JOURNAL OF BACTERIOLOGY, Nov. 1997, p. 6824–6830 Vol. 179, No. 21
0021-9193/97/$04.00+0

Expression of a Streptomycete Leaderless mRNA Encoding Chloramphenicol Acetyltransferase in *Escherichia coli*

CHI-JU WU AND GARY R. JANSSEN*

Department of Microbiology, Miami University, Oxford, Ohio 45056

Received 12 May 1997/Accepted 19 August 1997

The chloramphenicol acetyltransferase (*cat*) gene from *Streptomyces acrimycini* encodes a leaderless mRNA. Expression of the *cat* coding sequence as a leaderless mRNA from a modified *lac* promoter resulted in chloramphenicol resistance in *Escherichia coli*. Transcript mapping with nuclease S1 confirmed that the 5′ end of the *cat* message initiated at the A of the AUG translational start codon. Site-directed mutagenesis of the *lac* promoter or the *cat* start codon abolished chloramphenicol resistance, indicating that *E. coli* initiated translation at the 5′ terminal AUG of the *cat* leaderless mRNA. Addition of 5′-AUGC-3′ to the 5′ end of the *cat* mRNA resulted in translation occurring also from the reading frame defined by the added AUG triplet, suggesting that a 5′-terminal start codon is an important recognition feature for initiation and establishing reading frame during translation of leaderless mRNA. Addition of an untranslated leader and Shine-Dalgarno sequence to the *cat* coding sequence increased *cat* expression in a *cat*:*lacZ* fusion; however, the level of expression was significantly lower than when a fragment of the bacteriophage lambda *c*I gene, also encoding a leaderless mRNA, was fused to *lacZ*. These results indicate that in the absence of an untranslated leader and Shine-Dalgarno sequence, the streptomycete *cat* mRNA is translated by *E. coli*; however, the *cat* translation signals, or other features of the *cat* mRNA, provide for only a low level of expression in *E. coli*.

The translation frequency for procaryotic mRNA containing a 5′ untranslated leader region is determined, in part, by the extent of complementarity between a Shine-Dalgarno (SD) sequence within the leader region and the anti-Shine-Dalgarno (ASD) sequence located near the 3′ end of the 16S rRNA (8, 9, 28, 30). Small subunit rRNAs from procaryotic ribosomes contain the conserved ASD region, and most procaryotic mRNAs contain a readily identifiable SD sequence. While the contributions of the SD-ASD interaction to translation are likely to be mechanistically similar among all procaryotes, other features of the translation initiation region have been proposed to also contribute to translation levels (6, 18, 21, 29). The identification and analysis of mRNA features that influence translation levels are important for understanding the translation initiation process and for considerations of optimizing expression levels when genes are expressed within heterologous hosts.

Although untranslated leader regions and SD sequences are found at the 5′ ends of most procaryotic mRNAs, some genes encode leaderless mRNA whereby transcription and translation initiate at the same position. While genes that encode leaderless mRNA are relatively rare, more than 30 have been identified (11, 23, 32) since the *Escherichia coli* phage λ *c*I repressor was first reported in 1976 (22). Observations of leaderless mRNA (11) in *Bacteria*, *Archaea*, *Eucarya*, and eucaryotic organelles suggest that sequence and/or structural information contained within the coding sequence are sufficient to signal the translational start site and reading frame in these diverse biological systems. The widespread occurrence of leaderless mRNAs suggests that translation of leaderless mRNAs might represent a fundamental capability of all translation systems. Although the *E. coli* chromosome is not known to contain any genes that encode leaderless mRNA, the observation that it translates leaderless *c*I (22), *tetR* (15), gene *V* (4), and unleadered *vph* (31) mRNAs indicates that *E. coli* is able to translate mRNAs lacking upstream leader sequences.

The features that determine translation levels from leaderless mRNA are poorly characterized. In the absence of an untranslated leader region, information specifying translational efficiency must be contained within the coding region primary sequence or mRNA structure. Translation of leaderless mRNA might require interactions, or additional factors, unique to the translational machinery of the producing organism, thereby limiting expression to a narrow host range; alternatively, leaderless mRNA might contain translation signals near the 5′ terminus, including a terminal start codon, that are highly conserved and allow for expression of a variety of leaderless mRNAs within any host system. Characterization of leaderless mRNA expression in heterologous translation systems will help distinguish between these possibilities.

Chloramphenicol acetyltransferase (*cat*) genes are widespread among bacterial genera and represent the most common mechanism of microbial resistance to chloramphenicol (26, 27). Among the many known *cat* genes, only the *Streptomyces acrimycini cat* has been shown to encode a leaderless mRNA (20). In this report, we describe fusion of the *S. acrimycini cat* coding sequence to an *E. coli lac* promoter modified such that transcription initiated at the translational start codon. Chloramphenicol resistance resulting from *cat* expression indicated that *E. coli* ribosomes translated the leaderless *cat* mRNA; however, the observed low levels of resistance, and low expression from a *cat*:*lacZ* fusion, suggested that translation signals present on the *cat* mRNA did not function efficiently in *E. coli*.

MATERIALS AND METHODS

Strains, plasmids, and culture conditions. *E. coli* DH5α [F^- ϕ80d*lacZ*ΔM15 Δ(*lacZYA-argF*)*U169 deoR recA1 endA1 hsdR17*(r_K^- m_K^+)*supE44 thi-1 gyrA96 relA1*; Bethesda Research Laboratories] was used for all standard cloning procedures; *E. coli* CJ236 [*dut ung thi-1 relA1*/pCJ105(Cm^r) (16)] was used for site-directed mutagenesis, and *E. coli* RFS859 (F^- *thr-1 araC859 leuB6* Δ*lac74*

* Corresponding author. Mailing address: Department of Microbiology, Miami University, Oxford, OH 45056. Phone: (513) 529-1694. Fax: (513) 529-2431. E-mail: grjanssen@miavx1.muohio.edu.

tsx-274 λ⁻ *gyrA111 recA11 relA1 thi-1* [24]) was used as the host for measurements of β-galactosidase (*lacZ*) expression.

Plasmid pIU459 (*Plac*:*cat*) is a derivative of pTL61T (17) that contains the *cat* coding sequence from pIJ879 (20) and expressed from the *Plac* promoter (31). The unique *Nco*I site at the *cat* transcriptional/translational start site of pIU459 was end filled and religated, resulting in a 4-bp insertion (5′-ATGC-3′) at the *Plac* transcriptional start site to yield pIU460 (*Plac*:*4bp*:*cat*). The mutant *Plac** promoter (31) was fused to the *cat* sequence in pIU459 to generate pIU1010 (*Plac**:*cat*). A 1.8-kb *Eco*RI-*Sst*I fragment from pIU459 was cloned into pTZ18U (16) to yield pIU1015 (*Plac*:*cat*). Phagemid DNA derived from pIU1015 was used for site-directed mutagenesis of the *cat* start codon from ATG to ACC in pIU1018 [*Plac*(*ACC*):*cat*]. Plasmids pIU1037, pIU1038, pIU1039, and pIU1040 correspond to translational fusions of *cat* codons 1 to 16 from pIU459 or pIU460 to a *lacZ* reporter gene and are described in Fig. 5. Sixteen codons of the bacteriophage lambda *c*I gene are translationally fused to *lacZ* and expressed from *Plac* in pIU1041 (*Plac*:*cI*:*lacZ* [10]). A modified untranslated leader sequence from the *lac* operon (LL) or the β-D-glucuronidase (*gus*) gene (GL) was cloned between the *Plac* transcriptional start and the *cat* translational start to yield pIU1042 (*Plac*:LL:*cat*:*lacZ*) or pIU1043 (*Plac*:GL:*cat*:*lacZ*), respectively.

Culture conditions for *E. coli* were described by Schottel et al. (25). Chloramphenicol resistance was assayed by inoculating 10 μl of bacterial suspension (diluted to an A_{600} of 0.001) to 2 ml of 0.5× L broth (1× L broth contains, per liter, 10 g of Bacto Tryptone, 5 g of Bacto yeast extract, 5 g of NaCl, and 1 g of glucose) containing various concentrations (0, 0.3, 0.6, 0.9, 1.2, and 3 μg/ml) of chloramphenicol and measuring the A_{600} of cultures after growth for 24 h at 30°C.

DNA manipulation, sequencing, and transformation. All restriction endonucleases and DNA-modifying enzymes were used according to the manufacturers' recommendations. Dideoxynucleotide sequencing was done by using Sequenase (United States Biochemical) as recommended by the manufacturer. Transformation of *E. coli* DH5α and selection of transformants were performed as described previously (12).

Mutagenesis of the *cat* start codon from ATG to ACC. The mutagenesis procedure was a modification of the Kunkel method (16) and used a kinase-treated DNA fragment, resulting from a PCR amplification, instead of an oligonucleotide as the primer for in vitro DNA synthesis. PCR amplification using oligonucleotides 5′-CGGGGCGTCGGTGGACACATTATAC-3′ (underlined primer nucleotides identify the changed positions) and 5′-CGGGCAAGAGCA ACTCGGTCGCCGC-3′ as primers and pIU1015 as the template yielded a 1,942-bp *Plac*:*cat* fragment that was used to change the *cat* putative start codon from ATG to ACC.

Single-stranded phagemid DNA used for mutagenesis was prepared from *E. coli* CJ236 (*dut ung*) transformed with pIU1015 as described by Kunkel et al. (16). After annealing of the mutagenic PCR-amplified fragment to the phagemid template, in vitro synthesis of the complementary strand, and ligation, the synthesis reaction product was transformed into *ung*⁺ cells (DH5α). Transformants selected on L agar-ampicillin were pooled, and plasmid DNA was isolated. The miniprepped plasmid DNA was digested with *Nco*I to enrich for the desired mutation (i.e., ATG→ACC eliminates the *Nco*I restriction site [CCATGG]) and transformed into *E. coli* DH5α. Transformants containing a plasmid with the ACC mutation were identified by colony hybridization (31) and confirmed by DNA sequencing.

RNA isolation. *E. coli* DH5α cells containing pIU459, pIU460, pIU1010, or pIU1018 were harvested by centrifugation at an A_{600} not exceeding 0.7, and total RNA was isolated as described previously (7, 31).

Determination of the 5′ end of *cat* mRNA by S1 mapping. DNA probes (~250-bp *Eco*RI-*Apa*LI fragment) used in S1 mapping of *cat* mRNA start site were prepared by PCR as follows: Approximately 120 pmol of the oligonucleotide 5′-TGCACGGGACGCGCCGGCGG-3′ was phosphorylated by using 1.5 μl of [γ-^{32}P]ATP (150 mCi/ml, 6,000 Ci/mmol) and T4 polynucleotide kinase and purified by passage through a Sephadex G-50 NICK column (Pharmacia). Five microliters (approximately 0.6 to 0.7 pmol) of kinase-treated oligonucleotide eluted from the Sephadex G-50 NICK column was used as the primer in a PCR amplification; the second oligonucleotide primer (5′-CGGAATTCTGGCACG ACAGGTTTCC-3′) was not radiolabeled. Plasmids pIU459, pIU460, pIU1010, and pIU1018 were used as templates for PCR amplifications. One-tenth of the PCR product was used as the probe in each hybridization reaction.

^{32}P-labeled probe DNA was precipitated with 200 μg of total in vivo-synthesized RNA. The precipitate was washed with ethanol, dried, and then dissolved in 30 μl of hybridization solution [0.4 M NaCl, 40 mM piperazine-*N*,*N*′-bis(2-ethanesulfonic acid) (PIPES; pH 6.4), 1 mM EDTA, 80% (vol/vol) formamide] (5). After incubation at 85°C for 15 min, the temperature was lowered to 69.5°C and hybridization continued for 6 h. At the end of the hybridization period, samples were digested with nuclease S1 as described previously (31). Nuclease S1-resistant hybrids were analyzed by electrophoretic comparison to a DNA sequencing ladder prepared by dideoxynucleotide sequencing with the same primer (5′-TGCACGGGACGCGCCGGCGG-3′; complementary to nucleotide positions 75 to 95 of the *cat* coding sequence) used to prepare the ^{32}P-labeled probe DNA.

Translational fusions to *lacZ*. PCR amplification was used to introduce a *Sal*I restriction site after the 16th codon of the *cat* and *c*I coding regions. PCR-generated *Eco*RI-*Sal*I fragments, containing the *Plac* promoter and 16 codons of *cat* or *c*I, were ligated to plasmids containing a *Sal*I site at the fifth codon of a *lacZ* reporter gene (10). The *cat* sequence (codons 1 to 16) fused to *lacZ* is shown in Fig. 1. The *c*I DNA sequence fused to *lacZ*, beginning with the initiation codon (underlined), is 5′-ATGAGCACAAAAAAGAAACCATTAACACAAGAGCA GCTTGAGGACGCA-3′.

```
(1)
ATG GAC GCC CCG ATC CCG ACC CCG GCC CCG ATC GAC CTC GAC
Met Asp Ala Pro Ile Pro Thr Pro Ala Pro Ile Asp Leu Asp
1
      (16)
ACG TGG CCG CGC CGG CAG CAC TTC GAC CAC TAC CGC CGG CGC
Thr Trp Pro Arg Arg Gln His Phe Asp His Tyr Arg Arg Arg
43
                                (35)      (37)      (39)
GTC CCG TGC ACG TAC GCG ATG ACG GTG GAG GTG GAC ---
Val Pro Cys Thr Tyr Ala Met Thr Val Glu Val Asp
85
```

FIG. 1. DNA and inferred amino acid sequences at the 5′ end of the *S. acrimycini cat* coding sequence (20). Numbers below the sequence indicate nucleotide positions in the *cat* coding sequence; numbers in parentheses above the sequence indicate the codon, beginning with the initiation codon. Additional downstream potential start codons are boldfaced and underlined.

Modified untranslated leader regions from the *lac* operon or the *gus* gene were added to *cat*:*lacZ* constructs by ligating a DNA fragment, containing the *Plac* promoter and the *lac* or *gus* leader sequence, to the *cat* translational start site of a *cat*:*lacZ* fusion. The DNA sequences of the modified *lac* and *gus* untranslated leaders, beginning with the 5′-terminal nucleotide and extending to the initiation codon, are 5′-AATTGTGAGCGGATAACAATTTCACACAGGAAACAGC C-3′ and 5′-ATCGATCTGTGTGCAACACAGAATTGGTTAACTAATCAG ATTAAAGGTTGACCAGTATTATTATCTTAATGAGGAGTCCCCC-3′, respectively; the presumed SD sequences in the *lac* and *gus* leaders are underlined.

β-Galactosidase assay. Cultures were diluted from stationary-phase growth to a starting density (A_{600}) of 0.0125 in 2XYT medium (per liter, 16 g of Bacto Tryptone, 10 g of Bacto yeast extract, and 10 g of NaCl [pH 7.4]) containing 200 μg of ampicillin per ml and 0.2 mM isopropylthio-β-D-galactopyranoside (IPTG; Sigma Chemical Co.), and incubation with shaking continued at 37°C until the A_{600} reached approximately 0.4 (five doublings). For each strain, triplicate assays were performed on each of triplicate cultures. The β-galactosidase assay was performed by the method of Miller (19).

RESULTS

Fusion of the *cat* coding sequence to an *E. coli Plac* promoter. The *S. acrimycini cat* gene was subcloned from pIJ879 (20). The site-directed mutagenesis method of Kunkel et al. (16) was used to create *Nco*I restriction sites at the *cat* translational start site and the transcriptional start site of a modified *lac* promoter (*Plac* [31]). The *cat* coding sequence was then ligated to the *Plac* promoter via common *Nco*I restriction sites. Creation of an *Nco*I site at the *cat* translational start site (i.e., CGATGG→CCATGG [the *cat* start codon is underlined]) did not alter the amino acid sequence encoded by the *cat* sequence. An additional construct was prepared by end filling the *Plac*:*cat* (pIU459) *Nco*I restriction site, resulting in the *Plac*:*4bp*:*cat* (pIU460) construct containing a 4-bp (5′-ATGC-3′) insertion at the *Plac*:*cat* transcription/translation start site.

Determination of transcriptional start sites by nuclease S1 mapping. To ensure that *cat* was transcribed as a leaderless mRNA when expressed from the modified *E. coli Plac* promoter, nuclease S1 was used to identify the transcriptional start site. RNA was isolated from *E. coli* cells containing plasmids with the *Plac*:*cat* (pIU459) and *Plac*:*4bp*:*cat* (pIU460) fusions. The transcriptional start sites in *Plac*:*cat* mapped to the A and U of the AUG translational start codon (Fig. 2, lane 2). When four nucleotides (5′-ATGC-3′) were inserted by end filling the *Nco*I restriction site, the transcriptional start site shifted to the A and U positions within the inserted nucleotides (Fig. 2, lane 5); initiation at the A position results in a 5′-terminal AUG potential start codon, corresponding to the first three nucleo-

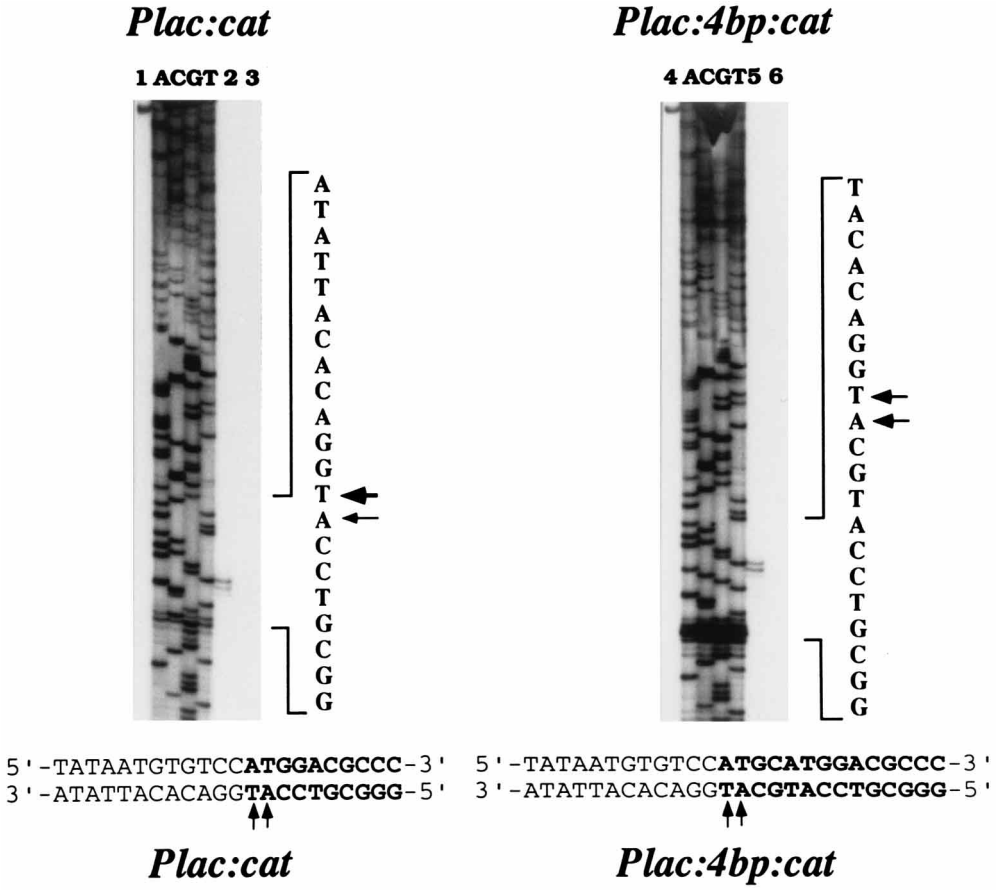

FIG. 2. Nuclease S1 mapping of *cat* mRNA initiated from *Plac*:*cat* (pIU459) and *Plac*:*4bp*:*cat* (pIU460). The PCR-amplified *Eco*RI-*Apa*LI fragment of *Plac*:*cat*, *Plac*:*4bp*:*cat*, uniquely labeled at the *Apa*LI 5′ end, was used as the probe (lanes 1 and 4) in hybridization reactions containing in vivo-synthesized RNA. Total RNA was isolated from *E. coli* DH5α cells lacking the *cat* gene (lanes 3 and 6) or containing *Plac*:*cat* (pIU459 [lane 2]) or *Plac*:*4bp*:*cat* (pIU460 [lane 5]). DNA sequencing ladders represent the template strand; arrows mark the transcriptional start sites identified by the migratory position of the nuclease S1-resistant fragments. The *cat* coding sequence, initiating with an ATG triplet, is boldfaced.

tides of the *Plac*:*4bp*:*cat* transcript which is out of frame with the *cat* coding sequence.

Although plasmids containing the *Plac*:*cat* and *Plac*:*4bp*:*cat* fusions had a transcriptional terminator upstream of the *Plac* promoter, it was possible that the *cat* mRNA observed (Fig. 2) resulted from processing of a larger vector-initiated transcript. To ensure that the observed *cat* leaderless mRNA resulted from initiation by *Plac*, we fused *cat* to a mutant *lac* promoter (*Plac** [31]). Fusion of *Plac** to *cat* resulted in loss of the *cat* mRNA (Fig. 3, lane 5), indicating that the leaderless *cat* mRNA observed in cells containing *Plac*:*cat* (Fig. 3, lane 2) initiated from the *Plac* promoter.

Expression of chloramphenicol resistance from the leaderless *cat* mRNA in *E. coli*. To determine whether the leaderless *cat* mRNA was translated to confer chloramphenicol resistance, growth of *E. coli* containing *Plac*:*cat*, *Plac*:*4bp*:*cat*, or *Plac**:*cat* was assessed in the presence of chloramphenicol. *E. coli* cells containing *Plac*:*cat* showed resistance to chloramphenicol, while cells containing *Plac*:*4bp*:*cat* and *Plac**:*cat* showed chloramphenicol sensitivity similar to that of the host cell control (Fig. 4). The observation that cells containing *Plac*:*4bp*:*cat* were sensitive to chloramphenicol indicated that the downstream in-frame AUG did not support sufficient translation to confer resistance at the concentrations tested. The lack of resistance in cells containing *Plac**:*cat* was consistent with the absence of *cat* mRNA (Fig. 3, lane 5) and suggested strongly that the leaderless *cat* mRNA detected by S1 mapping (Fig. 3, lane 2) was translated to provide chloramphenicol resistance.

Mutation (AUG→ACC) of the *cat* translational start site. The *S. acrimycini cat* start codon was identified by Murray et al. (20). However, additional in-frame potential start codons can be found downstream of the proposed translational start site (Fig. 1; codons 35, 37, and 39); initiation from a downstream site may account for the chloramphenicol resistance observed in *E. coli* (Fig. 4). To determine if the 5′-terminal AUG of the leaderless *cat* mRNA was necessary for chloramphenicol resistance, the AUG triplet was changed to ACC. Mutation of an AUG start codon to ACC would be expected to dramatically reduce translation initiation; however, initiation from a downstream site would be relatively unaffected by a two-nucleotide change near the 5′ end of the mRNA.

After the site-directed mutation of AUG to ACC, nuclease S1 was used to demonstrate the presence of *cat* mRNA and to verify that the transcriptional start site was unchanged (Fig. 3, lane 8). Chloramphenicol resistance in cells containing *Plac*(*ACC*):*cat* was at background levels (Fig. 4), suggesting that loss of resistance resulted from mutation of the 5′-terminal start codon. This result suggests strongly that alternate

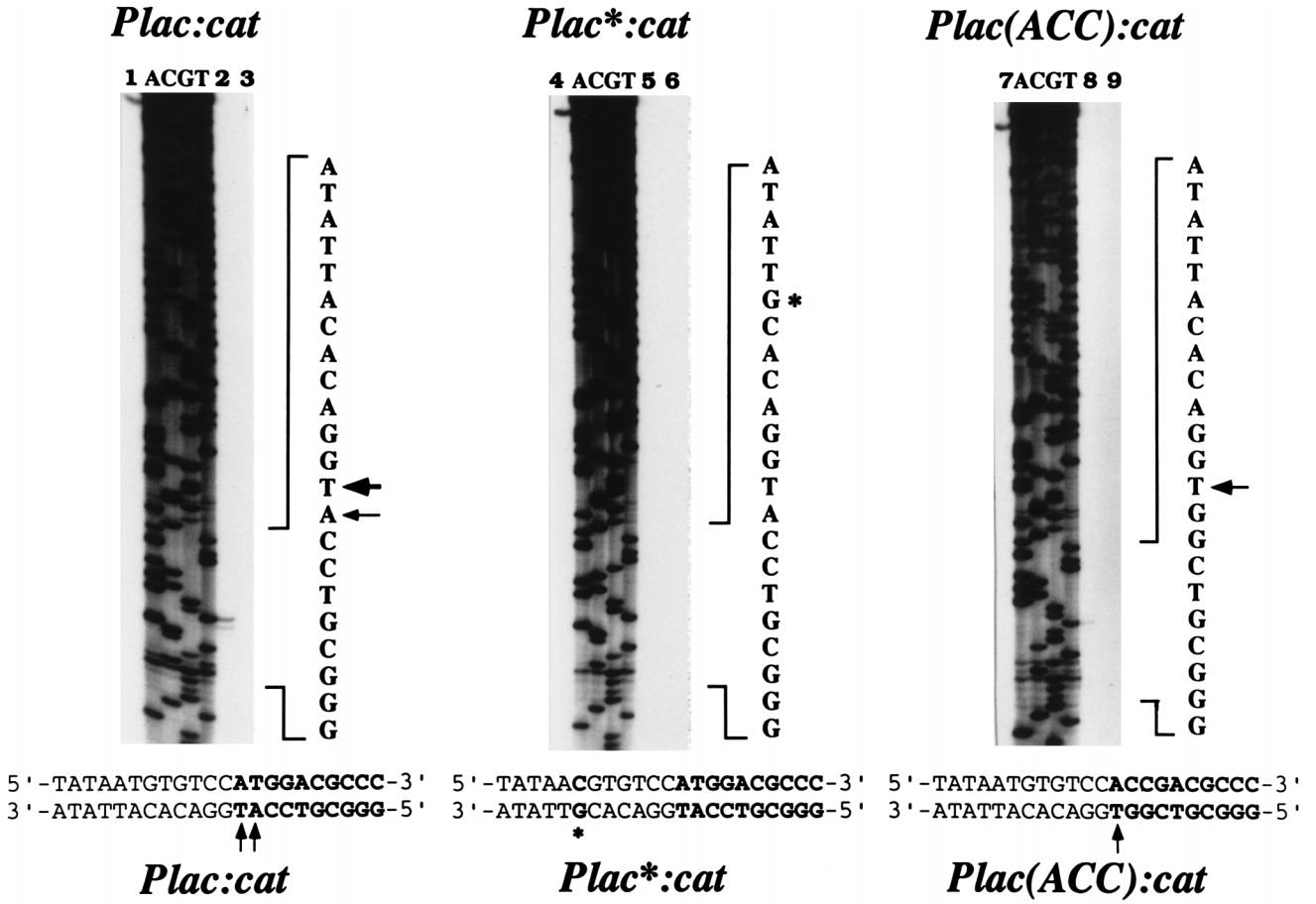

FIG. 3. Nuclease S1 mapping of *cat* mRNA initiated from *Plac*:*cat* (pIU459), *Plac**:*cat* (pIU1010), and *Plac*(*ACC*):*cat* (pIU1018). The PCR-amplified *Eco*RI-*Apa*LI fragments of *Plac*:*cat*, *Plac**:*cat* (pIU1010), and *Plac*(*ACC*):*cat* (pIU1018), uniquely labeled at the *Apa*LI 5′ end, were used as probes (lanes 1, 4, and 7) in hybridization reactions containing RNA synthesized in vivo. Total RNA was isolated from *E. coli* DH5α cells lacking the *cat* gene (lanes 3, 6, and 9) or containing *Plac*:*cat* (pIU459 [lane 2]), *Plac**:*cat* (pIU1010 [lane 5]), and *Plac*(*ACC*):*cat* (pIU1018 [lane 8]). Comparable amounts of the S1-digested reactions were loaded into all lanes. DNA sequencing ladders represent the template strand; arrows mark the transcription start sites identified by the migratory positions of the nuclease S1-resistant fragments. The *Plac* −10 region (5′-TATAAT-3′) has a sixth-position T→C mutation in the *Plac** promoter and is identified by an asterisk in the *Plac**:*cat* sequence. The *cat* coding sequence, initiating with an ATG triplet, is boldfaced.

in-frame start codons, downstream of the proposed start codon (Fig. 1), do not contribute to *cat* expression.

Start site utilization on *cat* mRNA containing two potential initiation codons near the 5′ end. Cells containing *Plac*:*4bp*:*cat* were sensitive to chloramphenicol, suggesting that the downstream in-frame AUG did not support significant levels of translation (Fig. 4). To quantify the translational activity of each of the two potential start codons, DNA fragments from *Plac*:*cat* and *Plac*:*4bp*:*cat* containing *cat* codons 1 to 16 were translationally fused to a *lacZ* reporter gene. Fusion of *Plac*:*cat* to *lacZ* (i.e., *Plac*:*cat*:*lacZ* [pIU1037]) resulted in 144.7 Miller units of β-galactosidase activity (Fig. 5). Cells containing a *Plac*:*4bp*:*cat*:*lacZ* fusion with the first (i.e., upstream) AUG in frame (pIU1038) resulted in 7.5 Miller units (Fig. 5) of activity, whereas the second (i.e., downstream) AUG in frame (pIU1039) produced 20.2 Miller units of activity. Fusion with neither AUG in frame (pIU1040) led to no β-galactosidase activity. These results indicated that both AUGs at or near the 5′ end of *Plac*:*4bp*:*cat*:*lacZ*-initiated mRNA were used for translation, but expression from each was lower than that from a single AUG (i.e., *Plac*:*cat*:*lacZ* in pIU1037).

In an effort to estimate the efficiency of translational signals contained in the leaderless *cat* mRNA, we compared the *lacZ* expression levels from *Plac*:*cat*:*lacZ* to the levels resulting from a *lacZ* fusion containing codons 1 to 16 of the bacteriophage lambda leaderless *c*I mRNA (i.e., *Plac*:*cI*:*lacZ* [pIU1041]) (Fig. 5). LacZ expression from *Plac*:*cat*:*lacZ* was only 1% of the activity measured from *Plac*:*cI*:*lacZ*, suggesting that the streptomycete leaderless *cat* mRNA functioned poorly in comparison to the leaderless *c*I mRNA.

Expression of *cat* after addition of an untranslated leader and SD sequence. Based on chloramphenicol resistance (Fig. 4) and LacZ activity from the *cat*:*lacZ* fusion (Fig. 5), the leaderless *cat* mRNA did not appear to be highly expressed in *E. coli*. In an effort to compare expression levels between a leaderless and a leadered *cat* mRNA, we prepared constructs *Plac*:LL:*cat*:*lacZ* (pIU1042) and *Plac*:GL:*cat*:*lacZ* (pIU1043) (see Materials and Methods). Relative to the leaderless *Plac*:*cat*:*lacZ* fusion, addition of the *lac* leader (pIU1042) increased expression 5.5-fold, whereas the *gus* leader (pIU1043) increased expression 9.3-fold (Fig. 5). Interestingly, addition of the *lac* or *gus* leader resulted in significantly less expression (5.0 and 8.4%, respectively) than measured for *Plac*:*cI*:*lacZ*. Although addition of an *E. coli* untranslated leader region stimulated expression, the translation signals within the lead-

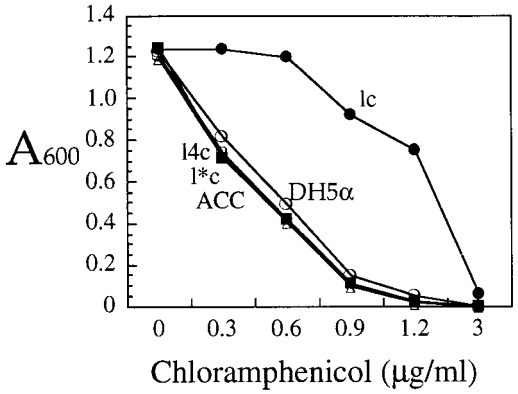

FIG. 4. Growth response (A_{600}) to chloramphenicol of *E. coli* DH5α cells lacking *cat* (DH5α) or containing *Plac*:*cat* (pIU459; lc), *Plac*:*4bp*:*cat* (pIU460; 14c) *Plac**:*cat* (pIU1010; 1*c), or *Plac*(*ACC*):*cat* (pIU1018; ACC).

ered *cat* sequence did not appear to function as well as the signals contained within the leaderless *c*I mRNA.

DISCUSSION

Translation of streptomycete leaderless *cat* mRNA in *E. coli*. The chloramphenicol resistance exhibited by cells containing *Plac*:*cat* indicated that *E. coli* was able to translate the streptomycete leaderless *cat* coding sequence. Chloramphenicol sensitivity resulting from *cat* fusion to a mutant *Plac* (i.e., *Plac**:*cat*) or from mutation of the *cat* mRNA 5′-terminal AUG→ACC [i.e., *Plac*(*ACC*):*cat*] indicated that the leaderless mRNA and a 5′-terminal start codon were both required for chloramphenicol resistance.

The observation that leaderless *cat* mRNA could be translated in *E. coli* suggests that information within the *cat* coding sequence is recognized by the translational machinery of both *Streptomyces* and *E. coli*. However, the low chloramphenicol resistance levels of cells containing *Plac*:*cat*, as well as the low β-galactosidase activities measured from *cat*:*lacZ* translational fusions, suggest that the streptomycete *cat* was not expressed well in *E. coli* even though it was under control of the strong *Plac* promoter. One possible explanation for these observations is that *E. coli* is very limited in its ability to translate leaderless mRNAs, a notion possibly consistent with the fact that only a small number of genes have been reported to encode leaderless mRNA in *E. coli*. Alternatively, efficient translation of leaderless mRNA may require specific mRNA-rRNA interactions, and because of differences between the *Streptomyces* and *E. coli* rRNA sequences, the complementary interactions needed for a high level of *cat* translation are not provided by *E. coli* ribosomes. It is also possible that additional factors, lacking from the *E. coli* translation system, are needed for efficient translation of the *cat* mRNA.

If the ribosome binding signals present on the *cat* mRNA do not function well in *E. coli*, then the *cat* mRNA might be

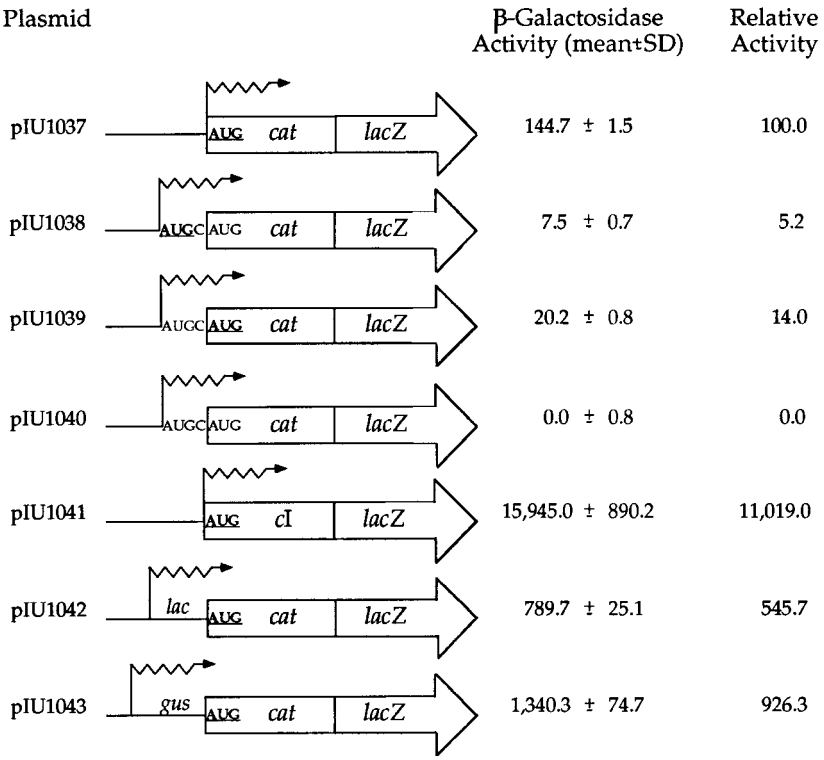

FIG. 5. Translational fusions of *cat* and *c*I coding sequences to a *lacZ* reporter gene. Start codons (AUG) in frame with *lacZ* are underlined. Codons 1 to 16 of *cat* (pIU1037) or *c*I (pIU1041) are fused to *lacZ* with the 5′-terminal AUG start codon in frame with *lacZ*. Plasmids pIU1038, pIU1039, and pIU1040 contain two potential start codons near the 5′ end of the mRNA. The upstream AUG of pIU1038 and the downstream AUG of pIU1039 are in frame with *lacZ*; neither AUG (i.e., 0) of pIU1040 is in frame with *lacZ*. Plasmids pIU1042 (*Plac*:LL:*cat*:*lacZ*) and pIU1043 (*Plac*:GL:*cat*:*lacZ*) contain, respectively, the untranslated leader regions from *lac* and *gus* upstream of *cat* codons 1 to 16 fused to *lacZ*. The untranslated leaders and coding sequences from *cat*, *c*I, and *lacZ* are not drawn to scale. Transcription is provided by the *Plac* promoter in all constructs and is indicated by a wavy line. β-Galactosidase activity is expressed as Miller units (19); relative activity is expressed relative to that of pIU1037 (*Plac*:*cat*:*lacZ*), which was set at 100%. *E. coli* RFS859, used as the host for *lacZ* fusion plasmids, does not contain measurable β-galactosidase activity.

relatively unprotected and quickly degraded (e.g., similar to observations made with *E. coli lacZ* [33]). A strong *Plac* combined with weak *E. coli* translation signals might result in an abundant leaderless *cat* mRNA with a short functional half-life and only a low level of translation.

Also, it is possible that the high G+C composition and biased codon usage observed with *Streptomyces* genes (1) cause *E. coli* ribosomes to stall during *cat* translation, resulting in less CAT protein and low resistance to chloramphenicol. However, comparison of the 16 *cat* codons present in the *cat*:*lacZ* fusion to the codon usage observed for the *E. coli lacZ* and *lacY* coding regions (14) revealed that only the *cat* CUC (Leu) codon was relatively infrequent, accounting for only 11% of the *lacZ* and *lacY* leucine codons. Applying a similar analysis to the *c*I sequence present in the *c*I:*lacZ* fusion, 6 of the 16 *c*I codons were present at a relative frequency of 11 to 16% of the total codons for a specific amino acid in *lacZ* plus *lacY*; even though the *c*I region contained six times as many underrepresented codons, the *c*I:*lacZ* expression levels were 110 times higher than those observed with the *cat*:*lacZ* fusion. By this analysis, it seems unlikely that the low expression observed with the *cat*:*lacZ* fusion resulted as a consequence of rare codons that impaired translation within the *cat* region.

Start site utilization on leaderless mRNA containing adjacent start codons. Addition of 5′-AUGC onto the 5′ terminus of the *cat* leaderless mRNA resulted in translation initiating from the authentic (downstream) start codon and the AUG triplet contained within the added sequence. In a similar experiment, Jones et al. (13) added a 5′-AUGC onto the 5′ terminus of a naturally leaderless *aph* mRNA and also observed that translation initiated from the added AUG triplet. Initiation from the upstream AUG suggests the possibility that ribosomes monitor, or examine, the mRNA 5′ terminus for a start codon. Examination of mRNA 5′ termini might result as a consequence of specific features of a leaderless mRNA that direct ribosomes to the 5′ end; alternatively, ribosomes, or other components of the translational machinery, might examine all nascent mRNA for the occurrence of a 5′-terminal initiation codon. Observations that ribosomes translated the *vph* mRNA from a 5′-terminal AUG (31), after removal of its untranslated leader sequence, support the notion that ribosomes do not limit their examination of 5′ termini to naturally leaderless mRNA.

It is interesting that *lacZ* expression in the *Plac*:*4bp*:*cat*:*lacZ* fusions, with either the upstream or downstream AUG in frame, was significantly lower than in the *Plac*:*cat*:*lacZ* fusion containing a single AUG. If ribosomes position themselves at one or the other of the two AUGs, one might expect the total number of initiation events from both AUGs to equal, or possibly exceed, the initiation events from an mRNA with a single AUG at the 5′ terminus. Contrary to expectation, the combined expression from the upstream and downstream AUGs present in the *Plac*:*4bp*:*cat*:*lacZ* fusions was less than expression from the single AUG present in *Plac*:*cat*:*lacZ*. A possible explanation for poor utilization of the downstream AUG might be a combination of its recessed position, relative to the 5′ end of the mRNA, and competition from the upstream AUG. The upstream added AUG might be improperly spaced relative to downstream signals that contribute to AUG placement into the ribosomal decoding site, thereby decreasing expression. Alternatively, these results might also suggest that the presence of two AUGs somehow interferes with initiation from either AUG; the interference could occur at the level of AUG recognition, competition for ribosomal P and A sites resulting in a paused initiation, an inhibitory mRNA structure, or other unknown mechanisms.

Efficiency of *cat* expression in *E. coli*. In the absence of information on the translational efficiency of the leaderless *cat* mRNA in *Streptomyces*, it is not possible to make comparisons to the apparent low efficiency observed for *E. coli*; however, the *lacZ* fusions suggested that the *cat* translational signals functioned poorly relative to the *c*I signals. Although addition of an untranslated leader increased expression of the *lacZ* reporter gene, the 5- or 9-fold increase seen with the *lac* or *gus* leader, respectively, was dramatically lower than the 110-fold increase observed when 16 codons of the *c*I gene were fused to *lacZ*. Low *cat*:*lacZ* activity in the presence of the *lac* or *gus* leader indicates that the mere addition of an untranslated leader was insufficient to confer high expression levels to the leaderless *cat* mRNA. Because translation signals have been localized upstream (2, 21, 30) and downstream (3, 29) of the start codon, the low *cat* expression may indicate that signals within the *lac* or *gus* leader are optimized for *E. coli* expression while signals within the cat coding sequence are optimized for the streptomycete translation system; *cat* expression as a leadered mRNA in *E. coli* might be low because of incomplete translation signals (i.e., upstream and downstream) needed for high expression. An alternative explanation is that sequence or structural features present in *cat* codons 1 to 16 are detrimental to efficient transcription or translation in *E. coli*, and the inhibitory effects of this region are not overcome by the addition of an *E. coli* untranslated leader sequence.

Expression features of leaderless mRNA. The results described here suggest that a 5′-terminal AUG, while a feature essential for expression, does not provide all of the information for determining translation levels from a leaderless mRNA. It seems likely that a combination of a 5′ AUG and downstream sequences, or mRNA structural features, specifies a leaderless mRNA's translational efficiency. The leaderless *cat* mRNA provided the 5′ AUG but only weak downstream signals, thereby resulting in only a low level of expression. Elucidation of how ribosomes initiate from the 5′ terminus of leaderless mRNA, combined with the identification of downstream features that influence expression levels, will contribute to our understanding of leaderless mRNA translation, as well as provide additional insights into translation initiation with leadered mRNA. Characterization of nucleotides within the coding sequence that contribute to formation of a translation initiation complex will further define the molecular events prerequisite to translation initiation and facilitate the design of coding sequences for maximal translation levels.

ACKNOWLEDGMENTS

We thank William Van Etten for comments on the manuscript, J. R. Liston for assistance with preparation of figures, David Hopwood for the *cat*-containing plasmid pIJ879, and Angela Walker for DNA fragments containing the *Plac* promoter with *lac* and *gus* leader sequences. We also extend a special acknowledgment and dedication to Sioni Maluf Barbieri (deceased) for her encouragement, enthusiasm, and support of our efforts.

This work was supported by Public Health Service grant GM45923 from the National Institutes of Health.

REFERENCES

1. **Bibb, M. J., P. R. Findlay, and M. W. Johnson.** 1984. The relationship between base composition and codon usage in bacterial genes and its use for the simple and reliable identification of protein-coding sequences. Gene **30:**157–166.
2. **Boni, I. V., D. M. Isaeva, M. L. Musychenko, and N. V. Tzareva.** 1991. Ribosome-messenger recognition: mRNA target sites for ribosomal protein S1. Nucleic Acids Res. **19:**155–162.
3. **Chen, H., L. Pomeroy-Cloney, M. Bjerknes, J. Tam, and E. Jay.** 1994. The influence of adenine-rich motifs in the 3′ portion of the ribosome binding site on human IFN-gamma gene expression in *Escherichia coli*. J. Mol. Biol. **240:**20–27.

4. **Christie, G. E., and R. Calendar.** 1985. Bacteriophage P2 late promoters II. Comparison of the four late promoter sequences. J. Mol. Biol. **181:**373–382.
5. **Favaloro, J., R. Triesman, and R. Kamen.** 1980. Transcription maps of polyoma virus-specific RNA: analysis by two-dimentional nuclease S1 gel mapping. Methods Enzymol **65:**718–749.
6. **Gallie, D. R., D. E. Sleat, J. W. Watts, P. C. Turner, and T. M. A. Wilson.** 1987. The 5′-leader sequence of tobacco mosaic virus RNA enhances the expression of foreign gene transcripts *in vitro* and *in vivo*. Nucleic Acids Res. **15:**3257–3272.
7. **Hopwood, D. A., M. J. Bibb, K. F. Chater, T. Kieser, C. J. Bruton, H. M. Kieser, D. J. Lydiate, C. P. Smith, J. M. Ward, and H. Schrempf.** 1985. Genetic manipulation of *Streptomyces*: a laboratory manual. John Innes Foundation, Norwich, England.
8. **Hui, A., and H. A. deBoer.** 1987. Specialized ribosome system: preferential translation of a single mRNA species by a subpopulation of mutated ribosomes in *Escherichia coli*. Proc. Natl. Acad. Sci. USA **84:**4762–4766.
9. **Jacob, W. F., M. Santer, and A. E. Dahlberg.** 1987. A single base change in the Shine-Dalgarno region of 16S rRNA of *Escherichia coli* affects translation of many proteins. Proc. Natl. Acad. Sci. USA **84:**4757–4761.
10. **Janssen, G. R.** Unpublished data.
11. **Janssen, G. R.** 1993. Eubacterial, archaebacterial, and eucaryotic genes that encode leaderless mRNA, p. 59–67. *In* R. H. Baltz, G. D. Hegeman, and P. L. Skatrud (ed.), Industrial microorganisms: basic and applied molecular genetics. American Society for Microbiology, Washington, D.C.
12. **Janssen, G. R., J. M. Ward, and M. J. Bibb.** 1989. Unusual transcriptional and translational features of the aminoglycoside phosphotransferase gene (*aph*) from *Streptomyces fradiae*. Genes Dev. **3:**415–429.
13. **Jones, R. L., III, J. C. Jaskula, and G. R. Janssen.** 1992. In vivo translational start site selection on leaderless mRNA transcribed from the *Streptomyces fradiae aph* gene. J. Bacteriol. **174:**4753–4760.
14. **Kalnins, A., K. Otto, U. Rüther, and B. Müller-Hill.** Sequence of the *lacZ* gene of *Escherichia coli*. EMBO J. **2:**593–597.
15. **Klock, G., and W. Hillen.** 1986. Expression, purification and operator binding of the transposon Tn*1721*-encoded Tet repressor. J. Mol. Biol. **189:**633–641.
16. **Kunkel, T. A., J. D. Roberts, and R. A. Zakour.** 1987. Rapid and efficient site-specific mutagenesis without phenotypic selection. Methods Enzymol. **154:**367–382.
17. **Linn, T., and T. St. Pierre.** 1990. Improved vector system for constructing transcriptional fusions that ensures independent translation of *lacZ*. J. Bacteriol. **172:**1077–1084.
18. **Loechel, S., J. M. Inamine, and P. Hu.** 1991. A novel translation initiation region from *Mycoplasma genitalium* that functions in *Escherichia coli*. Nucleic Acids Res. **19:**6905–6911.
19. **Miller, J. H. (ed.).** 1992. A short course in bacterial genetics. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
20. **Murray, I. A., J. A. Gil, D. A. Hopwood, and W. V. Shaw.** 1989. Nucleotide sequence of the chloramphenicol acetyltransferase gene of *Streptomyces acrimycini*. Gene **85:**283–291.
21. **Olins, P. O., and S. H. Rangwala.** 1989. A novel sequence element derived from bacteriophage T7 mRNA acts as an enhancer of the *lacZ* gene in *Escherichia coli*. J. Biol. Chem. **264:**16973–16976.
22. **Ptashne, M., K. Backman, M. Z. Humayum, A. Jeffrey, R. Maurer, B. Meyer, and R. T. Sauer.** 1976. Autoregulation and function of a repressor in bacteriophage lambda. Science **194:**156–161.
23. **Sabelnikov, A. G., B. Greenberg, and S. A. Lacks.** 1995. An extended −10 promoter alone directs transcription of the *Dpn*II operon of *Streptococcus pneumoniae*. J. Mol. Biol. **250:**144–155.
24. **Schlief, R.** 1972. Fine-structure deletion map of the *Escherichia coli* L-arabinose operon. Proc. Natl. Acad. Sci. USA **69:**3479–3484.
25. **Schottel, J. L., M. J. Bibb, and S. N. Cohen.** 1981. Cloning and expression in *Streptomyces lividans* of antibiotic resistance genes derived from *Escherichia coli*. J. Bacteriol. **146:**360–368.
26. **Shaw, W. V.** 1967. The enzymatic acetylation of chloramphenicol by extracts of R factor-resistant *Escherichia coli*. J. Biol. Chem. **242:**687–693.
27. **Shaw, W. V.** 1983. Chloramphenicol acetyltransferase: enzymology and molecular biology. Crit. Rev. Biochem. **14:**1–46.
28. **Shine, J., and L. Dalgarno.** 1974. The 3′-terminal sequence of *Escherichia coli* 16S ribosomal RNA: complementarity to nonsense triplet and ribosome binding sites. Proc. Natl. Acad. Sci. USA **71:**1342–1346.
29. **Sprengart, M. L., E. Fuchs, and A. G. Porter.** 1996. The downstream box: an efficient and independent translation signal in *Escherichia coli*. EMBO J. **15:**665–674.
30. **Steitz, J. A., and K. Jakes.** 1975. How ribosomes select initiator regions in mRNA: base pair formation between the 3′ terminus of 16S rRNA and the mRNA during initiation of protein synthesis in *Escherichia coli*. Proc. Natl. Acad. Sci. USA **72:**4734–4738.
31. **Wu, C.-J., and G. R. Janssen.** 1996. Translation of *vph* mRNA in *Streptomyces lividans* and *Escherichia coli* after removal of the 5′ untranslated leader. Mol. Microbiol. **22:**339–355.
32. **Wu, P. C., K. V. Srinivasan, and K. E. Kendrick.** 1995. Regulated expression of the histidase structural gene in *Streptomyces griseus*. J. Bacteriol. **177:**854–857.
33. **Yarchuk, O., N. Jacques, J. Guillerez, and M. Dreyfus.** 1992. Interdependence of translation, transcription and mRNA degradation in the *lacZ* gene. J. Mol. Biol. **226:**581–596.