Proc. Natl. Acad. Sci. USA
Vol. 84, pp. 1142–1146, March 1987
Biochemistry

Promoter mapping and cell cycle regulation of flagellin gene transcription in *Caulobacter crescentus*

(periodic gene expression/promoter sequences/nuclease S1 mapping)

SCOTT A. MINNICH AND AUSTIN NEWTON

Department of Molecular Biology, Princeton University, Princeton, NJ 08544

Communicated by Arthur B. Pardee, October 16, 1986 (received for review September 2, 1986)

ABSTRACT **_Caulobacter crescentus_ contains a 25- and a 27-kDa flagellin, which are assembled into the flagellar filament, and a 29-kDa flagellin, which is related in sequence but is of unknown function. We have used DNA sequence analysis and nuclease S1 assays to map the *in vivo* transcription start sites of the three flagellin genes and to study their regulation. These experiments lead to several conclusions. First, copies of the 29-, 25-, and 27-kDa flagellin genes are organized in a tandem array in the *flaEY* gene cluster of *C. crescentus*. Second, flagellin genes are under transcriptional control and each gene is expressed with a characteristic periodicity in the cell cycle. Third, flagellin gene promoters contain conserved nucleotide sequence elements at −13, −24, and −100 that are homologous to the *fla* genes in the hook gene cluster. The −13 and −24 sequences conform to a *fla* gene promoter consensus sequence (C/TTGGCC/GC-N_5-TTGC) that is similar in sequence to the −12, −24 consensus sequence of the *Klebsiella pneumonia nif* gene promoters. Fourth, the sequence element at approximately −100 in the 25- and the 27-kDa flagellin genes is homologous to a 19-base-pair sequence [designated previously as II-1; see Chen, L.-S., Mullin, D. M. & Newton, A. (1986) *Proc. Natl. Acad. Sci. USA* 83, 2860–2864] at −101 in the promoter of transcription unit II of the hook gene cluster; the two flagellin genes, like the *fla* genes examined in the hook gene cluster that contain the −100 element, are under positive control by transcription unit III of the hook gene cluster. This result supports a model in which the timing of *fla* gene transcription in the *C. crescentus* cell cycle is determined in part by a cascade of trans-acting regulatory gene products.**

The well-defined nature of asymmetric cell division in the aquatic bacterium *Caulobacter crescentus* has provided an attractive model for the study of mechanisms that regulate the temporal and spatial patterns of development. Unlike *Bacillus* and *Myxococcus*, in which differentiation is largely controlled by environmental conditions, morphogenesis in *C. crescentus* is driven by an internal clock that appears, from genetic results, to be linked to steps in DNA synthetic or cell division pathways (1–3). Flagellum biosynthesis has been of special interest in the study of cell-cycle timing because formation of this structure is stage-specific in the cell cycle, and the major flagellar proteins are synthesized periodically at the time of flagellum assembly (1, 2, 4). These proteins include the 70-kDa hook protein, the 27-kDa flagellin, which is assembled proximal to the hook, and the 25-kDa flagellin, which is assembled at the distal end of the filament. These two flagellins and a 29-kDa flagellin of undetermined function have extensive sequence homology (5–7).

Most *C. crescentus* flagellar (*fla*) genes are organized into three genetically unlinked clusters. The best characterized of these are the hook gene cluster (8, 9), which contains transcription units I–V, and the *flaEY* cluster (10, 11), which contains the 29- (10), the 25- (12), and as shown in this communication, the 27-kDa flagellin genes. Transcription units I, II, and III of the hook gene cluster (refs. 13 and 14; D. Mullin, L.-S. Chen, and A.N., unpublished work) and members of the flagellin gene family are transcriptionally regulated (15, 16). We also know that *de novo* RNA synthesis is required for the periodic synthesis of the hook protein and the flagellins (2), and it seems likely from these results that *fla* genes are differentially transcribed in the *C. crescentus* cell cycle and that promoter recognition plays a role in timing their expression (13, 14). To date, however, only the promoters of transcription units II and II.1 of the hook gene cluster have been identified and sequenced (14).

To extend these studies to flagellin gene expression in *C. crescentus*, we have used nuclease S1 assays to map the *in vivo* transcription start sites of the 29- (10), the 25-, and the 27-kDa flagellin genes and to determine their organization within the *flaEY* gene cluster. We show here that the three flagellin genes are transcriptionally regulated and that each gene is expressed with a characteristic periodicity in the cell cycle. The flagellin gene promoters contain conserved nucleotide sequences at −13, −24, and approximately −100, and we speculate that these sequence elements may be important for promoter recognition and cell-cycle regulation of *fla* gene expression in *C. crescentus*.

MATERIALS AND METHODS

Strain and Growth Conditions. Wild-type *C. crescentus* strain CB15 (American Type Culture Collection 19089) and motility mutants have been described (9). Deletion mutant PC7801 (S.A.M., N. Taylor, N. Ohta, and A.N., unpublished work) was constructed by homologous recombination between the *C. crescentus* chromosome and a nonreplicating hybrid plasmid in which the *Eco*RI(a)–*Eco*RI(b) fragment of the *flaEY* region (see Fig. 1) was replaced by the ω cassette (17), which confers streptomycin resistance to the recombinant strain. *C. crescentus* cultures were grown at 30°C in peptone yeast extract or in M2 minimal salts medium supplemented with 0.2% glucose as described previously (13). *Escherichia coli* strains were grown in LB medium at 37°C.

Isolation of Flagellin Genes. Genes in the *flaEY* cluster of *C. crescentus* were originally cloned on a 40-kilobase (kb) DNA fragment in cosmid vector pHC79 (18) by selecting for the kanamycin resistance of the *flaE*::Tn*5* mutation in strain SC1062 (S.A.M., N. Taylor, N. Ohta, and A.N., unpublished work). The same selection was used by Purucker *et al.* (11) to clone *flaEY* genes in plasmid pBR325. The pBR322 subclone used in these studies, pCN200, contains a 6.6-kb *Hin*dIII–*Hin*dIII DNA fragment (see Fig. 1) from the cosmid clone.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviation: nt, nucleotide(s).

Nuclease S1 Mapping and DNA Sequencing. DNA restriction fragments from hybrid plasmid pCN200 were eluted from agarose gels, 5′ end-labeled (19), and used as probes in nuclease S1 assays (13, 20). Cellular RNA was purified from *C. crescentus* cells as described (13). DNA fragments were subcloned from plasmid pCN200 into M13 vectors mp18 and mp19 for nucleotide sequencing of the sense and antisense strands by the method of Sanger *et al.* (21).

RESULTS

Identification of the Transcription Start Site of the 29-kDa Flagellin Gene. *C. crescentus* flagellin genes from the *flaEY* cluster were subcloned on a 6.6-kb *Hin*dIII–*Hin*dIII fragment in plasmid pBR322 (S.A.M., N. Taylor, N. Ohta, and A.N., unpublished work). The restriction map adjacent to the Tn*5* insertion element (Fig. 1) overlaps that published previously for the 29-kDa flagellin gene (10, 11). We confirmed that this portion of the cloned fragment contains the 29-kDa flagellin gene by determining the nucleotide sequence of the 300 base pairs (bp) to the right of the *Pst* I(d) site (Fig. 1) and comparing it to the published sequence (7).

The transcription start site of the 29-kDa flagellin gene was mapped using the 960-bp *Pst* I(c)–*Pst* I(d) fragment A (Fig. 1) as the 5′ end-labeled probe in nuclease S1 assays. RNA purified from wild-type *C. crescentus* strain CB15 protected a 104-nucleotide (nt) DNA fragment (Fig. 2*A*, lane 3). The direction of transcription was determined using probes B and C (Fig. 1) labeled only at the *Pst* I ends. The 104-nt fragment was protected when probe B was used but not when probe C was used (data not shown). Thus, the 29-kDa flagellin gene is transcribed from left to right with the *in vivo* transcription initiation site 104 bp to the left of the *Pst* I(d) site and approximately 60 bp upstream from the ATG start codon (Fig. 1; see ref. 7).

Identification of the 25-kDa Flagellin Gene and Transcription Start Site. Cloned fragments of *C. crescentus* DNA have been reported to contain the 25- and the 29-kDa flagellin genes (12). Our finding of significant sequence homology between the 950-bp *Pst* I(d)–*Sal* I(c) fragment, which includes most of the 29-kDa flagellin structural gene, and the 960-bp *Pst* I(c)–*Pst* I(d) fragment initially suggested that the 25-kDa flagellin gene lies upstream from promoter *P29* of the 29-kDa flagellin gene (Fig. 1). We have mapped an *in vivo* transcription start site in this region approximately 1200 bp 5′ to promoter *P29*. Partially protected fragments of 275 and 270 nt were obtained when the 1000-bp DNA fragment E was the labeled probe (Fig. 2*B*, lane 3), and fragments of 135 and 130 nt were obtained when the overlapping 1100-bp DNA fragment was the labeled probe (data not shown). Thus, the 5′ ends of the messenger RNAs map 130 and 135 bp to the left of the *Eco*RI(b) site, and transcription is from left to right (Fig. 1).

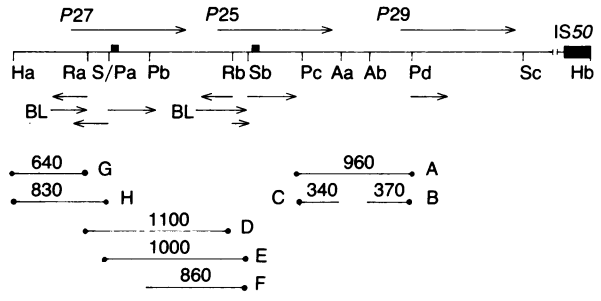

FIG. 1. Physical and transcription maps of flagellin genes in the *flaEY* cluster of *C. crescentus*. The 6.6-kb *Hin*dIII–*Hin*dIII fragment extends from within the *flaEY* cluster on the left to the *Hin*dIII(b) site on the right, which is within insertion sequence *50* of Tn*5*. The transcription start sites and the direction of transcription are indicated by the arrows above the restriction map. DNA fragments *Eco*RI(a)–*Eco*RI(b), *Sal* I(b)–*Sal* I(c), and *Pst* I(d)–*Sal* I(c) were used as probes in Southern blots (22) of genomic DNA digested with *Bam*HI, *Hin*dIII, *Pst* I, or *Eco*RI. Only the second probe, which contained most of the 25-kDa flagellin gene, hybridized strongly to chromosome sequences outside of the cloned *flaEY* gene segment. Boxes indicate a 62-bp direct repeat. Nucleotide sequence determinations were carried out on the DNA restriction fragments and BAL-31 digested fragments (BL) indicated by arrows below the restriction map. Sizes of DNA fragments are in bp with the 5′-labeled ends represented by closed circles. Restriction sites: H, *Hin*dIII; S, *Sal* I; A, *Ava* I; P, *Pst* I; R, *Eco*RI. *Pst* I(a) maps 18 bp to the right of *Sal* I(a).

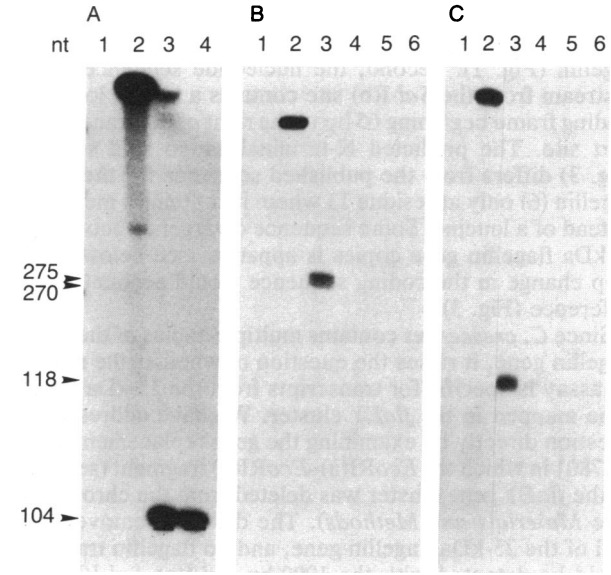

FIG. 2. Nuclease S1 mapping of promoters *P29*, *P25*, and *P27*. The 5′ end-labeled probes shown in Fig. 1 were incubated with nuclease S1 at 55°C. Control experiments are represented by lanes 1 and 2; other lanes contained nuclease S1, probe DNA, and *C. crescentus* RNA (100 μg) purified from exponentially growing cells of the strains indicated. (*A*) *P29* [960-bp *Pst* I(c)–*Pst* I(d) probe A]. Lanes: 1, probe plus nuclease S1; 2, probe alone; 3, strain CB15; 4, strain SC1032 (*flbD*198::Tn*5*, a mutation in transcription unit IV; ref. 9). (*B*) *P25* [1000-bp *Sal* I(a)–*Sal* I(b) probe E]. Lanes: 1, probe plus nuclease S1; 2, probe alone; 3, strain CB15; 4, strain PC7081 [*Eco*RI(a)–*Eco*RI(b) deletion mutant]; 5, strain SC1032; 6, strain SC1052 (*flaO*172::TN*5*, a mutation in transcription unit III; ref. 9). (*C*) *P27* [640-bp *Hin*dIII(a)–*Eco*RI(a) probe G]. Lanes: 1, probe plus nuclease S1; 2, probe alone; 3, strain CB15; 4, strain PC7081; 5, strain SC1032; 6, strain SC1052. The sizes of the S1-resistant DNA fragments were determined by electrophoresis on a 4% denaturing polyacrylamide gel containing 8 M urea and by comparison of their mobilities with standard DNA fragments from plasmid pBR322 digested with *Hin*fI, *Hae* II, or *Hae* III.

The location of the start site was confirmed by showing that DNA probe F that had been 5′ end-labeled at the *Sal* I site produced partially protected fragments of 275 and 270 nt (data not shown). The two messenger RNAs detected in these assays could result from either tandem start sites or messenger RNA processing since control experiments have shown that the heterogeneity is not due to incomplete nuclease S1 digestion.

Two observations support the conclusion that this transcription start site(s) locates promoter *P25* of the 25-kDa flagellin gene. First, hybridization experiments (see legend to Fig. 1) have shown that the coding sequence downstream from the start site is present in at least three genomic copies, which is consistent with the report that *C. crescentus* contains multiple copies of the 25-kDa flagellin gene (12); in similar experiments we have detected only one genomic sequence with strong homology to DNA sequences contain-

ing either the 29-kDa flagellin gene or the third flagellin gene identified in these studies, which may code for the 27-kDa flagellin (Fig. 1). Second, the nucleotide sequence 400 bp upstream from the *Sal* I(b) site contains a single, long open reading frame beginning 65 bp to the right of the transcription start site. The predicted N-terminal amino acid sequence (Fig. 3) differs from the published sequence for the 25-kDa flagellin (6) only at residue 13 where it contains a methionine instead of a leucine. Some sequence divergence between the 25-kDa flagellin gene copies is apparent (see below), and a 1-bp change in the coding sequence would account for this difference (Fig. 3).

Since *C. crescentus* contains multiple copies of the 25-kDa flagellin gene, it raises the question of whether the nuclease S1 assay is specific for transcripts from the 25-kDa flagellin gene mapped in the *flaEY* cluster. We have addressed this question directly by examining the gene replacement mutant PC7801 in which the *Eco*RI(a)–*Eco*RI(b) fragment (see Fig. 1) of the *flaEY* gene cluster was deleted from the chromosome (see *Materials and Methods*). The deletion removes the 5′ end of the 25-kDa flagellin gene, and no flagellin transcripts could be detected with the 1000-bp *Sal* I(a)–*Sal* I(b) probe (Fig. 2*B*, lane 4). The observation that the deletion mutant still synthesizes 25-kDa flagellin (unpublished results) attests to the specificity of the nuclease S1 assay used in these studies, and it supports the idea that some sequence divergence has occurred between copies of the 25-kDa flagellin gene.

Identification of the 27-kDa Flagellin Gene and the Transcription Start Site. The 25-kDa flagellin gene was found to contain a 62-bp nucleotide sequence 22 bp to the right of the *Sal* I(b) site that is repeated upstream 18 bp to the right of *Sal* I(a) site (boxed in Fig. 1). This redundant sequence indicated the possible presence of a third flagellin gene in the *flaEY* cluster. Nuclease S1 mapping using end-labeled DNA fragments G and H as probes (Fig. 1) produced partially protected fragments of 118 (Fig. 2*C*, lane 3) and 305 (data not shown) nt, respectively; this result places the transcription start site 118 bp to the left of the *Eco*RI(a) site and indicates that the three flagellin genes are transcribed in the same direction (Fig. 1).

The nucleotide sequence to the left of the *Pst* I(a) site contains an open reading frame in the direction of transcription beginning 56 bp downstream from the transcription start site. The sequence codes for a protein with extensive homology to the N-terminal amino acid sequence of the 25- and the 29-kDa flagellin (Fig. 3). The N-terminal amino acid sequence of the 27-kDa flagellin has not been published, but the sequence homologies shown in Fig. 3 and the regulation of this transcription unit discussed below suggest that this third *in vivo* transcription start site locates promoter *P*27 of the 27-kDa flagellin gene. Also supporting this gene assignment is the failure of the *Eco*RI(a)–*Eco*RI(b) deletion mutant PC7801 (see above and Fig. 1), which contains only the 5′-regulatory region and the first 16 codons of the gene, to synthesize the 27-kDa flagellin (S.A.M., N. Taylor, N. Ohta, and A.N., unpublished work). Nuclease S1 assays using the *Hin*dIII(a)–*Eco*RI(a) fragment G as the probe show that levels of transcription from the truncated flagellin gene are also greatly reduced in the deletion mutant (Fig. 2*C*, lane 4).

Regulation of Flagellin Genes by *fla* Genes in the Hook Gene Cluster. Genes in transcription units III, IV, and V of the hook gene cluster are required to be in trans for synthesis of the 25- and the 27-kDa flagellins (9), and the failure to detect transcripts from promoters *P*25 and *P*27 in mutants of transcription units III and IV (Fig. 2 *B* and *C*, lanes 5 and 6) indicates that the two flagellin genes are under positive, transcriptional regulation. By contrast, the level of transcription from the 29-kDa flagellin gene is not affected in *fla* mutants of transcription unit III (Fig. 2*A*, lane 4), although the rate of 29-kDa flagellin synthesis has been shown to increase by 5- to 10-fold in the same strains (9). These results suggest that the 29-kDa flagellin gene is subject to translational control, but the gene is apparently also under transcriptional control since the transcript is accumulated periodically in the cell cycle (see below).

Flagellin Genes Are Differentially Regulated in the *C. crescentus* Cell Cycle. The nuclease S1 assays described here are specific for individual flagellin gene transcripts (Fig. 2), and we have used them to measure flagellin gene expression in synchronous cell cultures (Fig. 4). Our findings are consistent with the results of RNA-hybridization blots (15, 16) that used the 29-kDa flagellin gene as a probe to show that a homologous transcript(s) first appears late in the S period of the cell cycle. More specifically, however, we have been able to distinguish between the three flagellin transcripts and to resolve their order of expression. Transcripts were detected in predivisional cells first from the 29-kDa flagellin gene, then from the 27-kDa flagellin gene, and finally from the 25-kDa flagellin gene (Fig. 4). Only the 25-kDa flagellin gene transcript was present in newly divided swarmer cells, presumably because of the greater stability of this messenger RNA (2).

The patterns of messenger RNA accumulation from promoters *P*29, *P*27, and *P*25 are significantly different from one another, yet they mirror almost exactly the synthetic periods of the corresponding flagellins determined in several laboratories by radioimmunoassay (1, 2, 4). This agreement between the messenger RNA and protein synthetic patterns is consistent with the assignment of the transcriptional start sites shown in Fig. 1.

```
29 kDa flagellin
   H2N-MET ALA LEU     SER VAL ASN THR ASN GLN PRO ALA LEU ILE ALA LEU GLN

25 kDa flagellin
       ATG GCG CTG AAC AGC ATC AAT ACG AAC GCG GGC GCG ATG ATC GCC CTG CAA
   H2N-MET ALA LEU ASN SER ILE ASN THR ASN ALA GLY ALA MET ILE ALA LEU GLN

27 kDa flagellin
       ATG     CTG AAC TCG ATC AAC ACC AAT CCT GGT GCG CTC GTG GCG CTC CAG
   H2N-MET     LEU ASN SER ILE ASN THR ASN PRO GLY ALA LEU VAL ALA LEU GLN
```

FIG. 3. N-terminal amino acid and coding sequences of flagellin genes. Nucleotide sequences of the 25- and the 27-kDa flagellin genes were analyzed for homology by comparison with the 29-kDa flagellin sequence in the National Biomedical Research Foundation's data bank,* and open reading frames of the two sequences were determined using the program of Delaney (23) modified by D. Welsch. The N-terminal amino acid sequence of the 29-kDa flagellin gene was taken from ref. 7. Residues conserved in three sequences are in bold-faced type and those conserved in two are underlined.

*Protein Identification Resource (1985) Protein Sequence Database (Natl. Biomed. Res. Found., Washington, DC), Release 6.

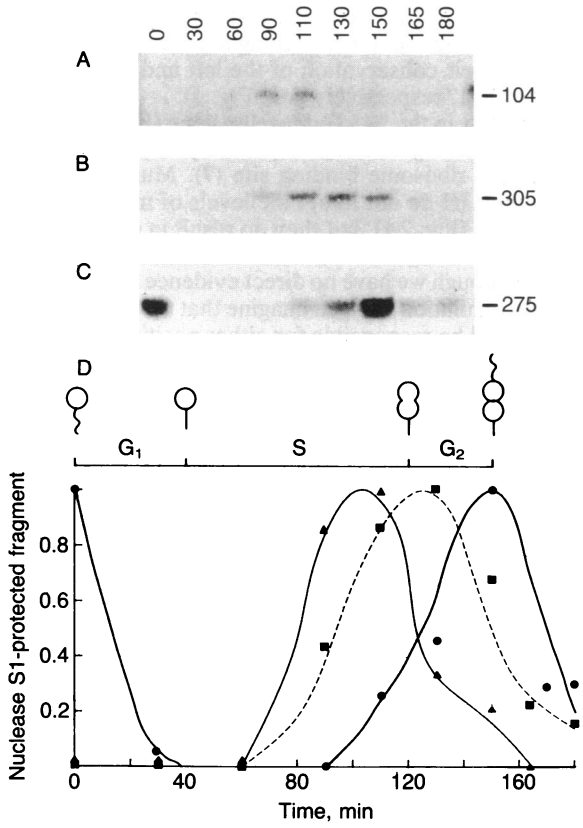

FIG. 4. Periodic expression of flagellin gene transcripts in the *C. crescentus* cell cycle. Cells of *C. crescentus* strain CB15F were synchronized as described (14). Total cellular RNA was isolated at the times indicated, and the same RNA preparation was used in parallel nuclease S1 assays with the three DNA probes: (*A*) *P29*, probe A. (*B*) *P27*, probe H. (*C*) *P25*, probe E. Autoradiograms were quantified as described (13), and the values were plotted after normalizing the peak amounts to 1 (*D*). ▲, *P29*; ■, *P27*; ●, *P25*. Cell division occurred at 150 min. The predominant cell type at each time is represented above the cell-cycle scale.

Nucleotide Sequences of Flagellin Gene Promoters. The nucleotide sequences of promoters *P25* and *P27* were determined on both strands (see Fig. 1) and compared with that of promoter *P29* (Fig. 5). The three promoters share regions of nucleotide sequence homology at −13 and −24 with *fla* gene promoters in the hook gene cluster (ref. 14; D. Mullin, S.A.M., L.-S. Chen, and A.N., unpublished work); no homology at −10 and −35 to the canonical *E. coli* promoter sequence (24) has been found.

A third region of nucleotide sequence homology is located in regions of dyad symmetry at −100 in promoter *P25* and at −115 in promoter *P27* (Fig. 5). This 19-bp sequence, II-1, is also present in *P27* at −4 and in the 29-kDa flagellin gene downstream from the transcription start site at +25. The recurrence of element II-1 in the *C. crescentus fla* gene promoters suggests that the sequence may be a cis-acting element important for the regulation of these genes in the cell cycle (see *Discussion*).

DISCUSSION

The results reported in this communication map the *in vivo* transcription start sites of the 29-, the 25-, and the 27-kDa flagellin genes and define the organization of the three genes in the *flaEY* gene cluster of *C. crescentus* (Fig. 1). The transcription start site of the 29-kDa flagellin gene was assigned using the results of the nuclease S1 assays and published sequence data (7). The two adjacent start sites cannot be definitively assigned until N-terminal amino acid-sequence data are available for all of the flagellins, but our results suggest that they correspond to the 25- and the 27-kDa flagellin genes, as shown in Fig. 1. This conclusion is supported by the predicted N-terminal amino acid sequences (Fig. 3), the characteristic patterns of cell-cycle regulation for the three flagellin genes (Fig. 4), and the presence of multiple genomic copies of the 25-kDa flagellin gene (Fig. 1). Also supporting these assignments is the observation that the deletion mutant PC 7801, which lacks the *Eco*RI(a)–*Eco*RI(b) segment of the *flaEY* cluster (Fig. 1), synthesizes 25-kDa flagellin, but not 27-kDa flagellin (S.A.M., N. Taylor, N. Ohta, and A.N., unpublished work).

Our principal interest in identifying the transcription start sites of the flagellin genes is to study the role of the 5′-regulatory regions in *fla* gene regulation. The *C. crescentus fla* genes examined to date, including the flagellin genes (Fig. 4) and those in the hook gene cluster (13, 14) are periodically expressed in the cell cycle. These genes are under transcriptional regulation, and in addition, there is some indication (see below) that the 29-kDa flagellin gene and perhaps other *fla* genes may also be under translational control.

How is this complex temporal pattern of gene regulation coordinated in the cell cycle with flagellar morphogenesis? Genetic studies in *C. crescentus* (2, 13, 14, 16) and *E. coli* (25) have suggested that *fla* genes are arranged in a regulatory hierarchy wherein expression of each *fla* gene requires the prior expression of a *fla* gene(s) located above it in the hierarchy. Thus, we have proposed that the timing of *fla* gene(s) expression required for formation of the hook and

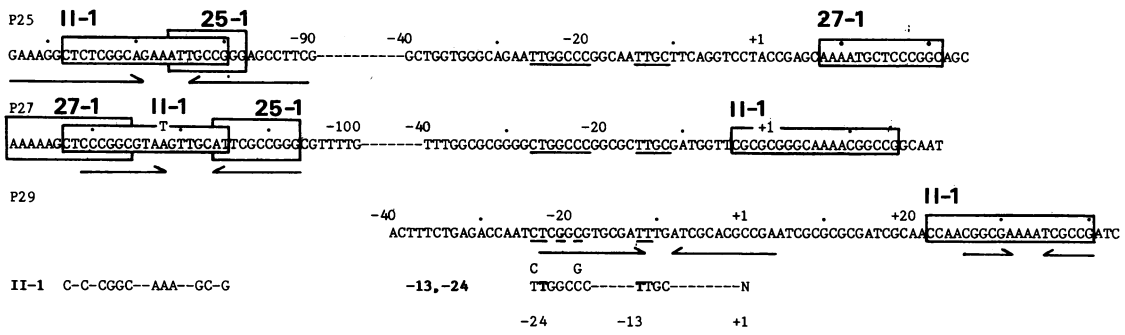

FIG. 5. Conserved sequence elements in flagellin gene promoters. The bases corresponding to the 5′-terminal nucleotides of nuclease S1-protected fragments are numbered +1 on the noncoding strand. In promoter *P25* the start site located 135 bp to the left of the *Eco*RI(b) site (Fig. 1) is indicated; the second 5′-messenger end (see text) maps at +5. The conserved −13 and −24 elements are underlined, and other regions of sequence homology are boxed. The sequences for II-1 and the −13 and −24 elements at the bottom of the figure are a consensus of six *fla* gene promoter sequences (D. Mullin, S.A.M., L.-S. Chen, and A.N., unpublished work). The nucleotide sequence of 29-kDa flagellin gene was taken from ref. 7. Sequences were analyzed for inverted repeats and regions of homology as described in Fig. 3.

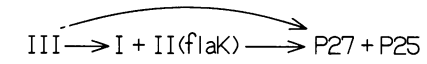

FIG. 6. Cascade regulation of *fla* genes in the hook operon and the *flaEY* cluster of *C. crescentus* (13, 14). Transcription units in the hook gene cluster are organized in the order I-II-II.1-III-IV-V, with *flaK* (hook protein structural gene) as the last gene in transcription unit II (13). Arrows indicate the requirement of one *fla* gene or transcription unit in trans for the expression of the next *fla* gene in the cascade.

flagellar filament is controlled by a regulatory cascade, as shown in Fig. 6 (13, 14). According to this model, the time at which a *fla* gene is expressed is determined in part by the position of the gene within the regulatory cascade, and the time of expression correlates in general with the location of the gene product in the morphogenic pathway of organelle assembly. Several results support this model: (*i*) the hook protein and flagellins are synthesized in the cell cycle at the time of their assembly into the flagellum (1, 4); (*ii*) transcription unit III, which is required to be in *trans* for expression of transcription units I and II, is expressed before transcription units I and II in the cell cycle (D. Mullin, L.-S. Chen, and A.N., unpublished work); (*iii*) expression of the hook protein structural gene (*flaK*) is required in trans for synthesis of the 25- and the 27-kDa flagellins; and (*iv*) transcripts from the 27- and the 25-kDa flagellin genes appear sequentially in the order of flagellin synthesis and assembly into the flagellar filament (Fig. 4).

In considering the mechanisms underlying a regulatory cascade of *fla* genes, we have assumed that differential promoter recognition plays an important role. The conserved sequence elements at −13 and −24 (Fig. 5) conform to a proposed *C. crescentus fla* gene consensus sequence (C/TTGGCC/GC-N_5-TTGC-N_{9-12}; A.N., D. Mullin, S.A.M., and L.-S. Chen, unpublished results) for the promoter sequences studied here (Fig. 5) and those in transcription units I, II, and III of the hook gene cluster (ref. 14; D. Mullin, S.A.M., L.-S. Chen, and A.N., unpublished work). The sequence, which is not homologous to the −10, −35 *E. coli* promoter sequence (24), is best conserved in *fla* genes under positive regulation by transcription unit III—e.g., *P25* and *P27* (see Fig. 5)—and it may not be representative of all *C. crescentus* promoters. This −13, −24 promoter sequence is, however, strikingly similar to the −12, −24 consensus sequence of *nif* and *ntr* gene promoters in *Klebsiella pneumonia* (26, 27), *E. coli* (28), and *Salmonella* (28). We do not yet know the significance of the −13, −24 promoter motif in *C. crescentus*, but the −12, −24 promoter of *E. coli* (28) and *Salmonella* (29) *glnA* genes require a specific σ factor for recognition. It would be interesting if a similar mechanism were involved in the differential expression of *C. crescentus fla* genes.

The third conserved sequence element II-1 at −100 in promoter *P25* and −115 in promoter *P27* (Fig. 5) was initially identified as one arm of a dyad upstream from transcription unit II of the hook gene cluster (14). We now know that this 19-bp sequence is located at −100, midway between divergent transcription units I and II, both of which require transcription unit III in trans (D. Mullin, S.A.M., L.-S. Chen, and A.N., unpublished work). Thus, four transcription units, which are under positive regulation by transcription unit III, contain element II-1 at approximately −100, and it seems reasonable to speculate that the transcriptional regulation of these genes is modulated by the interaction of a regulatory protein(s) with regions of dyad symmetry containing II-1. The copy of element II-1 at −4 in *P27*, and the partially homologous element 27-1 at +8 in the 25-kDa flagellin gene could be related in function. An intriguing feature of elements 27-1 and 25-1 is their conservation of the left and right segments of element II-1, respectively (see Fig. 5).

Element II-1 in the 29-kDa flagellin gene (Fig. 5) is at +23, and it contains a region of dyad symmetry 4 bp upstream from the proposed ribosome binding site (7). Mutations in transcription unit III do not affect the levels of messenger RNA from the gene (Fig. 2*A*), but they do result in overproduction of the 29-kDa flagellin (9), which indicates translational control. Although we have no direct evidence for a role of II-1 in *fla* gene regulation, we can imagine that the same sequence element could be responsible for either positive—e.g., in the 25-kDa flagellin gene—or negative regulation—e.g., in the 29-kDa flagellin gene—depending on its location relative to the transcription or translation start site.

fla genes are under complex temporal control in *C. crescentus*, and a molecular model for their regulation must account for the sequential staging of gene activation and product assembly in the cell cycle. It seems likely that the conserved 5′-sequence elements described in this communication will be important in formulating such a model.

We thank David Mullin for his comments and suggestions during the course of this work and Noriko Ohta for her suggestions and help in the preparation of this manuscript. This research was supported by Public Health Service Grants GM22299 and GM25644. Nucleotide sequence analysis was performed in part by resources provided by Bionet under National Institutes of Health Grant 1 U41 RR-168502.

1. Osley, M. A., Sheffery, M. & Newton, A. (1977) *Cell* **12**, 393–400.
2. Sheffery, M. & Newton, A. (1981) *Cell* **24**, 49–57.
3. Newton, A., Ohta, N, Huguenel, E. & Chen, L.-S. (1985) in *Microbial Differentiation*, eds. Setlow, P. & Hoch, J. (Am. Soc. Microbiol., Washington, DC), pp. 267–276.
4. Lagenaur, C. & Agabian, N. (1978) *J. Bacteriol.* **135**, 1062–1069.
5. Weissborn, A., Steinman, H. M. & Shapiro, L. (1982) *J. Biol. Chem.* **257**, 2066–2074.
6. Gill, P. R. & Agabian, N. (1982) *J. Bacteriol.* **150**, 925.
7. Gill, P. R. & Agabian, N. (1983) *J. Biol. Chem.* **258**, 7395–7401.
8. Ohta, N., Chen, L.-S. & Newton, A. (1982) *Proc. Natl. Acad. Sci. USA* **70**, 4863–4867.
9. Ohta, N., Swanson, E., Ely, B. & Newton, A. (1984) *J. Bacteriol.* **158**, 897–904.
10. Milhausen, H., Gill, P. R., Parker, G. & Agabian, N. (1982) *Proc. Natl. Acad. Sci. USA* **79**, 6847–6851.
11. Purucker, M., Bryan, R., Amemiya, K., Ely, B. & Shapiro, L. (1982) *Proc. Natl. Acad. Sci. USA* **79**, 6797–6801.
12. Gill, P. R. & Agabian, N. (1983) *Fed. Proc. Fed. Am. Soc. Exp. Biol.* **47**, 1969 (abstr.).
13. Ohta, N., Chen, L.-S., Swanson, E. & Newton, A. (1985) *J. Mol. Biol.* **186**, 107–115.
14. Chen, L.-S., Mullin, D. & Newton, A. (1986) *Proc. Natl. Acad. Sci. USA* **83**, 2860–2864.
15. Milhausen, M. & Agabian, N. (1983) *Nature (London)* **302**, 630–632.
16. Champer, R., Bryan, R., Gomes, S. L., Purucker, M. & Shapiro, L. (1985) *Cold Spring Harbor Symp. Quant. Biol.* **50**, 831–840.
17. Prentki, P. & Kirsch, H. M. (1984) *Gene* **29**, 303–313.
18. Hohn, B. & Collins, J. (1980) *Gene* **11**, 291–298.
19. Maxam, A. M. & Gilbert, W. (1980) *Methods Enzymol.* **65**, 499–560.
20. Berk, J. A. & Sharp, R. A. (1977) *Cell* **12**, 721–732.
21. Sanger, F., Nicklen, S. & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 5463–5467.
22. Southern, E. M. (1975) *J. Mol. Biol.* **98**, 503–517.
23. Delaney, A. D. (1982) *Nucleic Acids Res.* **10**, 61–67.
24. Hawley, D. K. & McClure, W. R. (1983) *Nucleic Acids Res.* **11**, 2237–2255.
25. Komeda, Y. J. (1982) *J. Bacteriol.* **150**, 16–26.
26. Beynon, J., Cannon, M., Buchanan-Wollanston, V. & Cannon, F. (1983) *Cell* **34**, 665–671.
27. Ausubel, F. M. (1984) *Cell* **37**, 5–6.
28. Hunt, T. P. & Magasanik, B. (1985) *Proc. Natl. Acad. Sci. USA* **82**, 8453–8457.
29. Hirschman, J., Wong, P.-K., Sei, J., Keener, J. & Kustu, S. (1985) *Proc. Natl. Acad. Sci. USA* **82**, 7525–7529.