Expression of a Streptomyces 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 a 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 cI 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 sequence to the 3′ end of the 16S rRNA (8, 9, 28, 30). Small subunit ribosomes from procaryotic ribosomes contain the conserved ASD region, and most procaryotic mRNAs contain a readily identifiable SD sequence. While the contributions of the SD-ASF 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 cI 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 region is 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 cI (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− Δ(lacZYA-argF)U169 relA1 endA1 hsdR17 (rK− mK−) supE44 thi-1 gyrA96 relA1; Bethesda Research Laboratories] was used for all standard cloning procedures. E. coli C1236 [supE44 thi-1 Δ(lacZYA-argF)U169 relA1 recA1 endA1 hsdR17 (rK− mK−) supE44 thi-1 gyrA96 relA1; Bethesda Research Laboratories] was used for site-directed mutagenesis, and E. coli RFS859 (F− thi-1 araC859 leuB6 Δlac74) was used for

* 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.
**RESULTS**

**Fusion of the cat coding sequence to an E. coli Plac promoter.** The S. acrincyni *cat* gene was subcloned from pIJ879 (20). The site-directed mutagenesis method of Kunkel et al. (16) was used to create *Neol* restriction sites at the cat transcriptional start site and the transcriptional start site of a modified *lacZ* promoter. The cat and *lacZ* leader sequences, with the 5'-terminal nucleotide and extending to the initiation codon, are 5'-ATGGTACGTCGAGAACAAGAGCCGCTTACATCTC-3' and 5'-ATGACGACAAAAAAGAACCATTACAAAAAGACCGAC-3'.

Modified untranslated leader regions from the *lac* operon or the *ｇｕｓ* gene were added to give *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 *catlacZ* 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'-ATGTGACGTCGAGAACAAGAGCCGCTTACATCTC-3' and 5'-ATGACGACAAAAAAGAACCATTACAAAAAGACCGAC-3'.

**RNA isolation.** E. coli DH5α cells containing pIU459, pIU460, pIU1010, or pIU1018 were harvested by centrifugation at 4°C 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; pIU459 or pIU460) were used in S1 mapping of cat mRNA start site were prepared by PCR: Approximately 120 pmol of the oligonucleotide oligonucleotide 5'-CGGGGAAGGAGCGCCCGGCGG-3' was phosphorylated by using 1.5 μl of [γ-32P]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'-CGGAATTCGCTGCGACGAGTTTCC-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.

**Tissue fusion to lacZ.** Cat mRNA hybridization was used to introduce a *Sal* restriction site into the 16th codon of the *cat* and *cl* coding region. PCR-generated EcoRI-SalII fragments, containing the Plac promoter and 16 codons of **cat** or *cl*, were ligated to plasmids containing a *Sal* 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. 5. The 5'-DNA sequence fused to *lacZ*, beginning with the initiation codon, added downstream potential start codons are boldface and underlined.

**RESULTS**

**Fusion of the cat coding sequence to an E. coli Plac promoter.** The *S. acrincyni* *cat* gene was subcloned from pIJ879 (20). The site-directed mutagenesis method of Kunkel et al. (16) was used to create *Neol* restriction sites at the cat transcriptional start site and the transcriptional start site of a modified *lacZ* promoter (Plac [31]). The cat coding sequence was then ligated to the Plac promoter via common restriction sites. Creation of an *Neol* 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 Placcat (pIU459) *Neol* restriction site, resulting in the Plac-4bp *cat* (pIU460) construct containing a 4-bp (5'-ATGC-3') insertion at the *Plac* transcription/translational start site.
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. 2.** Nuclease S1 mapping of cat mRNA initiated from Plac:cat (pIU459) and Plac:4bp:cat (pIU460). The PCR-amplified EcoRI-ApaLI fragment of Plac:cat, Plac:4bp:cat, uniquely labeled at the ApaLI 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.
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 codons 1 to 16 were translationally fused to a lacZ reporter gene. Fusion of Plac:cat to lacZ (i.e., Plac:catlacZ [plU1037]) resulted in 144.7 Miller units of β-galactosidase activity (Fig. 5). Cells containing a Plac:4bp:catlacZ fusion with the first (i.e., upstream) AUG in frame (pIU1038) resulted in 7.5 Miller units 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:catlacZ-initiated mRNA were used for translation, but expression from each was lower than that from a single AUG (i.e., Plac:catlacZ 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:catlacZ to the levels resulting from a lacZ fusion containing codons 1 to 16 of the bacteriophage lambda leaderless cI mRNA (i.e., Plac:4bp:catlacZ [plU1041]) (Fig. 5). LacZ expression from Plac:catlacZ was only 1% of the activity measured from Plac:Plac:catlacZ, suggesting that the streptomycete leaderless cat mRNA functioned poorly in comparison to the leaderless cI 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:catlacZ (pIU1042) and Plac:GL:catlacZ (pIU1043) (see Materials and Methods). Relative to the leaderless Plac:LL:catlacZ 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:catlacZ. Although addition of an E. coli untranslated leader region stimulated expression, the translation signals within the lead-
 Yorker 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

![Graph](http://jb.asm.org/)

**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).
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 ci sequence present in the ci::lacZ fusion, 6 of the 16 ci 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 ci region contained six times as many underrepresented codons, the ci::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′-AUGG 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′-AUGG 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::Aph::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::Aph::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 function poorly relative to the ci 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 ci 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 streptomyces 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′-AUGG 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 fragment 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**


