JOURNAL OF BACTERIOLOGY, Aug. 1995, p. 4372–4376 Vol. 177, No. 15
0021-9193/95/$04.00+0

A Consensus Promoter Sequence for *Caulobacter crescentus* Genes Involved in Biosynthetic and Housekeeping Functions

JALEH MALAKOOTI,† SHUI PING WANG, AND BERT ELY*

Department of Biological Sciences, University of South Carolina, Columbia, South Carolina 29208

Received 7 November 1994/Accepted 12 May 1995

Caulobacter crescentus **differentiates prior to each cell division to form two different daughter cells: a monoflagellated swarmer cell and a nonmotile stalked cell. Thus, one might expect that developmentally expressed genes would be regulated by mechanisms different from those used to regulate the expression of the biosynthetic genes. To determine a consensus promoter sequence for genes involved in biosynthetic or housekeeping functions, DNA fragments containing the regulatory regions of the *ilvD*, *ilvR*, *cysC*, *pleC*, and *fdxA* genes were cloned. S1 nuclease protection mapping and primer extension techniques were used to identify the transcription initiation sites. Comparison of the regulatory regions of these genes with those of the published sequences of the *ilvBN*, *rrnA*, *trpFBA*, *dnaA*, *dnaK*, *hemE*, and *rsaA* genes has resulted in the identification of a putative promoter consensus sequence. The −35 region contains the sequence TTGACGS, which is similar to the *Escherichia coli* −35 region, while the −10 region, GCTANAWC, has a more balanced GC content than the corresponding region in *E. coli*. Oligonucleotide-directed site-specific mutagenesis of both the *ilvBN* and *pleC* promoters indicates that mutations that make a promoter more like the consensus result in increased promoter activity, while mutations decreasing similarity to the consensus result in decreased promoter activity.**

Caulobacter crescentus is a gram-negative bacterium that differentiates prior to each cell division to generate two dissimilar progeny cells. The new cells differ from one another morphologically and developmentally. Much of what is known about gene regulation in *C. crescentus* has resulted from studies of periodically expressed genes, especially the genes involved in flagellar biogenesis and function (21, 27). The 5′ regulatory regions of some of the flagellar genes have been shown to contain a set of activated promoters that are transcribed by a σ^{54} RNA polymerase holoenzyme (3, 19). These promoters contain the consensus sequences recognized by the enteric σ^{54} at positions −12 and −24 (4, 19, 22). Other *fla* genes may be regulated by a second alternative sigma factor (5, 30, 35).

Very little information is available on the expression of genes that are expected to show no cell cycle-dependent regulation. For instance, biosynthetic genes are likely to be expressed throughout the cell cycle to fulfill the nutritional requirements of the bacterium. To understand the regulatory mechanisms involved in the expression of the biosynthetic and housekeeping genes, we investigated the *cis*-acting regulatory elements involved in the expression of a number of such genes. Analysis of the sequences within the 5′ regulatory regions led us to propose a consensus promoter sequence for biosynthetic and housekeeping genes. Site-specific mutagenesis of the *ilvBN* and the *pleC* promoter regions was used to evaluate the role of the nucleotides composing the proposed consensus.

MATERIALS AND METHODS

Bacterial growth. *C. crescentus* strains were grown in PYE medium (13), and *Escherichia coli* strains were grown in L broth (18). Growth media were supplemented with the following concentrations of antibiotics when required: for *C. crescentus*, tetracycline at 1 μg/ml, chloramphenicol at 1 μg/ml, and ampicillin at 20 μg/ml; for *E. coli* strains, tetracycline at 10 μg/ml, chloramphenicol at 25 μg/ml, ampicillin at 100 μg/ml, and kanamycin at 50 μg/ml.

Oligonucleotide mutagenesis and the construction of mutant of *ilvBN* promoter-*cat* fusion plasmids. A 260-bp *Sst*I-*Bam*HI DNA fragment containing the *ilvBN* promoter region (29) was cloned in pBluescript-II KS+ (Stratagene, La Jolla, Calif.). This phagemid was designated pJM2-89. A uracil-containing single-stranded template was prepared by introducing the plasmid pJM2-89 into *E. coli* CJ236 (14). The single-stranded template preparation and oligonucleotide mutagenesis procedure were those of the Bio-Rad Muta-Gene system essentially as described by Kunkel et al. (14). The oligonucleotides used for mutagenesis had the following designations and sequences: BE70, ACGCATGC**T**ATA**A**ACCC TTTC; BE71, ACGCATGCC**GG**ACCCTTTC; BE72, GAGCGCCT**CC**ACG CCCCATC; and BE73, ACGCATG**T**CATGA**T**CCTTTC. (The substituted nucleotides are in boldface.) Three oligonucleotides, BE70, BE71, and BE73, contained changes in the −10 region, and the fourth one, BE72, contained changes in the −35 region. The oligonucleotides were phosphorylated with T4 polynucleotide kinase and annealed to pJM2-89 single-stranded DNA. DNA polymerization and ligation were carried out concomitantly in the presence of all four deoxynucleoside triphosphates (dNTPs), T4 DNA polymerase, and T4 DNA ligase. The hybrid molecules were used to transform *E. coli* XL1-Blue (Stratagene). The mutated clones were identified by DNA sequence analysis using the Sequenase version 2.0 kit (Amersham/United States Biochemical Corp., Cleveland, Ohio). The mutated promoter fragments were isolated and cloned upstream from a promoterless *cat* gene in the pBluescript KS+ derivative pJM2-90. Next, the new plasmids were linearized at the single *Sst*I site and cloned into the *C. crescentus*-compatible plasmid pRK2L1 (20). To prevent read-through transcription from the *oriT* region of pRK2L1 into the *ilvBN* promoter, only recombinant plasmids in which the *ilvBN* sequences were distal to *oriT* were selected for further experiments. These plasmids were introduced into *C. crescentus*, and chloramphenicol acetyltransferase (CAT) activity was measured in whole-cell extracts by using [^{14}C]acetyl coenzyme A as directed by the manufacturer (New England Nuclear, Boston, Mass.).

Oligonucleotide mutagenesis and analysis of the *pleC* promoter. Two oligonucleotides were synthesized for oligonucleotide mutagenesis of the *pleC* promoter by PCR. Oligonucleotide BE235 (5′-CGGAATTCGAGCCCACCAAC CCCAATCGG-3′) matches the sequence of the region 480 bp upstream of the transcription start site of *pleC* (33). A second oligonucleotide, BE234 (5′-ACAG GATCCCTTCGCCAAGTCGTGRAYCYARCGTRCCGC-3′), was designed to introduce a variety of base pair substitutions in the −10 region of the *pleC* promoter. For cloning purposes, *Eco*RI and *Bam*HI restriction sites were included at the ends of BE235 and BE234, respectively. The oligonucleotides were used to amplify the *pleC* promoter region from pSCW401 (33) as a template. To obtain different combinations of mutations, the PCR was done in four combinations of Mg^{2+} concentrations and annealing temperatures. The PCR products were cloned into the pBluescript KS + vector after digestion with *Eco*RI and *Bam*HI. Mutant clones were identified by DNA sequence analysis. The *Eco*RI-*Bam*HI fragments of mutant clones chosen for further study were purified from polyacrymide gels and cloned in front of the *lacZ* gene in the transcription fusion vector, p*lac*/290 (9). The recombinant clones were confirmed by restriction analysis and Southern hybridization and introduced into *C. crescentus* CB15 by

* Corresponding author. Mailing address: University of South Carolina, Department of Biological Sciences, Columbia, SC 29208.

† Present address: Department of Genetics, College of Medicine, University of Illinois at Chicago, Chicago IL 60680.

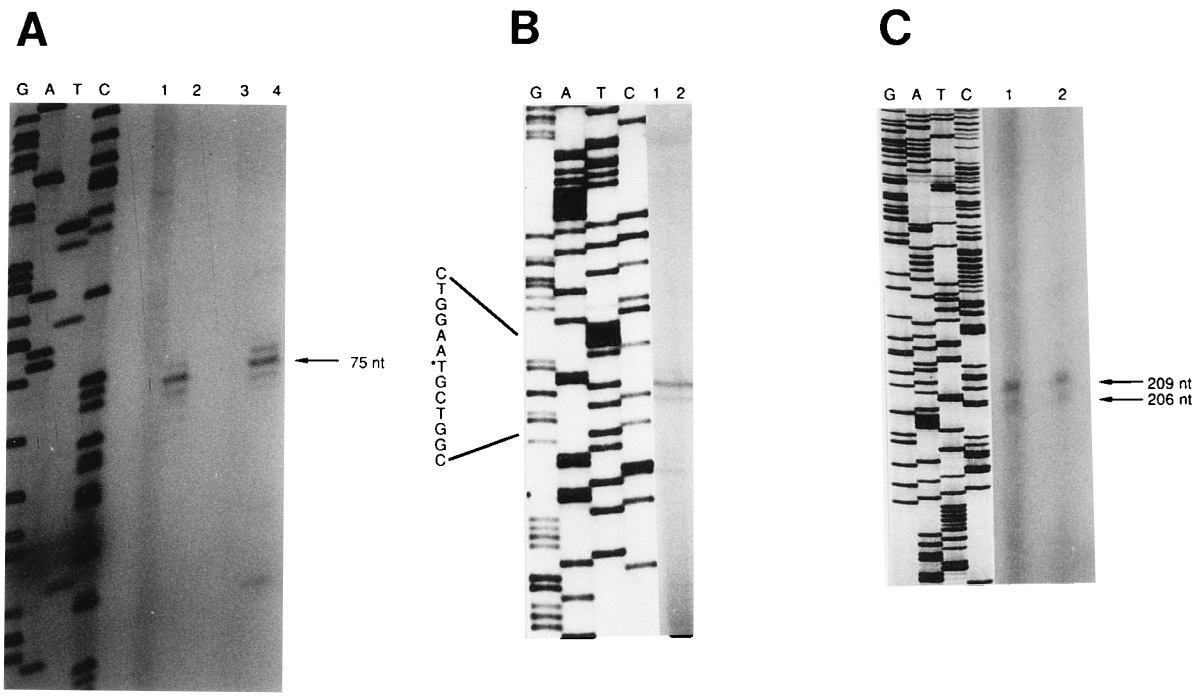

FIG. 1. (A) S1 nuclease protection of the *ilvD* transcript. A *Nar*I-*Xho*I fragment was labeled at the *Nar*I 5′ end and hybridized to 50 μg of in vivo RNA. The S1 digestion was carried out for 30 min at 37°C with 300 U of S1 nuclease. Lanes G, A, T, and C correspond to sequencing reactions of a known DNA fragment. The S1-protected fragments are indicated. (B) Primer extension mapping of the *ilvD* transcripts. A 30-nt primer complementary to the *ilvD* transcript was used for cDNA synthesis. The same oligonucleotide was used for sequence ladder shown on the right. Lane 1, RNA was prepared from cells grown in a rich medium; lane 2, RNA was prepared from cells grown in a minimal medium. The sequence spanning the transcription start site is shown, and the transcription start site is marked by an asterisk. (C) S1 nuclease mapping of the 5′ terminus of the *ilvR* transcript. A 360-bp *Xho*I-*Sst*I fragment which was labeled at the *Xho*I 5′ end was hybridized to 50 μg of in vivo RNA. The S1 digestion conditions were as described for panel B. The sizes of the protected fragments are indicated.

electroporation. The resultant strains were cultured in M_2 medium, and β-galactosidase activity was measured as described by Miller (18).

RNA purification. *C. crescentus* cells were grown in liquid PYE medium at 32°C with constant aeration. Cultures (100 ml) were centrifuged at 20,000 × *g* for 10 min at 4°C. The pellet was washed with 10 ml of STE buffer (10 mM NaCl, 10mM Tris-HCl [pH 7.6], 1 mM EDTA) and suspended in 10 ml of the same buffer. The cell suspension was transferred to a 65°C water bath, an equal volume of hot phenol (65°C) was added, and the mixture was shaken for 10 min. After centrifugation at 20,000 × *g* for 15 min, the hot phenol extraction was repeated, and the upper phase was extracted with phenol-chloroform-isoamyl alcohol (25:24:1). Nucleic acids were precipitated by the addition of 2 volumes of cold ethanol and centrifugation at 20,000 × *g* for 10 min. The pellet was washed with 75% ethanol, dried briefly, and suspended in 100 μl of DNase I buffer (40 mM Tris-HCl [pH 8.0], 6 mM $MgCl_2$, 10 mM NaCl) containing 1 U of RNase-free DNase I. After incubation at 37°C for 1 h and phenol-chloroform extraction, the RNA was precipitated by the addition of 2 volumes of ethanol. Following centrifugation, the pellet was washed with 75% ethanol, air dried, and suspended in distilled water. The RNA concentration was determined by optical density measurements at 260 nm.

Primer extension analysis. Total cellular RNA (50 μg) and 10^5 cpm of primer were coprecipitated with 2 volumes of ethanol and suspended in 30 μl of S1 hybridization buffer (25). The solution was incubated at 90°C for 3 min, then at 65°C for 5 min, and finally at 33°C overnight. After the RNA-primer annealing, the nucleic acids were precipitated, washed with 75% ethanol, and suspended in 25 μl of extension buffer (50 mM Tris-HCl [pH 8.3], 100 mM KCl, 10 mM $MgCl_2$, 10 mM dithiothreitol, 0.5 mM dNTPs, and 5 U of avian myoblastosis virus reverse transcriptase). After 1 h at 42°C, the reaction was stopped by the addition of 2 μl of 0.25 M EDTA and extracted with phenol-chloroform, and the nucleic acids were precipitated by the addition of 2 volumes of cold ethanol. After centrifugation, the pellet was washed with 75% ethanol, air dried, and suspended in loading buffer. The reaction products were fractionated on an 8% denaturing polyacrylamide gel and visualized by autoradiography.

RESULTS

Determination of the 5 terminus of the *ilvD* and the *ilvR* transcripts. The *ilvR* and *ilvD* genes of *C. crescentus* are divergently transcribed from a common regulatory region, and the *ilvR* gene product exhibits homology to the LysR family of transcriptional activators (15). The members of this family of transcriptional regulators often are activators of upstream genes that are divergently transcribed (12). Similarly, the IlvR protein was shown to be a transcriptional activator of the upstream and divergently transcribed *ilvD* gene (15). To map the transcription initiation site of the *ilvD* transcript, a *Nar*I-*Xho*I fragment containing the promoter regions for both the *ilvD* and *ilvR* genes was labeled at the *Nar*I 5′ end and hybridized to total cellular RNA. After nuclease S1 digestion, the DNA fragments protected from S1 nuclease were analyzed on a denaturing polyacrylamide gel. One major band of about 75 nucleotides (nt) was detected (Fig. 1A). The 3′ end of this fragment corresponds to an adenine residue 23 nt upstream from *ilvD* translation start codon. The transcription initiation site at this position was confirmed by a primer extension experiment. Primer BE-59 is complementary to the coding strand at a position 58 to 88 nt downstream from ATG start codon. This primer was used in both the primer extension assay and a set of sequencing reactions to obtain a precise transcription start point. This experiment yielded a 111-nt cDNA (Fig. 1B). The size of this signal placed the apparent transcription start site at the same adenine residue indicated by the S1 nuclease mapping experiment (Fig. 1A).

The transcription start site for the *ilvR* gene was determined by an S1 nuclease protection assay. A 360-bp fragment labeled at the 5′ end of the strand complementary to the *ilvR* RNA was isolated and hybridized with total cell RNA from *C. crescentus* cells with or without a plasmid harboring the *ilvR* gene. Two DNA probes of 209 and 206 nt were protected from nuclease S1 digestion (Fig. 1C). The sizes of the protected fragments

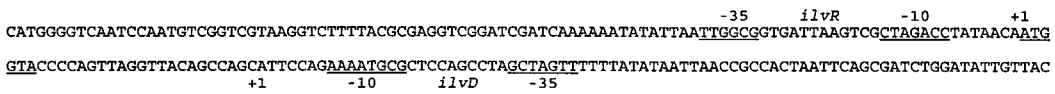

FIG. 2. DNA sequence of the *ilvD-ilvR* promoter region. The transcription start sites are indicated by +1. The proposed −10 and −35 regions are underlined.

were calculated on the basis of the sizes of the comigrating reference sequencing reactions. Fragments corresponding to these two start positions were present in additional experiments with different DNA probes and independent RNA isolations (data not shown). The 5′ end of the larger protected fragment (209 nt) coincides with the adenine residue of the ATG translation start codon of the *ilvR* gene. The start site of the smaller transcript (206 nt) corresponds to a guanine residue immediately downstream from the ATG codon. It is not known whether the smaller protected fragment represents an independent transcript or a degradation product of the larger transcript. The assignment of the ATG start codon to this position was based on the lack of any other potential start codon and the amino acid sequence homology of the *ilvR* protein with other LysR-type proteins (15). A summary of the promoter regions for *ilvD* and *ilvR* is shown in Fig. 2.

Determination of the 5 terminus of the *cysC* mRNA. The *cysC* gene is required for cysteine biosynthesis in *C. crescentus* (2). The cloned gene was identified by complementation of a cysteine auxotroph (26), and the DNA sequence was determined for part of the gene corresponding to the promoter region and the 5′ end of the coding region. The open reading frame corresponding to the *csyC* coding region was identified by codon preference analysis. The transcription initiation site of the *cysC* gene was defined by primer extension analysis using a 30-mer oligonucleotide specific for the *cysC* coding region. The 3′ end of this primer was 87 nt downstream from the *cysC* start codon. This experiment revealed a 193-nt fragment representing an extension product terminating at a position 76 nt upstream from the ATG codon (Fig. 3A).

Determination of the 5 terminus of the *fdxA* mRNA. The *fdxA* gene codes for the *C. crescentus* ferredoxin I (32). The promoter for the *fdxA* gene was defined by primer extension analysis using a 30-base oligonucleotide 77 nt downstream from the ATG translation initiation site. One major band, representing an extension product terminating 117 nt upstream from the ATG start codon, was observed on polyacrylamide gels (Fig. 3B).

Determination of the 5 terminus of the *pleC* mRNA. The *pleC* gene codes for a histidine protein kinase critical for polar development in *C. crescentus* (33). The transcriptional start site for *pleC* was determined by primer extension using a 30-base oligonucleotide 66 bp downstream of the ATG translational start codon of the *pleC* gene. As shown in Fig. 4, two major bands three bases apart were observed. The stronger, top band was designated the transcriptional start site and was 55 bases upstream from the ATG codon of *pleC*. The transcriptional start site was further confirmed by S1 mapping (data not shown).

Comparison of the promoter sequences for biosynthetic and housekeeping genes. There are a number of other constitutively expressed genes of *C. crescentus* for which the transcription initiation sites have been determined. An alignment of the known promoter sequences with the promoter sequences of the *ilvD*, *ilvR*, *cysC*, *fdxA*, and *pleC* genes allowed us to propose a consensus sequence for these *C. crescentus* promoters (Fig. 5). The resulting consensus is similar to that of *E. coli* in the −35 region but differs in the −10 region.

Promoter sequence mutagenesis. To demonstrate that the nucleotide sequences in the proposed promoter consensus sequence play a role in determining the efficiency of these promoters in vivo, oligonucleotide-directed mutagenesis was used to make nucleotide substitutions in the *ilvBN* promoter sequence (29). Mutations which make the *ilvBN* promoter sequence either more similar to or divergent from the proposed consensus promoter sequence were designed (Table 1). Double base substitutions were introduced at the flanking C nucleotides, the central nucleotides in the −10 region, and the highly conserved nucleotides in the −35 region. DNA fragments containing the wild-type *ilvBN* promoter and the mutated promoters were placed upstream from a promoterless *cat* gene on plasmid pRK2L1 and introduced into *C. crescentus*. The effect of base substitutions on promoter activity was ex-

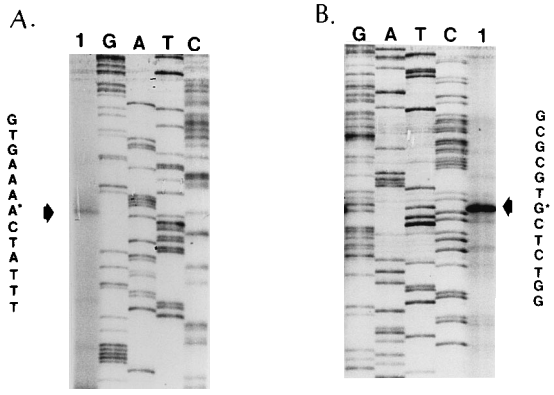

FIG. 3. Determination of the 5′ ends of mRNA from the *cysC* and *fdxA* genes by primer extension analysis. ^{32}P-radiolabeled primers complementary to the 5′ ends of *cysC* (A) and *fdxA* (B) were hybridized to 50 μg of total cell RNA. Primer extension products were analyzed on a 5% sequencing gel. The sequencing ladder (lanes G, A, T, and C) represents products of the sequencing reactions obtained by using the same oligonucleotides as used for primer extensions. The sequences shown are antisense strands. The asterisks show the bases representing the start of the transcription.

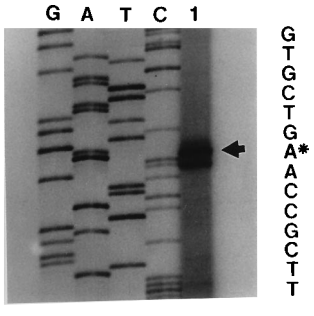

FIG. 4. Determination of the 5′ end of the *pleC* mRNA. A primer complementary to the *pleC* mRNA at a position 66 to 95 bases downstream from the ATG was labeled with [γ-^{32}P]ATP by using T4 polynucleotide kinase and hybridized to 100 μg of total *C. crescentus* mRNA. The primer was extended by reverse transcriptase, and the extended product was analyzed on a 5% sequencing gel adjacent to a sequencing ladder generated from the same primer. The sequences shown are antisense strands. The asterisk and arrow represent the start site for *pleC* mRNA.

C. crescentus

GENE	-35	-10	+1
ilvB	AAAGAGCGCCTTGACGCCCCATCTT	ACGCATGCCATGACCCTT	TCCG
ilvR	AATATATTAATTGGCGGTGATTAAG	TCGCTAGACCTATA	ACAA
ilvD	AATATATTTTTTGATCGATCCGACC	TCGCGTAAAAGAC	CTTA
fdxA	CGGAGGCGGCATGGCGCCGCCGGCC	ACCGCCCCAGAAATGCG	CGTG
cysC	CGGCGGCGCCATGCCGCCTCCGCGA	CAATCGGTTACCTCCGTG	AAAA
trpF	AACGCGCCGCTTGAGCGACTCGCCA	AGTCGGCCAAATCTTCT	GG
dnaA	AGTTTTCCGTTTGACCGGCCCCCTC	CGCTGGCTAGTTTAAGG	GTC
dnaK	CGAAGCGGGGTTGACGGGCTCGTCA	AAACTTCGCACAACGCAT	TCCG
rrnA	CGCGAGGGGGTTGCCAGTCGCCGGC	TGGCCGACTAGATCACCG	CC
hemE	ACGTCCGGGCTTGAGCCGGAGCCCG	CGCCGGGCTTTAACCCGG	TCCA
rsaA	AATTCGGCTATTGTCGACGCATGAC	GTTTGCTCTATAGCCATC	GCT
pleC	CGCTTCGGCGGTGACGCGAACGGCG	CGGTACGTTAGATTCACG	ACT
Consensus	**TTGaCgS** **10-14**	**GCtANAWC** **5-7**	**N**
E. coli	TTGACa 16-19	TAtAaT 5-9	A

FIG. 5. Compilation of sequences of the promoter regions of housekeeping and biosynthetic genes. The −10 and −35 promoter regions, which are separated by 13 to 16 nt, are indicated. The consensus promoter sequence is shown at the bottom. The promoter sequences (and references) are as follows: *ilvD*, *ilvR*, *cysC*, *fdxA*, and *pleC* (this study); *ilvBN* (29); *trpF* (24); *dnaA* (36); *dnaK* (10); *rrnA* (1); *hemE* (17); and *rsaA* (8). The proposed consensus sequence was based on the frequency of the occurrence of each nucleotide at a particular position. A base was considered part of the consensus if it was the primary one used at that position. Nucleotides that appeared in fewer than 70% of the promoter sequences are shown in lowercase letters. Nucleotides were assigned to each position in the consensus sequence on the basis of the frequency of occurrence of each nucleotide in the −10 and −35 promoter regions of the promoters. All promoter sequences are based on S1 nuclease mapping and primer extension analysis using in vivo RNA. An S in the consensus sequence indicates C or G, a W indicates A or T, and an N indicates any nucleotide. Lowercase letters indicate that the designated base was present in 7 of the 12 promoters examined; uppercase letters indicate that the designated base was present more than seven times. The *E. coli* promoter consensus sequence is also shown (11).

amined by assaying for CAT activity in cell extracts prepared from all clones carrying the hybrid plasmids (Table 1). A strain carrying plasmid pJM122, in which the internal AT dinucleotide of the −10 region was replaced with a CG dinucleotide, had activity threefold lower than that of a strain carrying the wild-type promoter sequence in pJM120. Nucleotide substitutions at the flanking C residues in pJM124 resulted in approximately a fourfold reduction in promoter activity. The replacement of a T at position −11 and deletion of an A at position −12, in pJM121, resulted in a 10-fold decrease in CAT activity. Similarly, base substitutions at the −35 region, which replace the highly conserved TG nucleotides with CC, resulted in a 10-fold decrease in the CAT activity. On the other hand, replacing two internal C and G residues of the −10 region with T and A, respectively, resulted in a 10-fold increase in promoter activity. These two substitutions change nonconsensus nucleotides to consensus nucleotides at positions −10 and −13.

TABLE 1. Activities of the mutant *ilvBN* promoters[a]

Gene or plasmid	DNA sequence at[b]: −35	DNA sequence at[b]: −10	CAT activity[c] (mean ± SD)
Consensus	TTG	GCtANAWC	
ilvBN (wild type)	TTG	GCCATGAC	
pJM120			100
pJM122		CG	36 ± 8
pJM121		▲G	10 ± 1
pJM125		T A	970 ± 140
pJM124		T T	23 ± 8
pLM126	CC		13 ± 20

[a] The mutated promoter fragments were fused to the *cat* gene, and CAT activity was measured in cultures of strains carrying different *cat* fusions.

[b] Sequence alterations resulting from in vitro site-specific mutagenesis of the *ilvBN* promoter region are shown. Substitutions from the wild type are shown at their appropriate positions. Dots indicate identity to the base at that position in the wild-type *ilvBN* promoter.

[c] Relative activity compared with that of pJM120, which has a DNA sequence identical to that of the wild-type *ilvBN* promoter.

TABLE 2. Relative activities of the mutant *pleC* promoters

Promoter or plasmid	−10 region DNA sequence[a]	β-Galactosidase activity[b] (mean ± SD)
Consensus	GCtANAWC	
pleC promoter	GTAC GTTAGATT	
pLEC2910		100 ± 3
pLEC2902	C	189 ± 3
pLEC2907	C	210 ± 4
pLEC2903	C C	155 ± 0.7
pLEC2908	C C	154 ± 1
pLEC2906	C C	239 ± 2
pLEC2904	C C C	156 ± 4
pLEC2901	C C C	34 ± 3
pLEC2912	C G C	43 ± 1
pLEC2911	C G G C	0 ± 7

[a] The DNA sequence of the wild-type *pleC* promoter at the −10 region was compared with the consensus shown in Fig. 5. Substitutions from the wild type are shown at their appropriate positions. Dots indicate identity to the base at that position in the wild-type *pleC* promoter.

[b] Relative activity compared with that of pLEC2910, which has a DNA sequence identical to that of the wild-type *pleC* promoter.

Further support for the consensus sequence was obtained by mutagenesis of the *pleC* promoter. The *pleC* promoter is a relatively weak promoter that lacks the conserved C residues at either end of the −10 region. Various changes were introduced into the −10 region or the spacer region by using a degenerate primer complementary to the promoter region (Table 2). When either or both of the terminal C residues was present in the −10 region, promoter activity increased approximately twofold. However, changing the conserved A located three residues from the end of the −10 region caused a 60% reduction in promoter activity. Furthermore, changing both of the conserved A residues to G's eliminated promoter activity entirely. Changing a T residue to a C in the spacer region had no effect on *pleC* promoter activity. In combination, these data indicate that the proposed consensus sequence is valid and suggest that its base composition is close to an optimal promoter sequence for constitutively expressed genes in *C. crescentus.*

DISCUSSION

Using high-resolution S1 nuclease protection and primer extension analyses, we have located the transcription initiation sites of the *ilvR*, *ilvD*, *cysC*, *fdxA*, and *pleC* genes of *C. crescentus*. The 5′ regulatory regions of these genes were compared with those of other previously published biosynthetic or housekeeping genes. These promoters were aligned on the basis of their obvious similarity at sequences similar to the −35 region of the *E. coli* σ^{70} consensus promoter sequence (11). The level of similarity at the 5′ regulatory sequences of these promoters suggested that the transcription was mediated via a common sigma factor. A comparison of these sequences allowed us to propose a consensus promoter sequence for biosynthetic and housekeeping genes in *C. crescentus*. The −35 region consensus promoter sequence (TTGACGS) is similar to the corresponding region of the promoters recognized by the *E. coli* RNA polymerase. However, the −10 region (GCTANAWC) is significantly different from the corresponding regions of σ^{70}

promoters. Also, the spacer region is slightly shorter in *C. crescentus* than in *E. coli*, and the base at +1 is not conserved in *C. crescentus*. In accordance with these differences, none of the *C. crescentus* gene promoters tested are expressed in *E. coli* (28, 31, 34). This finding suggests the involvement of a sigma factor with a specificity different from that of the σ^{70} of *E. coli*. Alternatively, an additional regulatory feature(s) may be required for the expression of some of these genes.

Prior to this analysis, the inability of *E. coli* to express genes from *C. crescentus* was rather puzzling since the *E. coli* transcriptional machinery recognizes transcriptional signals from a wide range of microorganisms as divergent as *Bacillus subtilis* (6) and *Saccharomyces cerevisiae* (23). In another report (16), we show that when the *C. crescentus rpoD* gene is fused to the *E. coli lacZ* promoter, it is expressed in *E. coli*. From sequence comparisons and other data, we concluded that *rpoD* encodes the principal sigma factor of *C. crescentus*. Thus, the barrier to *C. crescentus* gene expression in *E. coli* appears to be at the level of transcription (16). Furthermore, in the case of the *pleC* and *fdxA* promoters, which are not normally expressed in *E. coli* (31), we have demonstrated that expression occurs in *E. coli* strains which express the *C. crescentus rpoD* gene (16). Therefore, our results indicate that the inability of *E. coli* to transcribe the *C. crescentus* biosynthetic and housekeeping gene promoters is due to differences in the sigma factor subunits.

The results of in vitro mutagenesis of the *ilvBN* and *pleC* promoters indicated that (i) the proposed −10 and −35 regions of this promoter are involved in *C. crescentus* RNA polymerase interactions; (ii) base substitutions that decrease similarity to the consensus promoter sequence decrease promoter activity by 30 to 100%; and (iii) mutations that increase similarity to the consensus promoter sequence increase promoter activity. The sequence of the promoter for the *rsaA* gene, which codes for the paracrystallin surface array protein, matches the consensus sequence most closely. Since this gene encodes one of the most abundant proteins in *C. crescentus* (7), this concordance provides additional evidence that the proposed consensus is close to the optimal sequence for promoters recognized by the principal sigma factor of *C. crescentus.*

ACKNOWLEDGMENTS

We thank Betty Branham for help with preparation of the manuscript.

This work was supported by grants GM34765 and GM50547 from the National Institutes of Health.

REFERENCES

1. **Amemiya, K.** 1989. Conserved sequence elements upstream and downstream from the transcription initiation site of the *Caulobacter crescentus rrnA* gene cluster. J. Mol. Biol. **210:**245–254.
2. **Barrett, J. T., R. H. Croft, D. M. Ferber, C. J. Gerardot, P. V. Schoenlein, and B. Ely.** 1982. Genetic mapping with Tn*5*-derived auxotrophs of *Caulobacter crescentus*. J. Bacteriol. **149:**888–898.
3. **Brun, Y. V., and L. Shapiro.** 1992. A temporally controlled sigma-factor is required for polar morphogenesis and normal cell division in *Caulobacter*. Genes Dev. **6:**2395–2408.
4. **Dingwall, A., J. W. Gober, and L. Shapiro.** 1990. Identification of a *Caulobacter* basal body structural gene and a *cis*-acting site required for activation of transcription. J. Bacteriol. **172:**6066–6076.
5. **Dingwall, A., W. Y. Zhuang, K. Quon, and L. Shapiro.** 1992. Expression of an early gene in the flagellar regulatory hierarchy is sensitive to an interruption in DNA replication. J. Bacteriol. **174:**1760–1768.
6. **Ehrlich, C. D.** 1978. DNA cloning in *Bacillus subtilis*. Proc. Natl. Acad. Sci. USA **75:**1433–1436.
7. **Fisher, J. A., J. Smit, and N. Agabian.** 1988. Transcriptional analysis of the major surface array gene of *Caulobacter crescentus*. J. Bacteriol. **170:**4706–4713.
8. **Gilchrist, A., J. Fisher, and J. Smit.** 1992. Nucleotide sequence analysis of the gene encoding the *Caulobacter crescentus* paracrystallin surface layer protein. Can. J. Microbiol. **38:**193–202.
9. **Gober, J. W., and L. Shapiro.** 1992. A developmentally regulated *Caulobacter* flagellar promoter is activated by 3′ enhancer and IHF binding elements. Mol. Biol. Cell **3:**913–926.
10. **Gomes, S. L., J. W. Gober, and L. Shapiro.** 1990. Expression of the *Caulobacter* heat shock gene *dnaK* is developmentally controlled during growth at normal temperatures. J. Bacteriol. **172:**3051–3059.
11. **Hawley, D. K., and W. R. McClure.** 1983. Compilation and analysis of *Escherichia coli* promoter DNA sequences. Nucleic Acids Res. **11:**2237–2255.
12. **Henikoff, S., G. W. Haughn, J. M. Calvo, and C. Wallace.** 1988. A large family of bacterial activator proteins. Proc. Natl. Acad. Sci. USA **85:**6602–6606.
13. **Johnson, R., and B. Ely.** 1977. Isolation and spontaneously-derived mutants from *Caulobacter crescentus*. Genetics **86:**25–32.
14. **Kunkel, T. A., J. D. Robert, and R. A. Zakour.** 1987. Rapid and efficient site-specific mutagenesis without phenotypic selection. Methods Enzymol. **154:**367–382.
15. **Malakooti, J., and B. Ely.** 1994. Identification and characterization of the *ilvR* gene encoding a LysR-type regulator of *Caulobacter crescentus*. J. Bacteriol. **176:**1275–1281.
16. **Malakooti, J., and B. Ely.** The principal sigma subunit of the *Caulobacter crescentus* RNA polymerase. Submitted for publication.
17. **Marczynski, G., K. Lentine, and L. Shapiro.** Personal communication.
18. **Miller, J.** 1972. Experiments in molecular genetics. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
19. **Mullin, D., S. Minnich, L. 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.
20. **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.
21. **Newton, A., and N. Ohta.** 1990. Regulation of the cell division cycle and differentiation in bacteria. Annu. Rev. Microbiol. **44:**689–719.
22. **Ninfa, A. J., D. A. Mullin, G. Ramakrishnan, and A. Newton.** 1989. *Escherichia coli* σ^{54} RNA polymerase recognizes *Caulobacter crescentus flbG* and *flaN* flagellar gene promoters in vitro. J. Bacteriol. **171:**383–391.
23. **Ratzkin, B., and J. Carbon.** 1977. Functional expression of cloned yeast DNA in *Escherichia coli*. Proc. Natl. Acad. Sci. USA **74:**487–491.
24. **Ross, C. M., and M. E. Winkler.** 1988. Structure of the *Caulobacter crescentus trpFBA* operon. J. Bacteriol. **170:**757–768.
25. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
26. **Schoenlein, P., L. Gallman, and B. Ely.** 1988. Use of transmissible plasmids as cloning vectors in *Caulobacter crescentus*. Gene **70:**321–329.
27. **Shapiro, L.** 1993. Protein localization and asymmetry in the bacterial cell. Cell **73:**841–855.
28. **Tarleton, J., and B. Ely.** 1991. Isolation and characterization of *ilvA*, *ilvBN*, and *ilvD* mutants of *Caulobacter crescentus*. J. Bacteriol. **173:**1259–1267.
29. **Tarleton, J., J. Malakooti, and B. Ely.** 1994. Regulation of *Caulobacter crescentus ilvBN* gene expression. J. Bacteriol. **176:**3765–3774.
30. **Van Way, S. M., A. Newton, A. H. Mullin, and D. A. Mullin.** 1993. Identification of the promoter and a negative regulatory element, *ftr4*, that is needed for cell cycle timing of *fliF* operon expression in *Caulobacter crescentus*. J. Bacteriol. **175:**367–376.
31. **Wang, S. P., and B. Ely.** Unpublished data.
32. **Wang, S. P., P. J. Kang, Y. P. Chen, and B. Ely.** 1995. Synthesis of the *Caulobacter* ferredoxin protein, FdxA, is cell cycle controlled. J. Bacteriol. **177:**2901–2907.
33. **Wang, S. P., P. L. Sharma, P. V. Schoenlein, and B. Ely.** 1993. A histidine protein kinase gene family is implicated in polar organelle development in *Caulobacter crescentus*. Proc. Natl. Acad. Sci. USA **90:**630–634.
34. **Winkler, M. E., P. V. Schoenlein, C. M. Ross, J. T. Barrett, and B. Ely.** 1984. Genetic and physical analyses of *Caulobacter crescentus trp* genes. J. Bacteriol. **160:**279–287.
35. **Yu, J., and L. Shapiro.** 1992. Early *Caulobacter crescentus* genes *fliL* and *fliM* are required for flagellar gene expression and normal cell division. J. Bacteriol. **174:**3327–3338.
36. **Zweiger, G., and L. Shapiro.** 1994. Expression of *Caulobacter dnaA* as a function of the cell cycle. J. Bacteriol. **176:**401–408.