JOURNAL OF BACTERIOLOGY, Oct. 1990, p. 6066–6076
0021-9193/90/106066-11$02.00/0

Vol. 172, No. 10

# Identification of a *Caulobacter* Basal Body Structural Gene and a *cis*-Acting Site Required for Activation of Transcription

ANDREW DINGWALL, JAMES W. GOBER, AND LUCILLE SHAPIRO*

*Department of Developmental Biology, Beckman Center, Stanford University School of Medicine, Stanford, California 94305-5427*

Received 14 March 1990/Accepted 30 July 1990

**The genes that encode the components and regulatory proteins of the *Caulobacter crescentus* flagellum are transcribed at specific times in the cell cycle. One of these genes, *flbN*, is required early in the flagellar assembly process. The *flbN* gene was cloned and sequenced, and the time of transcription activation was determined. The derived amino acid sequence indicates that *flbN* encodes a 25-kilodalton protein with a cleavable leader peptide. The *flbN*-encoded protein has 30.8% identity with the protein encoded by the *Salmonella typhimurium* basal body L-ring gene, *flgH*. Site-directed mutagenesis and gel mobility shift assays identified a binding site at −100 from the transcription start site for a *trans*-acting protein, RF-2, that functions to partially activate *flbN* transcription at a defined time in the cell cycle. The RF-2 binding region is similar to a NifA binding site normally used in the activation of some $\sigma^{54}$ promoters involved in nitrogen fixation in other bacteria. Transcription of a *flbN*-reporter gene fusion in an *Escherichia coli* background was dependent on the presence of a NifA transcription factor supplied by a plasmid-borne *Rhizobium meliloti* gene encoding NifA. A deletion or base changes in the RF-2 binding region eliminated expression of the *flbN* gene in *E. coli* even when a NifA protein was provided in *trans*, suggesting that a $\sigma^{54}$ promoter with an upstream activator element is used by the *C. crescentus flbN* gene. A consensus sequence for a $\sigma^{54}$ promoter was found at the appropriate distance 5′ to one of two identified transcription start sites. Site-directed mutagenesis confirmed that a conserved nucleotide in this $\sigma^{54}$ promoter consensus sequence was required for transcription. Deletion of the region 5′ to the apparent $\sigma^{54}$ promoter caused a complete loss of transcription activation. Transcription activation of *flbN* in *C. crescentus* involves the combination of several elements: the NifA-like site is required for full activation, and other sequence elements 5′ to the promoter and 3′ to the transcription start site are necessary for the correct time of transcription initiation.**

The genes required for *Caulobacter crescentus* flagellar biogenesis are transcribed in an ordered temporal sequence, and the gene products are localized to the pole of the cell. Over 48 genes have been identified that are required for the formation and function of the polar flagellum and accompanying chemotaxis apparatus (15–17). In *Escherichia coli* (34, 35), *Salmonella typhimurium* (33), and *C. crescentus* (6, 8, 9, 48, 50, 60), the transcription of the flagellar and chemotaxis genes is regulated in a complex *trans*-acting hierarchy. In order to define the *cis*-acting sequence elements that are involved in transcriptional activation, we have identified and analyzed a flagellar gene, *flbN*, that is expressed early in the flagellum biogenesis pathway.

The *flbN* gene was identified on the basis of its nonmotile phenotype in insertion mutant strains SC1117 and AE8002 (9, 26). These mutants failed to form any assembled flagellar structure. Complementation analysis demonstrated that the *flbN* gene was located within a cluster of genes involved in flagellar basal body biogenesis. In this report, we present evidence that the *flbN* gene encodes the *C. crescentus* basal body L-ring protein.

Transcription of the *flbN* gene is subject to both positive and negative *trans*-acting control within the context of the regulatory hierarchy (60). The timing and extent of *flbN* gene transcription could be controlled by alternative sigma factors, as is the case in *Bacillus subtilis* sporulation (39), and/or by a unique combination of *trans*-acting factors that bind to upstream regulatory regions.

* Corresponding author.

At least two distinct promoter sequences for the *C. crescentus* flagellar and chemotaxis genes have been reported. These include those that conform to the *ntr* or $\sigma^{54}$ promoter (42, 46, 47) and a $\sigma^{28}$-like promoter sequence (20) that is used by many *E. coli* and *B. subtilis* flagellar and chemotaxis genes (2). The use of the $\sigma^{54}$ promoter by the *C. crescentus* hook operon was demonstrated by site-directed mutagenesis (47), and this promoter was shown to be recognized by purified *E. coli* $\sigma^{54}$ in vitro (49). The hook operon promoter has been shown to require the presence of an upstream sequence (*ftr*) for full transcription of the operon (47).

In other bacteria, all of the known genes with $\sigma^{54}$ promoters require the presence of an upstream regulatory sequence and its cognate *trans*-acting factor for transcription activation (for a review, see reference 37). We report here the sequence of the *flbN* gene and show that the expression of the gene is under temporal control. Deletion and site-directed mutagenesis have identified a *cis*-acting site required for *flbN* transcription, and gel mobility shift assays have shown that a protein, RF-2, specifically binds to this upstream site. The RF-2 binding site for *flbN* differs from the *ftr* site associated with the hook operon promoter. The RF-2 site is similar to NifA sites and is necessary for full activation of *flbN* in *C. crescentus* or in *E. coli*, where it is dependent on a NifA protein provided in *trans* from a plasmid-borne *Rhizobium meliloti nifA* gene. We found that *flbN* contains a consensus $\sigma^{54}$ promoter sequence at the requisite distance from one of two closely spaced start sites. In vitro mutagenesis that changed a conserved base in the −12 region of this promoter sequence abolished transcription, suggesting that

*flbN* is transcribed from a $\sigma^{54}$ promoter. Additional sequence elements in regions both 5′ and 3′ to the transcription start site are required for temporal activation of the *flbN* gene in *C. crescentus*.

## MATERIALS AND METHODS

**Materials.** Restriction enzymes were obtained from Boehringer-Mannheim or New England BioLabs. Calf intestine alkaline phosphatase, nuclease S1, polynucleotide kinase, DNA polymerase I, and Klenow were obtained from Boehringer-Mannheim. T4 DNA ligase was from Collaborative Research. [$\alpha$-$^{32}$P]dCTP (3,000 Ci/mmol), [$\alpha$-$^{32}$P]dATP (800 Ci/mmol), and [$\gamma$-$^{32}$P]dATP (3,000 Ci/mmol) were from Amersham Corp. Acrylamide, *N*,*N*-methylenebisacrylamide (BIS), and *N*,*N*,*N*′,*N*′-tetramethylethylenediamine (TEMED) were obtained from BioRad. [$^{35}$S]-methionine (1150 mCi/mmol) was obtained from ICN, and Formalin-fixed *Staphylococcus aureus* cells were purchased from Bethesda Research Laboratories. Amplify was obtained from Amersham. Reverse transcriptase was obtained from Life Sciences, Inc.

**Bacterial strains and growth conditions.** *C. crescentus* wild-type synchronizable strain CB15N and mutant strains were grown in PYE broth at 30°C (53) or in modified minimal M2 medium (10). Unless otherwise indicated, all *C. crescentus* strains were grown at 30°C and *E. coli* was grown at 37°C. Two *C. crescentus flbN* mutant strains, SC1117 (26) and AE8002 (9), were generated by a Tn*5* insertion and by insertion of a modified Tn*5* transposon that contained a promoterless *neo* gene, respectively. The acquisition of kanamycin resistance in strain AE8002 occurred as a result of insertion of the reporter gene downstream from the *flbN* promoter (see Fig. 1A). Kanamycin resistance screening was carried out by using PYE plates supplemented with antibiotic at 25 to 200 μg/ml. *R. meliloti* plasmid pRMB3.8H containing the *nifA* gene was kindly provided by F. M. Ausubel.

Synchronized populations were obtained by centrifuging 500 to 1,000 ml of *C. crescentus* grown in modified M2 medium at 5,000 × *g* for 20 min. The pellet was suspended in nonsupplemented M2 medium and combined with 1/3 volume of filtered Ludox (DuPont). The cells were then centrifuged, and swarmer populations were isolated (18). The swarmer cells (>95% pure) were suspended in fresh supplemented M2 medium (at an optical density at 660 nm [$OD_{660}$] of 0.2 to 0.4) and allowed to proceed through the cell cycle. Samples of the culture were removed at various times and either pulse-labeled or used for the preparation of RNA synthesized in vivo.

**Immunoprecipitations.** Cell extracts were prepared and immunoprecipitated with antibody as described previously (24). The synthesis of the flagellins and neomycin phosphotransferase II (NPT II) were measured by immunoprecipitating cell extracts obtained from cultures pulse-labeled with [$^{35}$S]methionine (3 μCi/ml) by using antibody raised against purified flagellins (57) or NPT II (8). The amount of radioactivity was normalized prior to addition of antibody such that equal counts per second of labeled cells were immunoprecipitated. Immunoprecipitated proteins were separated by electrophoresis through 10% polyacrylamide–sodium dodecyl sulfate gels. Labeled proteins were visualized by autoradiography following treatment of the gels with Amplify.

**In vitro mutagenesis.** Nested deletions in the *flbN* promoter region were made by using a modification of the procedure of Dale et al. (11) according to the manufacturer's specifications (Cyclone Rapid Deletion Subcloning Kit; IBI, Inc.). A 2.0-kilobase-pair *Sac*I-to-*Sal*I fragment containing the *flbN* promoter fused to a *neo* (NPT II) reporter gene was subcloned into M13mp19. Deletions were made by using T4 DNA polymerase, and endpoints were sequenced. Promoters carrying deletions were cloned into a broad-host-range plasmid pRK290 (14) and then mated into *C. crescentus*.

Site-directed mutagenesis was employed to alter or remove specific sequences according to the procedure of Kunkel and Roberts (36). Oligonucleotides were purchased from Operon Technologies, Inc. (Alameda, Calif.). A 1,100-base-pair (bp) *Pst*I-to-*Sal*I fragment that contained the entire *flbN* gene and upstream regulatory sequences was first cloned into M13mp18. Mutagenesis was confirmed by sequencing. Fusions were made to either luciferase (*lux*) or neomycin phosphotransferase (*neo*) reporter genes, as described in the text. Reporter gene fusions were cloned into pRK290.

**DNA sequence analysis.** The dideoxynucleotide chain termination method of Sanger et al. (54) was used for DNA sequence analysis. Subclones were prepared for sequence analysis by using M13-based vectors, and the complete sequences of both DNA strands were determined.

**Assay of luciferase activity.** Flagellar gene promoters were fused to a promoterless luciferase operon in the plasmid pUCD615 (28), which was kindly provided by Clarence Kado. This vector contains the genes required for luciferase activity in an operon composed of five genes, *luxC*, *luxD*, *luxA*, *luxB*, and *luxE*. These transcriptional fusions were tested for luciferase activity, as described previously (23), in both *C. crescentus* and *E. coli*. In all cases, cultures were grown at their normal temperatures and then downshifted to 25°C for 15 min. Luciferase activity is irreversibly lost at temperatures greater than 30°C, and incubation of cultures at 25°C for 15 min allows the assay of newly synthesized enzyme. A sample of the culture was placed in a scintillation vial, and *n*-decylaldehyde (2% emulsion in water) was added to a final concentration of 0.002%. Emitted light was measured immediately over a period of 6 s in a Packard Tri-Carb scintillation counter.

**Gel mobility shift assays.** Cell extracts were prepared from a log-phase culture of *C. crescentus*. At an $OD_{660}$ of 0.6 to 0.9, cultures were centrifuged, and the pellets were washed in buffer containing 20 mM Tris hydrochloride (pH 7.5), 0.1 mM EDTA, and 100 mM KCl. The cell pellet was resuspended in the same buffer and sonicated on ice by using six 15-s bursts. The cell lysate was centrifuged at 4°C, and the supernatant was removed from the debris. Extracts were stored at −80°C.

DNA-binding assays were performed by the method of Fried and Crothers (21). Single-stranded oligonucleotides, labeled with T4 polynucleotide kinase and [$\alpha$-$^{32}$P]ATP, were annealed and then used in binding reactions. Labeled probe and cell extracts (0.25 to 2.0 μg of protein) were mixed and incubated for 20 min at 25°C in binding buffer (19), which consists of 20 mM Tris hydrochloride (pH 8.0), 0.1 mM EDTA, 0.1 mM dithiothreitol, 100 mM KCl, and 100 μg of bovine serum albumin per ml. Poly(dI-dC) (0.25 μg) was added to each reaction mixture, followed by the addition of 50% (vol/vol) glycerol, 0.02% (wt/vol) bromophenol blue, and 0.02% (vol/vol) xylene cyanol. The samples were then loaded onto a 5% (wt/vol) acrylamide gel. Electrophoresis was carried out in 1× TBE buffer (40) by using 300 V for 3 h. The gels were run at 4°C. The dried gels were exposed to

autoradiographic film overnight at −80°C with intensifying screens.

**Primer extension and nuclease S1 protection assays.** The *flbN* transcript start site was determined by primer extension analysis (31) by using an oligonucleotide (Operon Technologies, Inc.) complementary to a portion of the *flbN* coding sequence (underlined in Fig. 5B). Products produced by reverse transcriptase were electrophoresed through 8% polyacrylamide sequencing gels and visualized by autoradiography. The same oligonucleotide was used in a dideoxy-sequencing reaction run alongside the primer extension lanes in order to align the extended product(s) with the DNA sequence.

Nuclease S1 protection assays were performed according to the procedure of Berk and Sharp (4). Probes used in the 5′ and 3′ transcript analyses were prepared from plasmid pKH1.7 (26). A 620-bp *Pst*I-to-*Ava*I fragment (see Fig. 1A) and a 1,364-bp *Sal*I-to-*Ava*I fragment (see Fig. 5A) were end labeled with [$\gamma$-$^{32}$P]dATP by using T4 polynucleotide kinase and used in the 5′ S1 analyses. A 264-bp *Not*I-to-*Sal*I fragment (see Fig. 5) was end labeled at the *Not*I site by using Klenow polymerase for the 3′ S1 analysis. RNA was extracted from samples of cells taken at various times in the *C. crescentus* cell cycle according to the procedure of Amemiya et al. (1).

**Nucleotide sequence accession number.** The sequence described in this paper has been assigned GenBank number M37270.

## RESULTS

**Identification and nucleotide sequence of the *flbN* gene.** The *flbN* gene maps to a cluster of genes required for basal body formation (26, 27). A schematic drawing of the basal body gene cluster is shown in Fig. 1A. The cluster contains at least six genes organized into several transcription units. The gene order shown above the partial restriction map is based on genetic complementation (26) and does not reflect actual gene boundaries.

The locations of two nonmotile *flbN* insertion mutations are indicated. The AE8002 strain was generated by insertion mutagenesis by using the transposon Tn*5*-VB32 (3) carried on a suicide vector (9). This transposon has a promoterless *neo* gene that renders a host cell resistant to kanamycin only if a chromosomal promoter has been accessed by the insertion event. The formation of a fusion protein is precluded by the presence of translation stop codons in all three reading frames in front of the *neo* gene. The Tn*5*-VB32 insertion in AE8002 was cloned, and the location and orientation of the insertion in the chromosome were determined by Southern blot analysis and DNA sequencing of the insert junction (data not shown). These analyses indicated that the insertion occurred at bp 721 in the sequence shown in Fig. 2. The orientation of the *neo* reporter gene insertion in *flbN*, shown schematically in Fig. 1A, revealed the direction of transcription of the *flbN* gene. The location of the Tn*5* insertion in strain SC1117 was determined by nuclease S1 analysis by using RNA obtained from this strain (data not shown).

In order to further define the boundaries of the *flbN* gene, subclones were made from the region and used in genetic complementation tests (Fig. 1B). Complementation was scored as the restoration of normal flagellar biogenesis and motility in the two *flbN* insertion mutants. It was previously shown that the subclone pKH2.1R (Fig. 1B) complements both the SC1117 and AE8002 mutant strains (26). A smaller, 1,100-bp subclone of the *flbN* region, from the *Pst*I to the

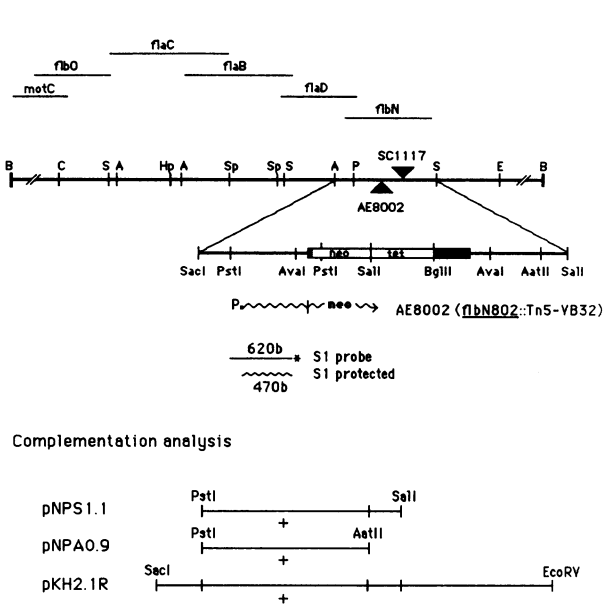


FIG. 1. Organization of the *C. crescentus* basal body cluster. (A) The order of genes within the basal body cluster was determined by genetic complementation (26). A partial restriction map of the region is shown below the genetic map: B, *Bam*HI; C, *Cla*I; S, *Sal*I; A, *Sac*I; Hp, *Hpa*I; Sp, *Sph*I; E, *Eco*RI. Two *flbN* insertion mutations, AE8002 and SC1117, are indicated (▲). The promoterless *neo* reporter gene in the transposon Tn*5*-VB32 (in AE8002 and derivative strains) is shown schematically within the *flbN* gene. Relevant restriction sites are indicated. A 620-base *Pst*I-to-*Ava*I probe used in nuclease S1 analysis and the 470-base protected mRNA are also shown. (B) Complementation analysis using subclones of the basal body region. Complementation was scored as restoration of flagellar biogenesis and motility.

*Sal*I site (pNPS1.1), was also found to complement both mutations, as did the smaller *Pst*I-to-*Aat*II subclone in pNPA0.9 (Fig. 1B). These results indicate that the entire *flbN* coding region is contained within a 935-bp region between the *Pst*I and *Aat*II sites.

The complete nucleotide sequence of an 1,100-bp region between the *Pst*I and *Sal*I sites was determined and is shown in Fig. 2. This region includes an open reading frame that could encode a protein of approximately 25 kilodaltons (kDa). This open reading frame is entirely contained within the 935-bp complementing clone pNPA0.9. The derived amino acid sequence is consistent with the *C. crescentus* preferred codon usage (32), and it shows a biased preference for either a G or C residue in the third position of the codon, a preference that has been observed in other cloned *C. crescentus* genes (35, 56).

The open reading frame identified in Fig. 2 is preceded by a consensus ribosome binding site (GAGG) eight bases from the initiating ATG codon. The predicted open reading frame also contains an amber (TAG) termination signal following codon 242. This termination signal is positioned 20 bases upstream from the *Aat*II restriction site in the complementing clone pNPA0.9 (Fig. 1B).

The derived *flbN* protein sequence contains a hydrophobic region at the amino terminus that conforms to the rules for predicted cleavable leader peptides (51), suggesting that the *flbN* protein may be transported out of the cytoplasm. A plot of the hydropathy profile of the *flbN* peptide according to the algorithm of Kyte and Doolittle (38) predicts that the *flbN*

```
1                                                   31
CTG CAG GGC AAG GCC ATG GCC GCC GCC GCC GCC GGT GAT GTG TTC GCC GTT CAG AAC ACT

61                                                  91
CTG TCG AAG AAG ATC ATC CAG GCT GTG GCC GTC GGC CCC GGC GCC GCC GCT GTC GGC CCG

121                                                 151
CAG GCC CAG AGC TTG CAA GCC CGC TCG CAA CCC CTC CGT TTC GCC GCT CGC TAA GAG GTC

181                                                 211
CAG ATC ATG CGT CGT CCT GCT ATC CTC GCC GCC GCC GTC CTC CTG GCC CCG CTG GCC GCC
       met arg arg pro ala ile leu ala ala ala val leu leu ala pro leu ala ala
241                                                 271
TGC TCG ACC GTC AAG GAG GCC GTC AAG GGT CCG GAT CTC GCG CCC GTC GGC TAC CCG GCC
cys ser thr val lys glu ala val lys gly pro asp leu ala pro val gly tyr pro ala
301    ↑                                            331
CCT CTG GCC CCC ATG CAA CAG CAG TAC GTG TCG GCC CGC GAG CCC GCC CCG CAG CGC TTC
pro leu ala pro met gln gln gln tyr val ser ala arg glu pro ala pro gln arg phe
361                                                 391
GGC CAA CTC GCT GTG GCG GGT CGG GGC CCG GCC TTC TTC AAC GAC CAG CGC GCC AGC CGC
gly gln leu ala val ala gly arg gly pro ala phe phe asn asp gln arg ala ser arg
421                                                 451
GTC GGC GAC ATC GTC ACC GTG ATG ATC GAC ATC GAC GAC AGC GCC AGG ACC AAG AAG CCG
val gly asp ile val thr val met ile asp ile asp asp ser ala arg thr lys lys pro
481                                                 511
ACC AAC AGC TCG CGC ACC GCC AAT ATG AAG GCC GGC GTG CGG CAT CTG CTG GGC ATG GAA
thr asn ser ser arg thr ala asn met lys ala gly val arg his leu leu gly met glu
541                                                 571
TCG AGC CTG GGC AAG TTC CTG CCG GGC GGC TTC GAT CCG GCC TCG GCG CTG GAG ACC AAC
ser ser leu gly lys phe leu pro gly gly phe asp pro ala ser ala leu glu thr asn
601                                                 631
TCG ACC ACC ACC AAC GCC GGC TCG GGC GGC TCA CGG TCG GAA AAG ATC AGC CTG ACG ATC
ser thr thr thr asn ala gly ser gly gly ser arg ser glu lys ile ser leu thr ile
661                                                 691
GCC GCT GTG GTC AGC CAG CTT CTG CCG AAC GGC AAC ATG GTC ATC CAG GGC ACC CAG GAA
ala ala val val ser gln leu leu pro asn gly asn met val ile gln gly thr gln glu
721                                                 751
GTG CGC ACC AAC GCC GAG CTG CGC CAG CTG ACC GTG GCC GGT ATC GTC CGC CCC GAG GAC
val arg thr asn ala glu leu arg gln leu thr val ala gly ile val arg pro glu asp
781                                                 811
ATC TCC TCG GCC AAC ACC ATC CGC CAC ACT CAG ATC GCC GAA GCC CGC ATC AGC TAC GGC
ile ser ser ala asn thr ile arg his thr gln ile ala glu ala arg ile ser tyr gly
841                                                 871
GGC CGC GGC GAC ATC AGC CGC GTG CAG AAG ACC CCG GCC GGC CAG TCG CTG GTC GAG AAG
gly arg gly asp ile ser arg val gln lys thr pro ala gly gln ser leu val glu lys
901                                                 931
TTC TCG CCC TTC TAG GCG CGG CTT TTG CGA CAC AGA CGT CGG GCG CGG CGG CCG AGA GCG
phe ser pro phe AMB
961                                                 991
TTC CGC GCC CGA TGC GTT TCA GGA GCC GTT CAG GTG CGC TCA GCG TTG ATC CGA TCA GGA

1021                                                1051
GGA TCA TGT CGC CCC GCA GTC TCG CCC TGG CCG CCC TTT CAC TAG CGA CCG CCC TCG CCG

1081
GCT GCG CGC ATG AAA GGA TGG CGT CGA
```

FIG. 2. Nucleotide sequence and the derived amino acid sequence of the *flbN* gene. The complete nucleotide sequence of an 1,100-bp *Pst*I-to-*Sal*I region including the *flbN* gene is shown. An open reading frame encoding a protein with a molecular size of 25 kDa was contained within a restriction fragment in pNPA0.9 that complemented *flbN* mutant strains. A predicted signal peptide cleavage site at the end of a hydrophobic leader sequence is indicated by a vertical arrow.

protein is composed of several alternating hydrophobic and hydrophilic regions and a hydrophilic carboxy terminus. The location of the open reading frame, together with the complementation data, predict that the *flbN* promoter and transcription start site are located within the 186-bp region between the *Pst*I site and the translation start codon.

A search of the available data base by using the BESTFIT parameters (12) revealed a strong similarity between the *flbN* protein and the *Salmonella typhimurium* basal body L-ring encoded by *flgH* (30). The comparison indicated a 30.8% identity (Fig. 3) and a 53.7% similarity between the two proteins. In addition, both proteins have predicted molecular sizes of approximately 25 kDa and they have hydrophobic leader sequences. There was no significant similarity between *flbN* and other flagellar genes in the data base.

**Transcription analysis of the *flbN* gene.** Nuclease S1 protection assays of the wild-type chromosomal copy of the *flbN* gene (Fig. 4) were used to determine the 5′ and 3′ ends of the *flbN* mRNA. Two protected fragments of approximately 470 bases were observed by using both a 1,364-base *Sal*I-to-*Ava*I probe (Fig. 4A, lane 1) and a 620-base *Pst*I-to-*Ava*I probe (data not shown). On the basis of migration distance, these protected products appeared to be approximately 10 bases apart. The relative abundance of the two

```
CcflbN    1 MRRPAILA...AAVLLAPLAACS..TVKEAVKGPDLA.PVGYPAPLAPMQ 44
StflgH    1 MQKYALHAYPVMALMVATLTGCAWIPAKPLVQGATTAQPIPGPVPVAN.G 49
CcflbN   45 QQYVSAREPAPQRFGQLAVAGRGPAFFNDQRASRVGDIVTVMIDIDDSAR 94
StflgH   50 SIFQSAQ...PINYGY.......QPLFEDRRPRNIGDTLTIVLQENVSAS 89
CcflbN   95 TKKPTNSSRTANMKAGVRHLLGMESSLGKFLPGGFDPASALETNSTTT.. 142
StflgH   90 KSSSANASRDGKTSFGF.......DTVPRYLQGLFGNSRADMEASGGNSF 132
CcflbN  143 NAGSGGSRSEKISLTIAAVVSQLLPNGNMVIQGTQEVRTNAELRQLTVAG 192
StflgH  133 NGKGGANASNTFSGTLTVTVDQVLANGNLHVVGEKQIAINQGTEFIRFSG 182
CcflbN  193 IVRPEDISSANTIRHTQIAEARISYGGRGDISRVQKTPAGQSLVEKFSPF 242
StflgH  183 VVNPRTISGSNSVPSTQVADARIEYVGNGYINEAQNMGWLQRFFLNLSPM 232
```

FIG. 3. Comparison between the derived *C. crescentus flbN* protein and the *S. typhimurium* basal body L-ring protein (*flgH*) (30). The sequences were aligned by using the BESTFIT analysis (University of Wisconsin Genetics Computer Group, Madison, Wis.; 12). The sequences were aligned starting with the initiation ATG codons; identity between residues (|) and similarities (. or :) are also shown. The protein sequences differ in length by 10 residues, indicated by gaps in the alignment.

transcripts differed, with the smaller transcript predominant. However, the two transcripts observed in wild-type cells were absent in a *flaS* mutant strain (Fig. 4A, lane 2). It was previously demonstrated that the *flaS* gene is required for the expression of *flbN* both on the chromosome (9) and on a plasmid (60).

The 3′ end of the *flbN* transcript was determined by using a 264-base *Not*I-to-*Sal*I fragment, labeled at the *Not*I site with Klenow polymerase. Nuclease S1 digestion produced two major protected fragments of approximately 122 and 128 bases, indicating that the transcript terminates imprecisely about 70 bases downstream from the end of the *flbN* open reading frame (Fig. 4B, lane 3). Similar analyses were performed by using longer upstream probes, and these produced the same 3′ termini identified in Fig. 4B (data not shown). A 3′ transcript analysis using RNA from a *flaS* mutant strain (Fig. 4B, lane 4) confirmed that *flbN* transcription was significantly diminished in the absence of the *flaS* gene product. Therefore, assays of both the 5′ and the 3′ region of the *flbN* transcript revealed that the gene is subject to positive control by the *flaS* gene.

Primer extension analysis was used to more precisely map the transcription start site for the chromosomal copy of the *flbN* gene (Fig. 5). Reverse transcriptase was used to extend the sequence from an annealed oligonucleotide that had been labeled by using polynucleotide kinase. The complement of the oligonucleotide sequence used in this experiment is underlined in Fig. 7B. The same oligonucleotide was used in a dideoxy-sequencing reaction, and electrophoresis was carried out next to the primer extension reaction in order to align the initiating nucleotide with the promoter sequence (Fig. 5). Two extended products, 10 bases apart, were consistently observed in the primer extension reactions. One of these transcription start sites was located 13 bases from a consensus $\sigma^{54}$ promoter sequence (see Fig. 7B). This sequence conforms to the $\sigma^{54}$ recognition site associated with nitrogen-regulated genes and nitrogenase genes from enteric bacteria and rhizobia, and the sequence has also been found in several flagellar genes in *C. crescentus*. The other transcription start site, which gave a stronger signal in the experiment shown in Fig. 5, begins 23 bases from the $\sigma^{54}$ promoter and has no apparent promoter consensus sequence.

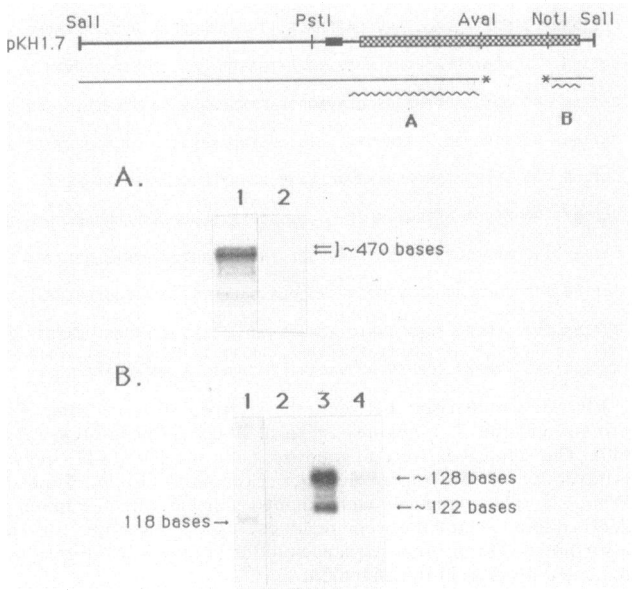


FIG. 4. Nuclease S1 protection assays of the 5′ and 3′ regions of the *flbN* transcript in wild-type and *flaS* mutant strains. Probes for both the 5′ and 3′ transcript analyses were prepared from plasmid pKH1.7 by end labeling either with T4 polynucleotide kinase or with Klenow polymerase. ▩, Putative *flbN* coding region; ■, promoter. A 1,364-base *Sal*I-to-*Ava*I fragment was end labeled as depicted in the schematic with T4 polynucleotide kinase for 5′ transcript analysis. A 264-base *Not*I-to-*Sal*I fragment was labeled with Klenow polymerase for 3′ transcript analysis. Protection assays were performed as described in Materials and Methods with 50 μg of whole-cell RNA from either wild-type strain CB15N or the *flaS* mutant strain SC508. Protected transcripts are shown as wavy lines below their respective probes; letters refer to gels below. (A) 5′ transcription analysis of *flbN* in wild-type CB15N (lane 1) and a *flaS* mutant, SC508 (lane 2). The protected transcript in lane 1 appears as a doublet of approximately 470 bases. (B) 3′ transcription analysis of *flbN*. Lane 1, Molecular weight marker; lane 2, no RNA; lane 3, CB15N RNA; lane 4, SC508 RNA. The two protected products are approximately 128 and 122 bases in length.

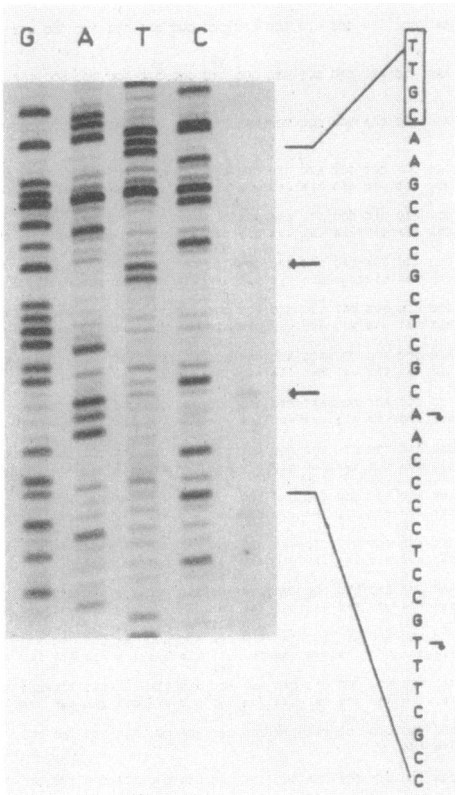


FIG. 5. Identification of the transcript initiation site(s) by primer extension. The initiating nucleotides for the *flbN* transcript were mapped by primer extension analysis using an oligonucleotide complementary to the sequence underlined in Fig. 5B. The 20-base oligonucleotide primer was end labeled and annealed to RNA from wild-type *C. crescentus*. Reverse transcriptase was then added in the presence of actinomycin D, and the extended products were analyzed on an 8% sequencing gel. Two start sites were routinely observed as shown by the arrows. The same primer was used in a dideoxy-sequencing reaction to align the extended products with the proper nucleotide sequence (left). A portion of a conserved $\sigma^{54}$ promoter sequence upstream of one of the start sites is boxed.

Because the transcript that initiates at an appropriate distance from the consensus $\sigma^{54}$ promoter sequence gave a very weak signal in the primer extension assay, we used in vitro mutagenesis to determine whether the $\sigma^{54}$ sequences were required for *flbN* transcription in vivo. This promoter sequence, highly conserved among a variety of gram-negative bacteria (25, 37), requires the conserved nucleotides in both the −24 and −12 regions shown in boldface in Fig. 6A. Alteration of the conserved nucleotides invariably leads to a loss of transcription from this promoter (25, 37). We altered one of the conserved nucleotides in the −12 region. The wild-type (Fig. 6B, lane 1) and mutagenized (Fig. 6B, lane 2) promoters were fused to a *neo* transcription reporter gene and cloned in both orientations in the vector pRK290. A G→T transversion, shown in Fig. 6A, resulted in a loss of transcription in vivo (Fig. 6B). Only one orientation of the fusion in the vector is shown in Fig. 6B, although similar results were observed in the opposite orientation. The single base change at −14 resulted in a significant (>10-fold) decrease in transcription activity. This result suggests that *flbN* transcription uses a promoter sequence in the $\sigma^{54}$ consensus region and that the shorter transcript (at −23) may be a product of rapid and efficient RNA processing, as has been suggested for the similar *flaN* promoter (47).

**Effect of deletions in the 5′ regulatory region on temporal expression of the *flbN* gene.** The time of transcription activation from the intact *flbN* promoter driving a promoterless *neo* reporter gene in strain AE8002 was assayed by pulse-labeling synchronized cells with [$^{35}$S]methionine at successive stages of the *C. crescentus* cell cycle (see Fig. 8, top). Cell extracts were immunoprecipitated with antibody directed against the product of the *neo* gene, NPT II. Because AE8002 contains a transcription fusion, the level of NPT II synthesis reflects the transcription of the *flbN-neo* chimeric mRNA. As a control to monitor synchrony, a portion of each labeled sample of cells was immunoprecipitated with anti-flagellin antibody (see Fig. 8). Transcription activation of the chromosomal *flbN* gene fusion occurred at approximately 0.75 division units. In this assay, transcription was not detected at the earlier time point, which was 0.5 division units. Nuclease S1 protection assays of the native *flbN* mRNA (see Fig. 10B) revealed that activation occurred at 0.6 division units with a peak of activity at 0.75 division units. Transcript levels of the intact gene dropped significantly after the peak of activity (see Fig. 10B), but the expression of the *flbN-neo* gene fusion did not decrease by 1.0 division unit. This may reflect altered mRNA turnover of the chimeric mRNA from the *flbN* gene fusion.

Deletions were made in the 5′ regulatory region to identify

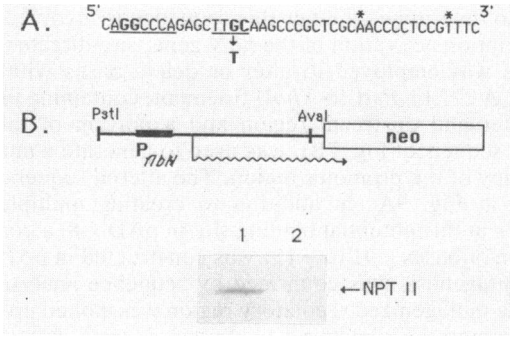


FIG. 6. In vitro mutagenesis of the conserved $\sigma^{54}$ promoter consensus sequence in the *flbN* 5′ region. (A) The conserved $\sigma^{54}$ promoter sequence and the identified transcription start sites (*). The partially conserved (underlined) and completely conserved (boldface) nucleotides in $\sigma^{54}$ promoters are indicated (37). A G→T substitution was made according to the mutagenesis procedure described in Materials and Methods. (B) In vivo transcription of the wild-type *flbN* promoter (lane 1) and the mutated *flbN* promoter with a G→T substitution (lane 2) fused to a promoterless *neo* reporter gene. Cells growing at mid-logarithmic phase ($OD_{660} \doteq 0.7$) were pulse-labeled with [$^{35}$S]methionine for 10 min. Cell extracts were prepared and immunoprecipitated with antibody directed against NPT II (*neo*). The labeled NPT II appears as a 25-kDa protein on 12% polyacrylamide–sodium dodecyl sulfate gels.

potential regulatory sequences necessary for transcription activation (Fig. 7A). The *flbN* gene containing a transposon insertion of a promoterless *neo* reporter gene was cloned from strain AE8002 (*flbN802*::Tn*5*-VB32). The clone, pADN-1, contains the entire promoter region for *flbN* up to the *Sac*I site and approximately two-thirds of the coding sequence. The *flbN-neo* fusion in pADN-1 was subcloned into M13mp19, and a series of nested deletions were made, as described in Materials and Methods. The promoter region deletions (pADN-2, pADN-3, and pADN-4), shown schematically in Fig. 7A, were sequenced to determine the endpoints indicated in Fig. 7B as N-2, N-3, and N-4, respectively. Each deletion fusion was then subcloned into pRK290 (14) and mated into *C. crescentus* CB15N in order to assay transcription activity. The deletion fusions in pADN-2, pADN-3, and pADN-4 were not transcribed above background levels in *C. crescentus*.

The activity of the pADN-2 deletion fusion in comparison with that of the intact upstream region in pADN-1 was assayed in synchronized cells (Fig. 8). The *flbN* promoter on the plasmid pADN-1 (three to five copies per cell) was temporally regulated in a manner analogous to the *flbN-neo* transcription fusion on the chromosome in strain AE8002 (Fig. 8). Therefore, proper cell cycle control of *flbN* transcription occurs whether the gene is on the chromosome or a plasmid. The higher background level of transcription of the *flbN-neo* fusion on pADN-1 could be due to a low level of transcription originating from plasmid sequences. Transcription from the plasmid-borne *flbN-neo* fusion (pADN-1) continued in the swarmer cell when *flbN* transcription is normally off (Fig. 8 and 10B). Although the time of transcription activation with pADN-1 was similar to that found for the chromosomal gene, the deletion in pADN-2 that removed all of the *flbN* upstream region but left the promoter intact displayed no detectable transcriptional activation at 0.75 division units. However, there appeared to be a transient increase in expression at 0.25 division units (Fig. 8) which corresponds to the time of initiation of chromosome replication (13). This small increase in transcription was consistently observed. Densitometry scanning of the autoradiograms showed that the fluctuation between the highest level of NPT II synthesis and the lowest varied by no more than twofold, whereas maximal activation at 0.75 division units usually exhibits an approximately 50-fold increase in the rate of transcription. The same pattern was observed when the construction was cloned in both orientations within the vector. Therefore, a *cis*-acting site resides between −47 and −246 that contributes to the temporal activation of *flbN*, demonstrating that the promoter alone is insufficient for transcription activation during the cell cycle.

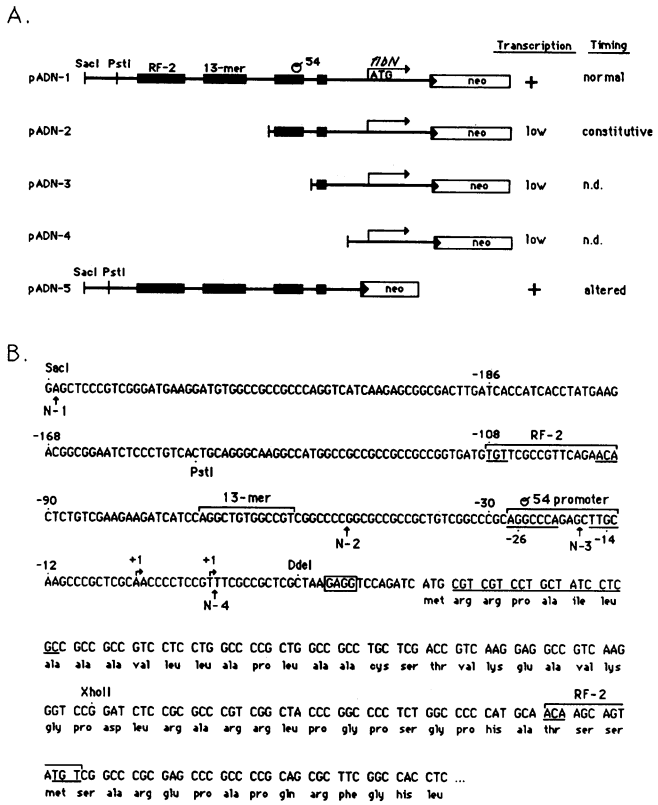


FIG. 7. Deletion analysis of the *flbN* 5′ region. (A) Schematic of the 5′ portion of the *flbN* gene and deletions of the 5′ region fused to a *neo* transcription reporter gene. Transcription from the *flbN* promoter was assayed by using a promoterless *neo* reporter gene. The deletions within the promoter region were generated by the method of Dale et al. (11) and were sequenced to determine the endpoints. A summary of transcription and timing for each of the *flbN* gene fusions is shown to the right of the schematic. n.d., Not determined. (B) Sequence of the *flbN* 5′ region. Potential regulatory protein binding sites are shown by brackets. The 5′ RF-2 site contains the sequence $TGTN_{12}ACA$, which is similar to the consensus NifA protein binding site (underlined nucleotides) for several $\sigma^{54}$ regulated promoters (37). A conserved 13-mer element found in the upstream region of several *fla* genes is indicated. The promoter region has a $\sigma^{54}$ promoter consensus sequence. The predicted Shine-Dalgarno sequence for *flbN* (boxed) is indicated, as is a partial coding region for *flbN*. Deletion endpoints (↑) are shown below the sequence. Mapped transcription start sites (⌐→) for *flbN* also are shown.

In order to determine whether sequence elements 3′ to the start site contribute to the time of transcription activation, sequences downstream from the *flbN* promoter, which include the entire *flbN* coding region, were deleted. The plasmid, pADN-5 (Fig. 7A), contains the promoter region from the *Dde*I site up to the *Sac*I site fused to a promoterless

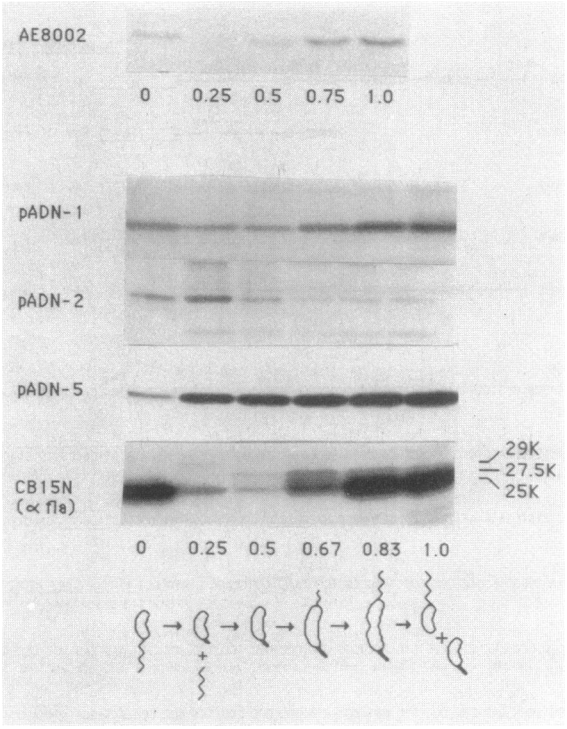


FIG. 8. Expression of *flbN-neo* transcription fusions as a function of the cell cycle. The transcription of *flbN* was measured by using promoterless reporter genes in both a single-copy chromosomal insertion (AE8002, top) and fusions carried on plasmids pADN-1, pADN-2, and pADN-5 in a wild-type strain, CB15N. Cell cultures were synchronized as described in Materials and Methods. Samples were pulse-labeled with [$^{35}$S]methionine at the times indicated. Cell extracts were prepared and immunoprecipitated with antibody to either NPT II or flagellins (bottom). The NPT II appears as a 25-kDa protein on 10% polyacrylamide–sodium dodecyl sulfate gels. The sizes of the three flagellins are indicated at the right. Numbers below the lanes represent fractions of the cell cycle expressed as division units, with 1.0 equivalent to one complete cell cycle. The drawing at the bottom of the figure is a schematic representation of the *C. crescentus* cell cycle that corresponds to the division units above.

*neo* reporter gene from Tn*5*. The 5′ end of the construction in pADN-5 is identical to that used in the construction of pADN-1. The expression of the pADN-5 fusion was at background level in the swarmer cell at 0 division units, similar to that observed with pADN-1. However, transcription activation occurred earlier, at 0.25 division units, and continued to increase until the end of the cell cycle (Fig. 8). The *flbN* promoter in pADN-5, while expressed earlier in the cell cycle, was still under positive control by the product of *flaS* (data not shown), indicating that positive control is exerted at sites upstream of the promoter. These experiments suggest that sequences 3′ to the promoter, together with an upstream site(s), contribute to correct timing of *flbN* expression.

**Identification of a *cis*-acting regulatory site.** $\sigma^{54}$ promoters are known to require an upstream *cis*-acting element for the activation of transcription (25, 37, 43, 44). Inspection of the *flbN* 5′ region revealed the presence of a sequence that is similar to a NifA protein binding site between −91 and −108 (Fig. 7B). Binding at a NifA site in *R. meliloti* and *Klebsiella pneumoniae* has been shown to be necessary for transcriptional activation of several $\sigma^{54}$ promoters (43, 44, 52, 55). In order to determine whether this sequence is involved in the transcription activation of the *flbN* gene, site-directed mutagenesis was employed to alter or delete bases within this region. A 277-bp *Pst*I-to-*Xho*II fragment, containing the *flbN* promoter and upstream region and a portion of the *flbN* coding sequence (Fig. 7B), was used to generate a mutagenized copy of the promoter region. The altered sequences are shown in Fig. 9A. In addition to creating multiple base changes in this potential binding site in pADN-8, a complete deletion of bases −91 to −111 was constructed in pADN-10. The mutagenesis was confirmed by sequence analysis, and then the mutagenized regulatory region was cloned upstream of a promoterless bacterial luciferase operon shown schematically in Fig. 9B. The 5′ end of the fusion was identical to that used in the *neo* transcription reporter gene fusion shown in Fig. 6.

Each construction (pADN-7, pADN-8, and pADN-10) was mated into *C. crescentus* CB15N, and the synthesis of luciferase was measured as a function of the cell cycle by counting photon emission at discrete time intervals (Fig. 10A). The expression of the reporter gene transcription fusions was compared with the temporal transcription pattern of the wild-type chromosomal *flbN* gene (Fig. 10B). Transcription from the chromosomal copy of the *flbN* gene was determined by nuclease S1 analysis of the mRNA taken from different stages of the *C. crescentus* cell cycle, as described in Materials and Methods. The 620-base *Pst*I-to-*Ava*I fragment shown in Fig. 1A was used as a probe in an S1 protection assay. A band that migrated as a 470-base fragment was detected. An autoradiograph of the protection analysis was scanned by densitometry, and the results were plotted as shown in Fig. 10B. Transcription activation of the wild-type *flbN* gene on the chromosome occurred at 0.6 division units and peaked at 0.75 division units. Following the activation period, there was a significant decrease in the synthesis of the *flbN* mRNA at the end of the cell cycle prior to division. The time of transcription activation of the *flbN* promoter in plasmid pADN-7 was similar to that observed for the wild-type *flbN* gene measured directly by nuclease S1 assays. As observed previously, there was a slight increase in activity at 0.25 division units. A major increase in activity occurred between 0.6 and 0.75 division units, and then activity decreased at cell division. In contrast, cell cycle transcription from the *flbN* promoter was only partially activated when sequences within the predicted binding site were changed or deleted (pADN-8 and pADN-10; Fig. 10A). Although pADN-8 and pADN-10 showed an increase in transcription at 0.7 division units, it was half the level observed with the intact *flbN* 5′ region (pADN-7). The partial activation of the mutagenized promoter, in comparison to the wild-type pattern of *flbN* transcription both on the chromosome and on a plasmid, indicates that a regulatory site, which we call RF-2, is necessary for full temporal activation of the *flbN* gene.

**The RF-2 site is functionally homologous to a NifA binding site.** The RF-2 binding site within the *flbN* upstream region ($TGTN_{12}ACA$) is similar to the consensus NifA binding site ($TGTN_{9-11}ACA$) found upstream of some $\sigma^{54}$ promoters (25, 43, 44). In order to test whether a NifA protein could recognize and activate *flbN* transcription, the intact *flbN* regulatory region fused to a promoterless *lux* operon (pADN-7) and derivatives altered in the RF-2 site (pADN-8 and pADN-10) were mated into *E. coli* strains that either lacked or contained the *R. meliloti nifA* gene on plasmid pRMB3.8H (58). The intact *flbN* promoter on pADN-7 showed no luciferase activity in an *E. coli* strain that lacked

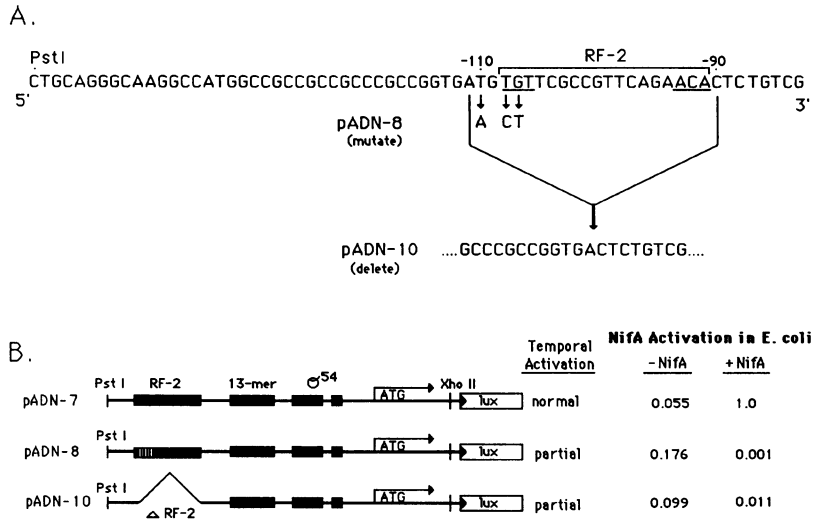


FIG. 9. Schematic of site-directed mutagenesis of the RF-2 site. (A) The base changes and the deletion of the predicted binding site for RF-2 are shown. (B) Transcriptional activity of the altered 5' region was determined by fusing the *flbN* promoter region to a promoterless luciferase operon, *luxCDABE*, and then mating the constructions into *C. crescentus* CB15N. Luciferase activity was assayed as described in Materials and Methods. A summary of the temporal activation of the *flbN* gene fusions in *C. crescentus* and the expression of the fusions in *E. coli* in the presence and absence of a *nifA* gene is shown to the right of the gene fusions. The time of reporter gene expression was tested in synchronized cultures of *C. crescentus*. The three constructions were also placed in *E. coli*, and expression of the *flbN* promoter was determined in both the presence and absence of the *R. meliloti nifA* gene contained on a second plasmid (pRMB3.8H).

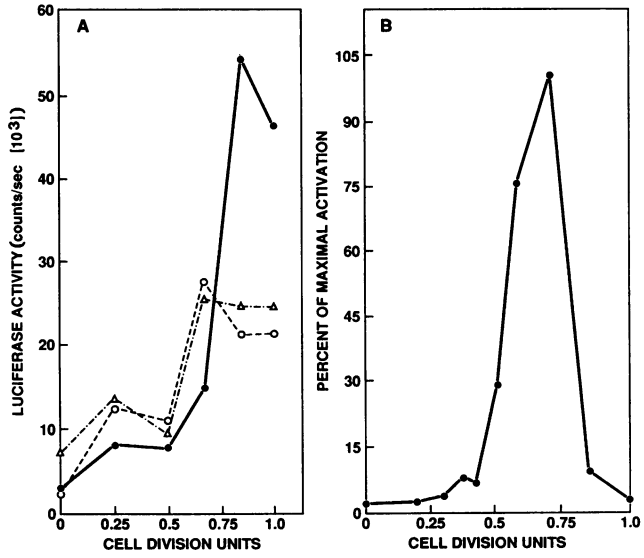


FIG. 10. Temporal expression of the *flbN-lux* reporter gene fusions and the wild-type gene. (A) Synchronized cultures of CB15N carrying plasmids pADN-7 (●), pADN-8 (○), and pADN-10 (△) were tested for luciferase activity (counts per second based on an $OD_{660}$ of 1.0) as a function of the cell cycle. (B) Detection of *flbN* mRNA from the intact chromosomal gene as a function of the cell cycle by nuclease S1 protection assays. RNA was prepared from cells within a single cell cycle from wild-type strain CB15N at the stages indicated. A 620-base *Pst*I-to-*Ava*I fragment (Fig. 1A) was end labeled at the *Ava*I site and used in the protection experiment. Densitometry tracings were made from the autoradiographs to determine the relative levels of *flbN* mRNA. The cell division units correspond to the same stages of the *C. crescentus* cell cycle shown in Fig. 6.

a *nifA* gene. However, in the presence of a plasmid containing the *nifA* gene, *flbN* transcription was stimulated approximately 20-fold (Fig. 9B). The NifA protein provided in *trans* from the *R. meliloti* gene failed to stimulate transcription when the RF-2 binding site was altered or deleted (pADN-8 and pADN-10; Fig. 9B). Therefore, the RF-2 site can be recognized and used by the *R. meliloti* NifA to activate *flbN* transcription in *E. coli*, lending further support to the conclusion that the *flbN* promoter is of the $\sigma^{54}$ class.

**Protein complex formation with an RF-2 oligonucleotide.** In order to determine whether *C. crescentus* cell extracts contain a protein that binds to the RF-2 site, a $^{32}$P-labeled oligonucleotide with the sequence shown in Fig. 11A was mixed with crude cell extracts in a gel mobility shift assay (Fig. 11B). Mixtures of probe and cell extract were subjected to electrophoresis under nondenaturing conditions, as described in Materials and Methods. Nonspecific competitor DNA (pBR325) or cold probe were added, as indicated, in order to determine whether the observed gel retardation complexes were due to specific binding. Two complexes were detected by the mobility shift assay. Complex II was the predominant form, although both complexes were competed away with the same efficiency by the addition of cold oligonucleotide. Neither complex I nor complex II was competed away with nonspecific pBR325 DNA (lanes 2 through 6). These results suggest that a protein(s) in *C. crescentus* binds to the RF-2 region. The protein that binds to the RF-2 site has been shown by Southwestern analysis of the RF-2 complexes to be a single protein of approximately 70 kDa (J. W. Gober, unpublished results).

## DISCUSSION

The biogenesis of the *C. crescentus* flagellum occurs in a tightly controlled synthetic pathway in which the time of transcription initiation of individual flagellar genes reflects the order of assembly of the flagellar components. Furthermore, the time of flagellum synthesis is tightly coupled to the

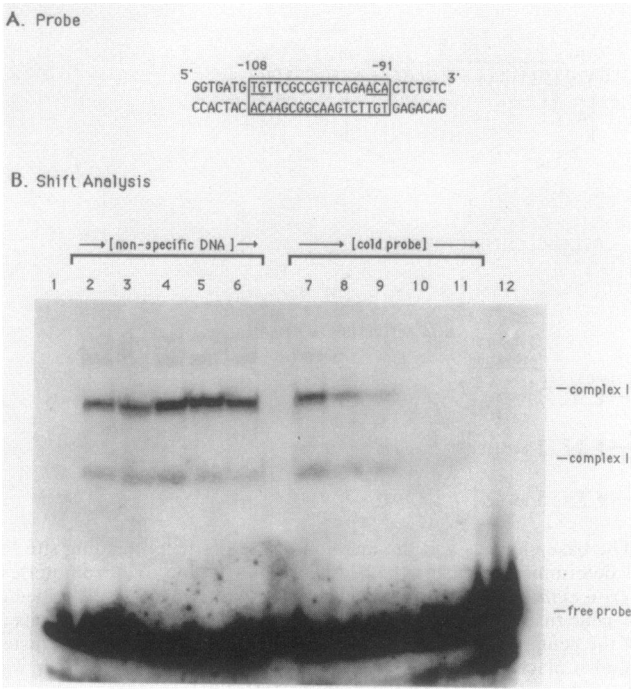


FIG. 11. DNA-protein complex formation between an oligonucleotide containing the RF-2 site and *C. crescentus* cell extracts. Whole-cell extracts from a wild-type strain (CB15N) were incubated with the $^{32}$P-labeled probe that contains the putative RF-2 binding site (boxed in panel A). Underlined sequences conform to the consensus NifA binding site. Binding mixtures were analyzed for specific association with the labeled probe by mobility shift assays shown in panel B. Labeled probe (2.25 ng) was incubated with 2 μg of protein and mixed with increasing amounts of either nonspecific competitor DNA (pBR325; lanes 2 to 6), or cold probe DNA (lanes 7 to 11): lane 2, 250 ng; lane 3, 500 ng; lane 4, 1,000 ng; lane 5, 1.75 μg; lane 6, 3.5 μg; lane 7, 0.8 ng; lane 8, 1.75 ng; lane 9, 3.5 ng; lane 10, 7.0 ng; lane 11, 14 ng. Lanes 1 and 12, Controls with no protein. Complexes were separated by electrophoresis through nondenaturing 5% (wt/vol) acrylamide gels in 1× TBE buffer.

cell cycle. To understand the regulatory signals that control the activation of *fla* gene transcription, we have studied a flagellar gene, *flbN*, that is required early in the assembly pathway.

The *flbN* gene was cloned and sequenced. An open reading frame was contained within a restriction fragment that complemented two chromosomal *flbN* insertion mutations. This open reading frame could encode a 25-kDa protein with a cleavable hydrophobic signal sequence often associated with transported proteins. The derived *flbN* protein sequence had 30.8% identity with the *S. typhimurium* basal body L-ring protein (*flgH*). The L-ring is the outermost ring of the basal body and appears to be associated with the outer membrane.

Nuclease S1 analysis was used to determine the location of both the 5′ and 3′ ends of the *flbN* transcript. These experiments indicated that *flbN* transcription initiated approximately 35 bases 5′ to the putative open reading frame and that transcription terminated imprecisely about 70 bases downstream of the translation stop codon. We previously established that *flbN* transcription was subject to positive control by the *flaS* gene (9, 60). This was confirmed by assays of both the 5′ and 3′ termini showing that the chromosomal *flbN* transcript is regulated by the product of the *flaS* gene.

The nuclease S1 assays revealed two apparent transcription start sites, approximately 10 bases apart, that were both down regulated in a *flaS* mutant background. Primer extension analysis confirmed these two apparent start sites and revealed that the larger transcript initiated 13 bases downstream from a consensus $\sigma^{54}$ promoter sequence. Although the shorter transcript was more abundant, there was no recognizable promoter sequence at the appropriate distance from its putative start site. The shorter transcript appeared to initiate 23 bases downstream from the consensus $\sigma^{54}$ promoter. Possible explanations for these two apparent start sites are that the more abundant transcript is read from a new promoter class or that it results from the rapid and efficient processing of the larger transcript that could be read from the consensus $\sigma^{54}$ promoter. The fact that both transcripts were identically regulated in a *flaS* background (Fig. 4) and that both were temporally controlled in an identical manner (A. Dingwall, unpublished results) suggest that they originated from a single promoter. Mullin and Newton (47) have recently demonstrated that another flagellar gene, *flaN*, is transcribed from a $\sigma^{54}$-like promoter, yet the only detectable transcript initiates 23 bases downstream of the promoter sequence. Brun et al. (5) have also reported that processing of the *E. coli gltX* primary transcript produced a shorter, more-abundant transcript.

There are several lines of evidence to support the possibility that the *flbN* transcript is read from the $\sigma^{54}$ promoter consensus sequence. (i) The larger transcript is initiated at the correct distance from this highly conserved promoter sequence that is nearly identical to several *C. crescentus* flagellar gene promoters (*flbG*, *flaN*, *flgK*, and *flgL*) that have been demonstrated previously to be recognized by *E. coli* $\sigma^{54}$ RNA polymerase holoenzyme in an in vitro transcription assay (49). (ii) A mutation generated at −14 in the promoter sequence, which changed a completely conserved G residue to a T, abolished *flbN* transcription. Similar analyses have previously been used to establish the identity of $\sigma^{54}$ promoters in a variety of organisms (25), including *C. crescentus* (47). (iii) In addition to conserved nucleotides in the promoter sequence, $\sigma^{54}$ promoters require an upstream activator site for transcription initiation (25, 37). Mutation and deletion analysis of the *flbN* 5′ region and gel retardation assays showed that an upstream protein binding site was necessary for full activation of *flbN* transcription (Fig. 7 and 9). The RF-2 site, located approximately 100 bases upstream of the transcript start site, is similar to the NifA binding sites from both *R. meliloti* (7) and *K. pneumoniae* (43, 44), although there is a difference in the spacing of the consensus half-sites, 12 versus the more usual 10 nucleotides. NifA is known to activate transcription from several $\sigma^{54}$-dependent promoters (7, 37, 44, 55). (iv) The *flbN* gene cannot be transcribed in an *E. coli* host unless the *R. meliloti* NifA protein is provided in *trans*. In the presence of a plasmid-borne *nifA* gene, transcription of the *flbN* gene was stimulated approximately 20-fold (Fig. 9). In similar experiments performed in *E. coli*, Morett et al. (44) showed that the *R. meliloti* NifA protein could substitute for the *K. pneumoniae* NifA in its ability to activate several *K. pneumoniae* $\sigma^{54}$ promoters. An unexpected result was the apparent repression of *flbN* transcription by NifA when the RF-2 site was altered or deleted. One possibility is that these mutations may result in NifA binding at other potential sites in the *flbN* promoter region, thereby occluding recognition and transcription from its promoter.

In vitro mutagenesis of the RF-2 *cis*-acting site demonstrated that this region is required for full activation of *flbN*

transcription at 0.7 division units; approximately 50% of the activity remained after alteration or deletion of the RF-2 site. The high level of remaining activity was somewhat surprising in that most $\sigma^{54}$ promoters require an intact upstream activator site (UAS) for full transcription. However, there are several examples of situations in which either no UAS is required for NifA-dependent activation (59) or an altered UAS can substitute (44). Given that an intact UAS is not absolutely required for $\sigma^{54}$ promoter transcription, we reason that in the case of *flbN*, other potential imperfect binding sites for RF-2 may be utilized or that the transcription observed in pADN-8 and pADN-10 may be due to activation by RF-2 in *trans* when bound at a different location (45), perhaps on the chromosome. The second possibility is consistent with the postulated mechanism of action for activator proteins that require a DNA loop to be formed (22, 29, 41). Deletion analysis showed that there was no activation of the *flbN* promoter in the absence of the entire upstream region from position −47 (pADN-2). Therefore, other regulatory elements must work in concert with the RF-2 site and the promoter for full activation.

Sequences downstream from the transcription start site also appear to play a role in the temporal control of *flbN* transcription. Transcription from the *flbN* promoter in pADN-5, a construction which retained the entire 5′ upstream region but deleted the *flbN* coding region, was activated at an earlier stage in the cell cycle than the intact gene. One possible explanation for the altered temporal control is that the pADN-5 construct results in a deletion of a sequence between +161 and +174 that is similar to the upstream RF-2 binding site (Fig. 7B). Since the effect of the deletion in pADN-5 is to allow transcription at an earlier time in the cell cycle, we reason that the putative 3′ RF-2 binding site might be involved in modulating transcription from the *flbN* promoter. This modulation may be accomplished by the interaction of the RF-2 protein bound to both 3′ and 5′ sites through the formation of a DNA loop. Such loop formation would be functionally analogous to the system of alternate loops required for the regulation of the *ara* operon in *E. coli* by the AraC protein (29).

## ACKNOWLEDGMENTS

We thank F. M. Ausubel for providing the plasmid pRMB3.8H and C. Kado for plasmid pUCD615. We also thank B. Ely and Idene Saam for performing the data base search.

This investigation was supported by U.S. Public Health Service grant GM32506 from the National Institutes of Health. A.D. received predoctoral support from Public Health Service Training Grant GM-07491-07 from the National Institutes of Health, and J.W.G. is a Helen Hay Whitney Fellow.

## LITERATURE CITED

1. **Amemiya, K., B. Raboy, and L. Shapiro.** 1980. Involvement of host RNA polymerase in the early transcription program of *Caulobacter crescentus* bacteriophage φCd1 DNA. Virology **104**:109–116.
2. **Arnosti, D. N., and M. J. Chamberlin.** 1989. Secondary σ factor controls transcription of flagellar and chemotaxis genes in *Escherichia coli*. Proc. Natl. Acad. Sci. USA **86**:830–834.
3. **Bellofatto, V., L. Shapiro, and D. Hodgson.** 1984. Generation of a Tn*5* promoter probe and its use in the study of gene expression in *Caulobacter crescentus*. Proc. Natl. Acad. Sci. USA **81**: 1035–1039.
4. **Berk, A. J., and P. A. Sharp.** 1977. Sizing and mapping of early adenovirus mRNAs by gel electrophoresis of S1 endonuclease-digested hybrids. Cell **12**:721–732.
5. **Brun, Y. V., R. Breton, P. Lanouette, and J. Lapointe.** 1990. Precise mapping and comparison of two evolutionarily related regions of the *Escherichia coli* K-12 chromosome. J. Mol. Biol. **214**:845–864.
6. **Bryan, R., M. Purucker, S. L. Gomes, W. Alexander, and L. Shapiro.** 1984. Analysis of the pleiotropic regulation of flagellar and chemotaxis gene expression in *Caulobacter crescentus* using plasmid complementation. Proc. Natl. Acad. Sci., USA **81**:1341–1345.
7. **Buck, M., S. Miller, M. Drummond, and R. Dixon.** 1986. Upstream activator sequences are present in the promoters of nitrogen fixation genes. Nature (London) **320**:374–378.
8. **Champer, R., R. Bryan, S. L. Gomes, M. Purucker, and L. Shapiro.** 1985. Temporal and spatial control of flagellar and chemotaxis gene expression during *Caulobacter* cell differentiation. Cold Spring Harbor Symp. Quant. Biol. **50**:831–840.
9. **Champer, R., A. Dingwall, and L. Shapiro.** 1987. Cascade regulation of *Caulobacter* flagellar and chemotaxis genes. J. Mol. Biol. **194**:71–80.
10. **Contreras, I., L. Shapiro, and S. Henry.** 1978. Membrane phospholipid composition of *Caulobacter crescentus*. J. Bacteriol. **135**:1130–1136.
11. **Dale, R. M. K., B. A. McClure, and J. P. Houchins.** 1985. A rapid single-stranded cloning strategy for producing a sequential series of overlapping clones for use in DNA sequencing: application to sequencing the corn mitochondrial 18S rDNA. Plasmid **13**:31–40.
12. **Devereaux, J., P. Haerlberli, and O. Smithies.** 1984. A comprehensive set of sequence analysis programs for the VAX. Nucleic Acids Res. **12**:387–395.
13. **Dingwall, A., and L. Shapiro.** 1989. Rate, origin, and bidirectionality of *Caulobacter* chromosome replication as determined by pulsed-field gel electrophoresis. Proc. Natl. Acad. Sci. USA **86**:119–123.
14. **Ditta, G., S. Stanfield, D. Corbin, and D. R. Helinski.** 1980. Broad host range cloning system for gram-negative bacteria: construction of a gene bank for *Rhizobium meliloti*. Proc. Natl. Acad. Sci. USA **77**:7347–7351.
15. **Ely, B., R. H. Croft, and C. J. Gerardot.** 1984. Genetic mapping of genes required for motility in *Caulobacter crescentus*. Genetics **180**:523–532.
16. **Ely, B., and T. W. Ely.** 1989. Use of pulsed field gel electrophoresis and transposon mutagenesis to estimate the minimal number of genes required for motility in *Caulobacter crescentus*. Genetics **123**:649–654.
17. **Ely, B., C. J. Gerardot, D. L. Fleming, S. L. Gomes, P. Frederikse, and L. Shapiro.** 1986. General non-chemotactic mutants of *Caulobacter crescentus*. Genetics **114**:717–730.
18. **Evinger, M., and N. Agabian.** 1977. Envelope-associated nucleoid from *Caulobacter crescentus* stalked and swarmer cells. J. Bacteriol. **132**:294–301.
19. **Fisher, R. F., T. T. Egelhoff, J. T. Mulligan, and S. R. Long.** 1988. Specific binding of proteins from *Rhizobium meliloti* cell-free extracts containing NodD to DNA sequences upstream of inducible nodulation genes. Genes Dev. **2**:282–293.
20. **Frederikse, P. H., and L. Shapiro.** 1989. An *Escherichia coli* chemoreceptor gene is temporally controlled in *Caulobacter*. Proc. Natl. Acad. Sci. USA **86**:4061–4065.
21. **Fried, M., and D. M. Crothers.** 1981. Equilibria and kinetics of *lac* repressor-operator interactions by polyacrylamide gel electrophoresis. Nucleic Acids Res. **9**:6505–6525.
22. **Gober, J. W., and L. Shapiro.** 1990. Integration Host Factor (IHF) is required for the activation of developmentally regulated genes in *Caulobacter*. Genes Dev. **4**:1494–1504.
23. **Gomes, S. L., J. W. Gober, and L. Shapiro.** 1990. The expression of the *Caulobacter* heat shock gene *dnaK* is developmentally controlled during growth at normal temperatures. J. Bacteriol. **172**:3051–3059.
24. **Gomes, S. L., and L. Shapiro.** 1984. Differential expression and position of chemotaxis methylation proteins in *Caulobacter crescentus*. J. Mol. Biol. **178**:551–568.
25. **Gussin, G. N., C. W. Ronson, and F. M. Ausubel.** 1986. Regulation of nitrogen fixation genes. Annu. Rev. Genet. **20**: 567–591.
26. **Hahnenberger, K. M., and L. Shapiro.** 1987. Identification of a

gene cluster involved in flagellar basal body biogenesis in *Caulobacter crescentus*. J. Mol. Biol. **194**:91–103.
27. **Hahnenberger, K. M., and L. Shapiro.** 1988. Organization and temporal expression of a flagellar basal body gene in *Caulobacter crescentus*. J. Bacteriol. **170**:4119–4124.
28. **Hirooka, T., P. M. Rogowsky, and C. I. Kado.** 1987. Characterization of the *virE* locus of *Agrobacterium tumefaciens* plasmid pTiC58. J. Bacteriol. **169**:1529–1536.
29. **Huo, L., K. J. Martin, and R. Schleif.** 1988. Alternative DNA loops regulate the arabinose operon in *Escherichia coli*. Proc. Natl. Acad. Sci. USA **85**:5444–5448.
30. **Jones, C. J., M. Homma, and R. M. Macnab.** 1989. L-, P-, and M-ring proteins of the flagellar basal body of *Salmonella typhimurium*: gene sequences and deduced protein sequences. J. Bacteriol. **171**:3890–3900.
31. **Jones, K. A., K. R. Yamamoto, and R. Tijan.** 1985. Two distinct transcription factors bind to the HSV thymidine kinase promoter *in vitro*. Cell **42**:559–572.
32. **Kaplan, J. B., A. Dingwall, R. Bryan, R. Champer, and L. Shapiro.** 1989. Temporal regulation and overlap organization of two *Caulobacter* flagellar genes. J. Mol. Biol. **205**:71–83.
33. **Katsukake, K., Y. Ohya, and T. Iino.** 1990. Transcriptional analysis of the flagellar regulon of *Salmonella typhimurium*. J. Bacteriol. **172**:741–747.
34. **Komeda, Y.** 1982. Fusions of flagellar operons to lactose genes on a Mu *lac* bacteriophage. J. Bacteriol. **150**:16–26.
35. **Komeda, Y.** 1986. Transcriptional control of flagellar genes in *Escherichia coli* K-12. J. Bacteriol. **168**:1315–1318.
36. **Kunkel, T. A., and J. D. Roberts.** 1987. Rapid and efficient site-specific mutagenesis without phenotypic selection. Methods Enzymol. **154**:367–382.
37. **Kustu, S., E. Santero, J. Keener, D. Popham, and D. Weiss.** 1989. Expression of $\sigma^{54}$ (ntrA)-dependent genes is probably united by a common mechanism. Microbiol. Rev. **53**:367–376.
38. **Kyte, J., and R. F. Doolittle.** 1982. A simple method for displaying the hydropathic character of a protein. J. Mol. Biol. **157**:105–132.
39. **Losick, R., L. Kroos, J. Errington, and P. Youngman.** 1989. Pathways of developmentally regulated gene expression in the spore-forming bacterium *Bacillus subtilis*, p. 221–240. *In* D. A. Hopwood and K. F. Chater (ed.), Genetics of bacterial diversity. Harcourt, Brace and Jovanovich, London.
40. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
41. **Minchin, S. D., S. Austin, and R. A. Dixon.** 1989. Transcriptional activation of the *Klebsiella pneumoniae nifLA* promoter by NTRC is face-of-the-helix dependent and the activator stabilizes the interaction of sigma 54-RNA polymerase with the promoter. EMBO J. **8**:3491–3499.
42. **Minnich, S. A., and A. Newton.** 1987. Promoter mapping and cell cycle regulation of flagellin gene transcription in *Caulobacter crescentus*. Proc. Natl. Acad. Sci. USA **84**:1142–1146.
43. **Morett, E., and M. Buck.** 1988. NifA-dependent *in vivo* protection demonstrates that the upstream activator sequence of *nif* promoters is a protein binding site. Proc. Natl. Acad. Sci. USA **85**:9401–9405.
44. **Morett, E., W. Cannon, and M. Buck.** 1988. The DNA-binding domain of the transcriptional activator protein NifA resides in its carboxy terminus, recognises the upstream activator sequences of *nif* promoters and can be separated from the positive control function of NifA. Nucleic Acids Res. **16**:11469–11488.
45. **Müller, H.-P., J. M. Sogo, and W. Schaffner.** 1989. An enhancer stimulates transcription in *trans* when attached to the promoter via a protein bridge. Cell **58**:767–777.
46. **Mullin, D., S. Minnich, L. S. Chen, and A. Newton.** 1987. A set of positively regulated flagellar gene promoters in *Caulobacter crescentus* with sequence homology to the *nif* gene promoters of *Klebsiella pneumoniae*. J. Mol. Biol. **195**:939–943.
47. **Mullin, D. A., and A. Newton.** 1989. Ntr-like promoters and upstream regulatory sequence *ftr* are required for transcription of a developmentally regulated *Caulobacter crescentus* flagellar gene. J. Bacteriol. **171**:3218–3227.
48. **Newton, A., N. Ohta, G. Ramakrishnan, D. Mullin, and G. Raymond.** 1989. Genetic switching in the flagellar gene hierarchy of *Caulobacter* requires negative as well as positive regulation of transcription. Proc. Natl. Acad. Sci. USA **86**:6651–6655.
49. **Ninfa, A. J., D. A. Mullin, G. Ramakrishnan, and A. Newton.** 1989. *Escherichia coli* $\sigma^{54}$ RNA polymerase recognizes *Caulobacter crescentus flbG* and *flaN* flagellar gene promoters in vitro. J. Bacteriol. **171**:383–391.
50. **Ohta, N., L. S. Chen, E. Swanson, and A. Newton.** 1985. Transcriptional regulation of a periodically controlled flagellar gene operon in *Caulobacter crescentus*. J. Mol. Biol. **186**:107–115.
51. **Oliver, D.** 1985. Protein secretion in *Escherichia coli*. Annu. Rev. Microbiol. **39**:615–648.
52. **Pasloske, B. L., D. S. Drummond, L. S. Frost, and W. Paranchych.** 1989. The activity of the *Pseudomonas aeruginosa* pilin promoter is enhanced by an upstream regulatory site. Gene **81**:25–34.
53. **Poindexter, J. S.** 1964. Biological properties and classification of the *Caulobacter* group. Bacteriol. Rev. **28**:231–295.
54. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74**:5463–5467.
55. **Santero, E., T. Hoover, J. Keener, and S. Kustu.** 1989. *In vitro* activity of the nitrogen fixation regulatory protein NIFA. Proc. Natl. Acad. Sci. USA **86**:7346–7350.
56. **Schoenlein, P. V., L. S. Gallman, and B. Ely.** 1989. Organization of the *flaFG* gene cluster and identification of two additional genes involved in flagellum biogenesis in *Caulobacter crescentus*. J. Bacteriol. **171**:1544–1554.
57. **Shapiro, L., J. Mansour, P. Shaw, and S. Henry.** 1982. Synthesis of specific membrane proteins is a function of DNA replication and phospholipid synthesis in *Caulobacter crescentus*. J. Mol. Biol. **159**:303–322.
58. **Szeto, W. W., J. L. Zimmerman, V. Sundaresan, and F. M. Ausubel.** 1984. A *Rhizobium meliloti* symbiotic regulatory gene. Cell **36**:1035–1043.
59. **Thöny, B., D. Anthamatten, and H. Hennecke.** 1989. Dual control of the *Bradyrhizobium japonicum* symbiotic nitrogen fixation regulatory operon *fixR nifA*: analysis of *cis*- and *trans*-acting elements. J. Bacteriol. **171**:4162–4169.
60. **Xu, H., A. Dingwall, and L. Shapiro.** 1989. Negative transcriptional regulation in the *Caulobacter* flagellar hierarchy. Proc. Natl. Acad. Sci. USA **86**:6656–6660.