the hook gene cluster (see below). They do, however, make elevated levels of the 29-kd flagellin, a protein made at very low levels in strain CB15 and in many *fla*$^-$ strains (13).

The spontaneous insertions that are within the hook protein gene (*flaK155*) and immediately adjacent to the 5′ end of the gene (*fla-144*) abolish hook protein synthesis and generally show the same pattern of flagellin synthesis just described for the 3′ Tn*5* insertion mutants.

Other Tn*5* insertions that are 5′ to the hook protein gene are more typical of previously examined *fla* mutations in their effects. These mutants generally make reduced, but significant levels of hook proteins, flagellin A, and flagellin B. They also make detectable levels of the 29-kd flagellin. Strains PC5503 (*fla-612*) and PC5506 (*fla-602*) are unusual in producing only the 68-kd protein; the incomplete polarity of these two insertions, which are presumably in the same transcriptional unit as the hook protein gene, is considered below. Strain PC5502 (*fla-176*) is also unusual in overproducing the 68-kd protein.

Complementation patterns of insertion mutations in the hook gene cluster. Ditta et al. (4) have described a broad host range cloning vector pRK290 which can be mobilized by plasmid pRK2013 and transferred from *E. coli* to *Rhizobium* sp. in triparental crosses. We have found that pRK290 can be transferred to *C. crescentus* cells in analogous crosses and that the plasmid, which carries a tetracycline resistance gene, replicates stably in *C. crescentus* when the drug is

TABLE 2. Rates of hook protein and flagellin synthesis in Tn*5* insertion mutants of the hook gene cluster

Strain	Genotype	Synthesis rate[a]				
		Hook proteins		Flagellin		
		70 kd	68 kd	A (25 kd)	B (27 kd)	(29 kd)
CB15	Wild type	1	0.5	1	1	0
PCM103	*flaJ303*	9.5	0.2	0.07	0.1	0.1
PC5501	*fla-166*	1.2	0.5	0.19	0.14	<0.02
PC5502	*fla-176*	1	10.	0.25	0.19	0.36
PC5505	*fla-603*	0.89	1.2	0.12	0.05	<0.02
PC5506	*fla-602*	0	0.2	0.06	0.1	0.16
PC5503	*fla-612*	0	0.5	0.1	0.1	0.2
SC298	*fla-144*	0	0	0.05	<0.02	0.33
SC511	*flaK155*	0	0	0.05	<0.02	0.26
PC5511	*fla-172*	0	0	0.03	0	0.24
PC5514	*fla-173*	0	0	0	0	0.11
PC5504	*fla-607*	0	0	0	0	0.15
PC5509	*fla-188*	0	<0.1	0.02	<0.02	0.33
PC5510	*fla-198*	0	0.1	0.03	<0.02	0.29
PC5512	*fla-609*	0	0	0	0	0.44
PC5515	*fla-177*	0	<0.1	0.03	0	0.18
PC5516	*fla-608*	0	<0.1	0	0	0.1

[a] For the 70-kd hook protein, flagellin A, and flagellin B, the rates of synthesis in strain CB15 were taken as 1.0, and rates of synthesis in the mutants were normalized to this value. The rates of the 68-kd protein synthesis were normalized to the wild-type level of 70 kd protein synthesis, and the rates of 29-kd synthesis were normalized to the wild-type level of flagellin B synthesis. Tn*5* insertion mutants beginning with *fla-166* are arranged according to their map positions in Fig. 1.

present (data not shown). *Eco*RI and *Bam*HI fragments of various lengths from hybrid clones of the hook gene cluster were subcloned into the unique *Eco*RI or *Bgl*II sites of pRK290 (4) and tested for complementation of the *fla*::Tn*5* mutants described above. Since the insertion of Tn*5* in both orientations is usually polar (1, 2), complementation between insertion mutants and cloned fragments of different lengths can be used to define transcriptional units or operons rather than genes: for complementation to occur, a fragment must carry the entire operon, including the promoter region. A similar strategy has been used in complementation studies of the *nif* region of *Rhizobium* sp., except that the cloned fragment also carried selected Tn*5* insertions (25).

The hybrid plasmids of pRK290 carrying the fragments shown in Fig. 3 were transferred to a set of *fla*::Tn*5* insertion mutants of the hook gene cluster by selection of prototropic transconjugates on tetracycline plates. To test for complementation, the merodiploid strains were purified and then stabbed into motility agar (12) that contained tetracycline. The extent of swarming after 48 h was compared with the wild-type *fla*$^+$ strain, the *fla*::Tn*5* parent, and the mutant strain carrying the the pRK290 vector.

Subclone S339 complements the hook protein gene mutation *flaK155* and several mutations 5′ to the hook protein gene (*fla-144*, *fla-612*::Tn*5*, *fla-315*::Tn*5*, and *fla-602*::Tn*5*) to wild-type levels of motility, but it does not complement *fla-311*::Tn*5* (Fig. 2). This pattern of complementation and that exhibited by subclones with overlapping *Bam*HI fragments S312, S304, S350, and S1397 (see below) define the extent of transcription unit II carried by S339. It is bounded on the right by *Bam*HI site d, which is at the end of the hook protein gene, and on the left by a site upstream from *fla-602*::Tn*5*. Since the direction of hook protein gene transcription is known (21), the promoter can be placed between *fla-602*::Tn*5* and *fla-311*::Tn*5* (Fig. 3). We also infer from these results that *fla-311*::Tn*5*, which is not complemented by S339, and the adjacent insertions not covered by S339 belong to at least one other transcriptional unit (I; Fig. 3).

In contrast to the strong complementation of *fla-612*::Tn*5* and *fla-602*::Tn*5* by S339, we observed very small swarms in merodiploids containing the truncated fragment S312 (Fig. 3). This may be explained by incomplete polarity of the insertions (Table 2) due to weak promoters downstream from the sites of insertion, weak outward promoters at the ends of the Tn*5* insertions themselves (2, 3), or a small number of *fla*$^+$ recombinants (see below). However, the failure of S1397 to complement *flaK155* argues against the presence of a strong promoter in the 0.67-kb *Bam*HI fragment.

A similar analysis of complementation by the subcloned *Eco*RI and *Bam*HI fragments downstream from the 2.3-kb hook protein gene was also carried out. The ability of S1505, as well as S385, to complement insertions 3′ to the hook protein gene supports the conclusion that the hook protein gene is part of another transcription unit (II; Fig. 3). Comparison of the complementation patterns of overlapping subclones S1505 and S1835 places the downstream insertions in at least three transcriptional units (Fig. 3). Subclone S1505 apparently contains units III and IV, and subclone S1835 contains units IV and V. The boundary between units III and IV has not been precisely determined, and the analysis of the additonal subclones will be required to determine whether each of the three units contains only a single transcriptional unit.

Complementation of spontaneous and UV light-induced mutations. In the genomic screen of spontaneous *fla* mutants

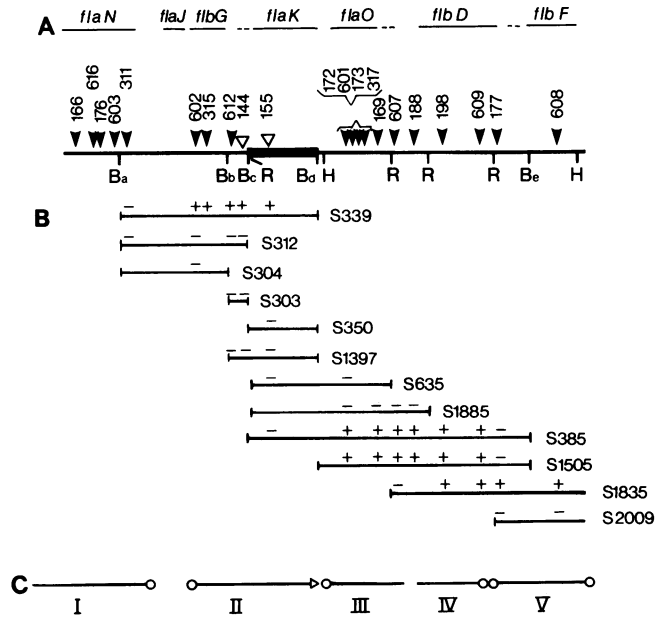

FIG. 3. Complementation analysis of insertion mutants by cloned sequences. (A) The positions of Tn*5* and spontaneous insertion mutants examined for complementation and definitions are given in the legend to Fig. 1. (B) Restriction fragments subcloned in pRK290 are shown. Each subclone was introduced into the *fla* mutants by triparental crosses (4) and selected by tetracycline resistance. Transconjugants were tested for motility (shown above lines representing subclones; +, good motility; −, no motility). The pattern of complementation observed by motility has been confirmed by radioimmunoassay of flagellin production and electron microscopy of flagellum assembly (N. Ohta and A. Newton, unpublished data). (C) Proposed transcriptional units. This represents the minimum number of transcriptional units which might be present in the region indicated (see the text): the direction of transcription is indicated only for unit II. The boundary between transcription units III and IV has not been specified because of ambiguous results in complementation between *fla-188* and S1835 (see the text).

defective in hook production (13), we were able to physically map only the *fla-144* and *flaK155* insertions to the hook protein gene cluster. This set of mutations plus two polyhook mutations were also examined for complementation by the subclones from this region (Table 3). When complementation of the mutations occurred, the swarms in the test plates were similar to those observed for wild-type controls. However, small swarms were observed in a number of cases. When these swarms were analyzed further, it was discovered that plasmid-free segregants retained their motility, indicating that the observed motility was due to recombination rather than complementation. Using both the complementation and the recombination data from Table 3, we were able to determine that mutations previously designated *flaL* and *flaM* are located in the 2.3-kb *Bam*HI fragment that contains the hook protein gene (*flaK*). Since this protein is large enough to occupy nearly the entire fragment, and since both the *flaL* and the *flaM* mutants produce barely detectable amounts of the hook protein (13), we conclude that these mutations are both in the hook protein gene. Consequently, we have redesignated the two alleles *flaK* (Table 3).

Table 3 also indicates that *flaJ303* (27) and *flaJ125* (13), which result in the overproduction of hook protein and the

TABLE 3. Complementation of UV light-induced and spontaneous *fla* mutants

Recipient strain	Mutant strain	Origin	Complementation with the following donor plasmids[a]:					
			S303	S304	S312	S339	S385	S635
PCM103	*flaJ303*	UV	−	+	+	+	−	−
SC269	*flaJ125*	Spontaneous	−	+	+	+	−	−
SC298	*fla-144*	Spontaneous		NT	−[b]	+	−	−[b]
SC511	*flaK155*	Spontaneous	−	−	−	+	−[b]	−
SC300	*flaK146*	Spontaneous	−	NT	−	+	R	R
SC266	*flaK124*	Spontaneous	−	NT	−	+	R	R
SC272	*flaN128*	Spontaneous	−	R	R	R	−	−
SC290	*flaO138*	Spontaneous	−	NT	−	−	+	R
SC307	*flaQ150*	Spontaneous	−	NT	−	−	−	−
SC259	*flaW120*	Spontaneous	−	NT	−	−	−	−
SC297	*flaW143*	Spontaneous	−	NT	−	−	−	−

[a] The donor plasmid was introduced into the recipient strain by selection for tetracycline resistance. After purification, single colonies were stabbed into swarming medium (12) containing tetracycline. Bacteria from positive swarms were purified, grown in the absence of tetracycline, and screened for the loss of plasmid. Plasmid-free isolates were stabbed into motility agar to determine whether concurrent loss of motility had occurred. +, True complementation with concurrent loss of motility and tetracycline resistance; R, recombination, with motility retained independent of drug resistance; −, no swarming was observed; NT, not tested.

[b] Insertion covered by plasmid insert but recombination not observed.

formation of polyhooks, are in the 3.7-kb *Bam*HI fragment. Although these mutations are complemented by S304 and S339, it is not yet clear whether they are nonpolar mutations in the hook gene transcriptional unit or in another, adjacent transcriptional unit (unit 1; see below).

We also map the *flaN* gene to this same fragment in the vicinity of the Tn*5* insertion in PC5001 (Fig. 3); *flaN128* recombines with S339, but it is not complemented by the fragment. These mutations and the cluster of adjacent Tn*5* insertions are in transcriptional unit I, and they are all assigned provisionally in the *flaN* gene. However, given the different phenotypes of the mutations (compare hook gene expression of *fla-166*, *fla-176*, and *fla-603* in Table 2), it is likely that the cluster contains several genes.

We have been able to place the *flaO* gene on the 3′ side of the hook protein gene on the basis of its complementation with S385 and recombination with S635. Therefore, the *flaO* gene must be located in the 7-kb *Bam*HI fragment within the transcriptional unit III. The other two spontaneous *fla* genes, *flaQ* and *flaW*, may not be present on any of the subclones tested, since no recombination or complementation was observed.

Location of the hook gene cluster on the *C. crescentus* genetic map. Physical mapping of the *flaK155* insertion to the 2.3-kb *Bam*HI fragment (Fig. 1) indicates that the hook protein gene is the *flaK* gene. Previous experiments have shown that the *flaK* gene is part of a multigene cluster that is located in the vicinity of *cysC* on the *C. crescentus* genetic map (Ely et al., submitted for publication). The complementation experiments described above allowed us to determine the physical location of a number of spontaneous *fla* mutations and to determine the gene order for the region, *flaN-flaJ-flbG-flaK-flaO-flbD-flbF* (Fig. 4). Genes *flaQ* and *flaW* are located in this region, since they exhibit linkage value similar to those of nearby *hunB* mutation (Ely et al., submitted for publication), but their relative positions could not be demonstrated, since no recombination or complementation was observed with any of the cloned DNA fragments.

Since *fla*::Tn*5* insertions were shown to be interspersed with the spontaneous *fla* mutations, we used the kanamycin resistance of the Tn*5* insertions to help orient the gene cluster with respect to nearby markers. Phage grown on SC1132 (*flaD608*), SC1135 (*flbG602*), and SC1038 (*flaN166*) were used to transduce SC1721 (*cysB102 aar-127 str-196*) to kanamycin resistance. When the resulting colonies were screened for prototrophy and amino acid resistance, it was determined that only the *flaN* gene was linked to the *cysB* (1%) and that *flbD*, *flbG*, and *flaN* alleles were linked to *aar* with linkage values of 27, 60, and 59%, respectively. These results indicate that *flaN* is closest to *aar* and *cysB*. Furthermore, since *cysC* had been shown previously to be unlinked to *aar* (16), the hook gene cluster must be between *cysC* and *aar*. The gene order in this region and the position of the hook gene cluster relative to the *flaE* gene cluster, which contains genes for the 25- and 29-kd flagellins (7, 23) are shown in Fig. 4.

DISCUSSION

As part of a study of flagellar gene expression in *C. crescentus*, we have identified and physically mapped a set of Tn*5* insertion mutations that are linked to the hook protein gene (21; Ohta et al., Fed. Proc.). The insertions define a major cluster of *fla* genes, and complementation analysis using DNA fragments from the region subcloned in broad host range plasmid pRK290 has shown that these *fla* genes are organized into at least five transcriptional units. The results also show that a number of the genes downstream

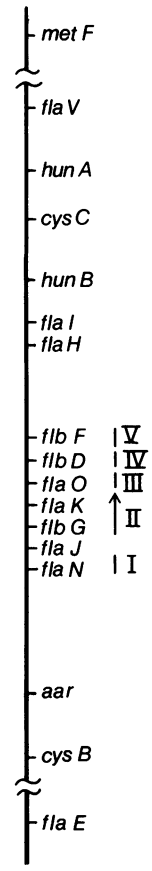

FIG. 4. Genetic map of the hook gene cluster and its position on the *C. crescentus* chromosome relative to the *flaE* gene cluster. Proposed locations of transcriptional units I to IV (Fig. 3) are indicated in the hook gene cluster.

from the hook protein gene are required for expression of both the hook protein gene and the flagellin A and B genes, which are not linked to the hook gene cluster on the *C. crescentus* chromosome (Fig. 3). The result suggests that these genes play a central role in regulation of the flagellar gene family.

The complementation and recombination results (Table 3) indicate that four previously identified genes *flaK*, *flaJ*, *flaN*, and *flaO* map in this cluster. Furthermore, genes designated *flaL* and *flaM* are actually in the *flaK* gene, which is identified here as the hook protein gene. We have placed the *flaK*, *flbG*, and *flaJ* genes in transcriptional unit II, although it is possible that the *flaJ303* and *flaJ125* mutations could map to the region designated *flbG*. This analysis also identifies at least three *fla* genes (*flbF*, *flbD*, and *flbG*) not observed in earlier mutational analyses. Our assignment of genes in this gene cluster is conservative, however, and except for unit II, only one gene has been designated per transcriptional unit. Since the *flaO* and *flbD* genes span almost 6 kb of DNA that contains numerous Tn*5* insertions, it seems likely that additional *fla* genes are located in the hook gene cluster.

The mutations *flaK155* and *fla-144* are the first spontaneous insertions of *C. crescentus* identified. In fact, a large fraction of spontaneous *fla* mutations previously mapped in the *flaE* region have been characterized as deletions (23). Mapping the *flaK155* insertion in the 2.3-kb fragment identifies *flaK* as the hook protein gene; this conclusion has been confirmed by comparison of the N-terminal amino acid sequence of the hook protein gene with the DNA sequence of the 2.3-kb fragment (Vaaler et al., in preparation). *fla-144* maps to the .67 *Bam*HI fragment, and since this fragment is not part of the hook protein coding sequence (Vaaler et al., in preparation), the mutation may be in another gene. This association places the hook gene cluster within the *flaJ*, *K*,*N*,*O*,*Q*,*W* linkage group mapped previously on the *C. crescentus* chromosome (Ely et al., submitted).

Three subclones, S339, S1505, and S1835, together complement most of the *fla*::Tn*5* mutations in the hook gene cluster, and they help to define the five transcriptional units discussed in this paper (Fig. 3). S339 apparently contains a single transcription unit of approximately 6 kb, but low levels of complementation were also observed with overlapping subclone S312, suggesting the presence of additional, weaker promoters to the right of *fla-144* or, alternatively, transcription from outward promoters on Tn*5* elements of *fla-612*::Tn*5* and *fla-602*::Tn*5* (2, 3). Either explanation is consistent with the low levels of the 68-kd protein synthesized by SC1049 and SC1135 (Table 2).

Transcriptional units III and IV each contain approximately 3 kb of DNA, and they are carried by S1505 (Fig. 3). The directions of transcription have not been determined, and the possibility that they contain more than one transcriptional unit has not been eliminated. Somewhat weaker complementation was observed for several of the insertions by these two transcriptional units, e.g., *fla-609* and *fla-198*; this may result from the relative instability of plasmids in these strains (data not shown) or from the overproduction of gene products by the multicopy plasmid, which in turn has a deleterious effect on flagellum assembly or motility in merodiploid strains, as described for the *flaA* gene in *E. coli* (D. O. Clegg and E. Koshland, Fed. Proc., **42**:2137, 1983).

The complementation analysis described here also permits several significant conclusions about the regulation of flagellar gene expression. First, cloned hook protein gene carried on an autonomously replicating plasmid is expressed and allows normal flagellum assembly and function in complementing strains like SC511(pRK290:S339) (Fig. 3). Recent results also show that hook protein synthesis in this merodiplid occurs with normal periodicity in the cell cycle (Ohta et al., manuscript in preparation). Thus, although induction of the hook protein gene appears to require a defined stage of chromosome replication (28), localization of this particular transcription unit to the chromosome is not a necessary feature in the regulation of the hook protein gene. Second, transcriptional unit II contains at least the hook protein gene *flaK* and *flbG*, and possibly *flaJ* as well; this is the first description of an operon in *C. crescentus*. Third, downstream mutations with pleiotropic effects on flagellin and hook proteins fall into at least three transcription units, indicating that more than one gene or operon is somehow involved in expression of the hook protein and flagellin genes. Some of these genes must code for positive regulatory proteins, since Tn*5* insertions in these genes are complemented in *trans* by S1505 and S1835. This conclusion also follows from the proposed upstream location of promoter for transcription unit II and unlinked chromosomal locations of the flagellin genes. Four flagellin genes have been cloned in *C. crescentus* (19; P. R. Gill and N. Agabian, Fed. Proc. **42**:1969, 1983). The 29-kd flagellin gene and one copy of flagellin A gene are located near *flaE* (7; Fig. 4), and a second copy of the flagellin A gene and presumably the flagellin B gene are located at another, genetically unmapped location (Gill and Agabian, Fed. Proc.).

Previous studies of polyhook mutant PCM103 (*flaJ*303) have shown that flagellin genes are not expressed when the assembly of hook and hook protein synthesis are not terminated normally in the cell cycle (28). Radioimmunoassay of strain SC511 (*flaK155*) suggests that hook protein gene expression is also required for flagellin synthesis (Table 2). Consistent with this conclusion is the failure of all Tn*5* insertion mutants downstream from the hook protein gene to synthesize either hook protein or flagellins A and B. This latter set of mutants is unusual in their effect on flagellar gene regulation, since most other *fla* mutants produce reduced but significant levels of flagellin (13).

Thus, the regulation of *fla* gene expression in *C. crescentus* seems at least as complex as that described for other bacteria. The requirement of some *fla* gene products for expression of other *fla* gene indicates that a hierarchical regulation as described for *E. coli* and *Salmonella typhimurium* (10, 16, 29, 30) may operate in *C. crescentus* as well (see discussion in reference 28). In addition, the requirement of DNA synthesis for expression of flagellar genes in these bacteria (20a, 28) must interact at some level of this regulatory circuit. The analysis of individual transcription units like those described in this study provides one approach to the study of regulation of flagellar gene expression.

ACKNOWLEDGMENTS

We thank Melissa Bucuk for constructing the transductants of the Tn*5* insertion mutants listed in Table 2.

This work was supported by Public Health Service grants GM 22299 and GM 25644 (A.N.) and 08-R1GM-33580A01 (B.E.) from the National Institutes of Health and National Science Foundation grant PCM 8003729 (B.E.). Whitehall Foundation also provided support.

LITERATURE CITED

1. **Berg, D. E.** 1977. Insertion and excision of the transposable kanamycin-resistance determinant Tn5, p. 205–212. *In* A. I. Bukhari, J. A. Shapiro, and S. L. Adhya (ed.), DNA insertion

elements, plasmids, and episomes. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
2. **Berg, D. E., A. Weiss, and L. Crossland.** 1980. Polarity of Tn5 insertion mutations in *Escherichia coli*. J. Bacteriol. **142**:439–446.
3. **Ciampi, M. S., M. B. Schmid, and J. R. Roth.** 1982. Transposon Tn10 provides a promoter for transcription of adjacent sequences. Proc. Natl. Acad. Sci. U.S.A. **79**:5016–5020.
4. **Ditta, G., S. Stanfield, D. Corbin, and D. R. Helinski.** 1980. Broad host range cloning system for gram-negative bacteria: construtions of a gene bank for *Rhizobium meliloti*. Proc. Natl. Acad. Sci. U.S.A. **77**:7347–7351.
5. **Ely, B., and R. H. Croft.** 1982. Transposon mutagenesis in *Caulobacter crescentus*. J. Bacteriol. **149**:620–625.
6. **Ferber, D. M., and B. Ely.** 1982. Resistance to amino acid inhibition in *Caulobacter crescentus*. Mol. Gen. Genet. **187**:446–452.
7. **Gill, P. R., and N. Agabian.** 1983. The nucleotide sequence of Mr = 28,500 flagellin genes of *Caulobacter crescentus*. J. Biol. Chem. **258**:7395–7401.
8. **Grunstein, M., and D. S. Hogness.** 1975. Colony hybridization: a method for the isolation of cloned DNAs that contain a specific gene. Proc. Natl. Acad. Sci. U.S.A. **72**:3961–3965.
9. **Huguenel, E., and A. Newton.** 1982. Localization of surface structures during procaryotic differentiation: role of cell division in *Caulobacter crescentus*. Differentiation **21**:71–78.
10. **Iino, T.** 1977. Genetics of structure and function of bacterial flagella. Annu. Rev. Genet. **11**:161–182.
11. **Johnson, R. C., and B. Ely.** 1977. Isolation of spontaneously derived mutants of *Caulobacter crescentus*. Genetics **86**:25–32.
12. **Johnson, R. C., and B. Ely.** 1979. Analysis of nonmotile mutants of the dimorphic bacterium *Caulobacter crescentus*. J. Bacteriol. **137**:627–634.
13. **Johnson, R. C., D. M. Ferber, and B. Ely.** 1983. Synthesis and assembly of flagellar components by *Caulobacter crescentus* motility mutants. J. Bacteriol. **154**:1137–1144.
14. **Jorgensen, R. A., S. J. Rothstein, and W. S. Reznikoff.** 1979. A restriction enzyme cleavage map of Tn5 and location of a region encoding neomycin resistance. Mol. Gen. Genet. **177**:65–72.
15. **Katz, L., D. J. Kingsbury, and D. R. Helinski.** 1973. Stimulation by cyclic adenosine monophosphate of plasmid deoxyribonucleic acid replication and catabolite repression by the plasmid deoxyribonucleic acid protein relaxation complex. J. Bacteriol. **114**:577–591.
16. **Komeda, Y.** 1982. Fusions of flagellar operons to lactose genes on a *mu lac* bacteriophage. J. Bacteriol. **150**:16–26.
17. **Maniatis, J., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual, p. 90. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
18. **Milhausen, M., and N. Agabian.** 1983. *Caulobacter* flagellin mRNA segregates asymmetrically at division. Nature (London) **302**:630–632.
19. **Milhausen, M., P. R. Gill, G. Parker, and N. Agabian.** 1982. Cloning of developmentally regulated flagellin genes from *Caulobacter crescentus* via immunoprecipitation of polyribosomes. Proc. Natl. Acad. Sci. U.S.A. **79**:6847–6851.
20. **Newton, A.** 1972. Role of transcription in the temporal control of development in *Caulobacter crescentus*. Proc. Natl. Acad. Sci. U.S.A. **69**:447–451.

20a. **Newton, A.** 1984. Temporal and spatial control of the *Caulobacter* cell cycle, p. 51–75. *In* P. Nurse and E. Streiblova (ed.), The microbial cell cycle. CRC press.

21. **Ohta, N., L.-S. Chen, and A. Newton.** 1982. Isolation and expression of cloned hook protein gene from *Caulobacter crescentus*. Proc. Natl. Acad. Sci. U.S.A. **79**:4863–4867.
22. **Poindexter, J. S.** 1964. Biological properties and classification of the *Caulobacter* group. Bacteriol Rev. **28**:231–295.
23. **Purucker, M., R. Bryan, K. Amemiya, B. Ely, and L. Shapiro.** 1982. Isolation of a *Caulobacter* gene cluster specifying flagellum production by using nonmotile Tn5 insertion mutants. Proc. Natl. Acad. Sci. U.S.A. **79**:6797–6801.
24. **Rigby, P. W., M. Dieckmann, C. Rhodes, and P. Berg.** 1977. Labeling deoxyribonucleic acid to high specific activity *in vitro* by nick translation with DNA polymerase I. J. Mol. Biol. **113**:236–251.
25. **Ruvkun, G. B., V. Sundaresan, and F. M. Ausubel.** 1982. Direct transposon Tn5 mutagenesis and complementation analysis of *Rhizobium meliloti* symbiotic nitrogen fixation genes. Cell **29**:551–559.
26. **Shapiro, L.** 1976. Differentiation in the *Caulobacter* cell cycle. Annu. Rev. Microbiol. **30**:377–407.
27. **Sheffery, M., and A. Newton.** 1979. Purification and characterization of a polyhook protein from *Caulobacter crescentus*. J. Bacteriol. **138**:575–583.
28. **Sheffery, M., and A. Newton.** 1981. Regulation of periodic protein synthesis in the cell cycle: control of initiation and termination of flagellar gene expression. Cell **24**:49–57.
29. **Silverman, M.** 1980. Building bacterial flagella. Quart. Rev. Biol. **55**:395–408.
30. **Silverman, M., and M. I. Simon.** 1977. Bacterial flagella. Annu. Rev. Microbiol. **31**:397–419.
31. **Southern, E. M.** 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. J. Mol. Biol. **98**:503–517.