Proc. Natl. Acad. Sci. USA
Vol. 83, pp. 8506–8510, November 1986
Biochemistry

Translation of a synthetic two-cistron mRNA in *Escherichia coli*

(bovine growth hormone/human growth hormone/runaway replicon)

BRIGITTE E. SCHONER, RAMA M. BELAGAJE, AND RONALD G. SCHONER

Division of Molecular and Cell Biology, Lilly Research Laboratories, Indianapolis, IN 46285

Communicated by Phillip A. Sharp, July 24, 1986

ABSTRACT **A synthetic two-cistron expression system was constructed for the high-level expression of eukaryotic genes in *Escherichia coli*. This system was designed to overcome translational inhibition of mRNAs containing eukaryotic sequences. The first cistron in this system is a 31-base A+T-rich synthetic sequence that provides for efficient translation initiation. The second cistron contains the protein coding sequence for the eukaryotic gene. Insertion of the first cistron between the 5′ untranslated region of the mRNA and the protein coding region separates the two and thereby potentially minimizes the formation of local secondary structures that might prevent ribosomes from binding and initiating translation. The 31-base cistron contains three nonsense codons (TAA), one in each of the three translational reading frames, and an 8-base Shine–Dalgarno sequence that is complementary to the 3′ end of the 16S rRNA. The effects of translation of the first cistron in all three reading frames on the expression of the second cistron was examined. The most efficient expression of the second cistron seemed to occur when the stop codon that terminates translation of the first cistron is located 3′ to the Shine–Dalgarno sequence and close to the AUG start codon for the second cistron. When the Shine–Dalgarno sequence was deleted from the first cistron, no detectable expression of the second cistron was observed. This two-cistron system has been used to express the gene encoding methionylalanyl bovine growth hormone with its native codons and the gene encoding methionyl human growth hormone at a level greater than 20% of total cell protein. In the case of human growth hormone, we show that the amount of gene product is not significantly diminished by placing a "functional" first cistron in front of a gene that can be expressed without a cistron.**

In recent years, many prokaryotic and eukaryotic genes have been expressed at high levels in *Escherichia coli*. The general approach has been to use a multicopy cloning vector with a strong promoter and an efficient ribosome binding site for the transcription and translation of the cloned gene (1, 2). However, the level of gene expression with these vectors varies widely for different eukaryotic genes. Low-level expression has been attributed to protein degradation by *E. coli* proteases (3) or to inefficient translation initiation of mRNAs containing heterologous gene sequences (4–8). Several studies suggested that the efficiency of translation initiation depends on the degree of complementarity between the Shine–Dalgarno (SD) sequence and the 16S rRNA, the distance between the SD sequence and the initiation codon, and the nucleotide sequence of this "window" region (9–16). There is evidence that the translational efficiency also depends on the sequence of the 5′ untranslated region of the mRNA outside the SD sequence and the 5′ end of the protein coding region (17–19).

To reconcile these observations, it has been proposed that translation is inhibited when local secondary structures form with regions containing the SD sequence and/or the AUG start codon such that the ribosomes cannot initiate translation (20–25). The formation of such secondary structures may explain failures to express methionyl bovine growth hormone (Met-bGH) with its native codons at high levels (26, 27). To overcome this potential problem, Seeburg *et al.* (27) have introduced several base changes into the 5′ end of the bovine growth hormone (bGH) gene to create a sequence that is similar to the 5′ end of the highly expressed human growth hormone (hGH) gene. Likewise, George *et al.* (26) reported high-level expression (15% of total cell protein) after changing 13 codons in the 5′ end of the bGH gene. These approaches are limited by the need to preserve the amino acid sequence of the protein. To avoid such limitations, we have demonstrated that a two-cistron expression system can be used to synthesize Met-bGH at a high level with its native codons (28).

In this paper we have further examined translation of a synthetic two-cistron mRNA encoding methionylalanyl bGH (Met-[Ala]bGH) and other eukaryotic gene products in *E. coli*. In this system the first cistron provides a sequence for the efficient initiation of translation and the second cistron contains the protein coding information. The first cistron also separates the protein coding region from the 5′ untranslated region and thus reduces the formation of local secondary structures that might interfere with translation initiation. For this study, we chemically synthesized a 31-base first-cistron sequence that contains the following features: (*i*) several unique restriction sites; (*ii*) a translational initiation codon (ATG) followed by three translational stop codons (TAA), one in each reading frame; and (*iii*) a SD sequence with an 8-base complementarity to the 16S rRNA. This cistron sequence was inserted initially into Met-[Ala]bGH expression vectors to analyze the effects of sequence changes in the first cistron on Met-[Ala]bGH expression. Subsequently, we tested this system for the expression of other mammalian genes.

MATERIALS AND METHODS

Bacterial Strains and Media. *E. coli* K-12 RV308 [*su, ΔlacX74, gal IS II::OP308, strA*] (29) was the host strain for all two-cistron expression plasmids. Some of the intermediate plasmids in the constructions were propagated in *E. coli* K-12 RR1 (*hsdS20, ara-14, leuB6, proA2, lacY1, galK2, rpsL20, xyl-5, mtl-1, supE44*; ATCC 31343). Plasmid pIMIA (1) was obtained from M. Inouye. In all experiments, RV308 cells containing recombinant plasmids were grown in TY broth (Difco) with 50 μg of kanamycin per ml (Sigma) at temperatures indicated in the text.

Chemical Synthesis of DNA Linker Sequences. All deoxyribo-oligonucleotides were prepared with a DNA synthesizer [model 380A; Applied Biosystems (Foster City, CA)]

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

Abbreviations: bGH, bovine growth hormone; Met-bGH, methionyl bGH; Met-[Ala]bGH, methionylalanyl bGH; hGH, human growth hormone; Met-hGH, methionyl hGH; SD, Shine–Dalgarno.

according to the procedures recommended by the manufacturer. The oligonucleotides were purified by HPLC using a Whatman Partisil-10 SAX column or by gel electrophoresis (30). The linkers were prepared by joining oligonucleotides enzymatically using T4 DNA ligase (31).

Plasmid Constructions. Conditions for all enzymatic reactions were those recommended by the manufacturer (New England Biolabs or Boehringer Mannheim). For plasmid constructions, restriction fragments purified from agarose gels (Sigma) (32) were ligated with T4 DNA ligase and transformed into competent $CaCl_2$-treated *E. coli* cells (33). Transformants were selected at 25°C on TY agar plates containing 50 μg of kanamycin per ml. Plasmid DNA from the transformants was extracted (34) and examined by restriction analysis. Appropriate restriction fragments containing synthetic linkers were end-labeled with [γ-^{32}P]ATP and polynucleotide kinase and sequenced by the method of Maxam and Gilbert (35).

Polyacrylamide Gel Electrophoresis. Polyacrylamide gels (36) were used to analyze cells for Met-[Ala]bGH accumulation. Cell pellets were dissolved in modified sample buffer [0.125 M Tris·HCl, pH 6.8/2% $NaDodSO_4$/30% (vol/vol) glycerol/1 M 2-mercaptoethanol/6 M urea] and boiled for 3 min before loading. After staining with Coomassie blue (Sigma), gels were scanned with a Shimadzu 930 dual-wavelength TLC scanner (Columbia, MD) that integrates the areas under the peaks.

RNA Isolation and Blotting. Cells grown at 25°C to early logarithmic phase were shifted to 37°C for 3 hr, and total cellular RNA was isolated as described by Young and Furano (37). Serial dilutions of the RNA were applied directly to nitrocellulose paper (38). The conditions for nick-translation, hybridization, and washing have been described (39).

RESULTS

Design and Construction of a Two-Cistron Expression System. We have demonstrated previously (28) that Met-bGH with its native codons can be overproduced in *E. coli* with a two-cistron expression system. To determine what features of this system are required for high-level expression, we chemically synthesized a different first-cistron sequence that is A+T-rich (23/31 bases) and contains several unique restriction sites to facilitate subsequent plasmid constructions. Fig. 1 shows the DNA sequence corresponding to the 5′ end of the two-cistron mRNA when transcribed from the *E. coli lpp* promoter. The synthetic DNA sequence extends from the *Xba* I to the *Nde* I restriction sites and contains the 31-base region that we refer to as the first cistron. We have designed this sequence such that translation can start at the ATG codon (for Met 1) and potentially terminate at any one of 3 TAA stop codons. In the sequence shown in Fig. 1, the S+0 stop codon is in frame with the ATG; a shift to the reading frames containing the S+1 and the S+2 stop codons can be accomplished by inserting or deleting bases between the ATG codon and the *Cla* I restriction site. The 8-base sequence designated SD 2 is complementary to the 3′ end of the 16S rRNA. The ATG in the *Nde* I restriction site can serve as a translational initiation codon for coding sequences inserted at the *Nde* I site. The unique *Cla* I site near the 5′ end of the first cistron was included to facilitate subsequent sequence manipulations.

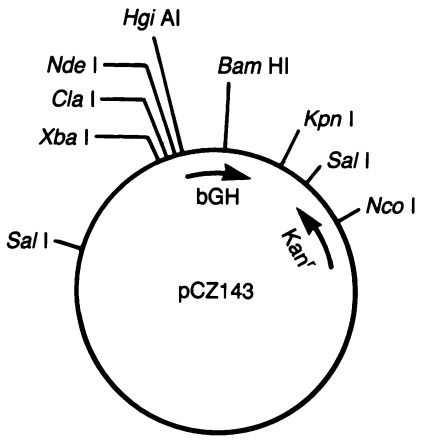

FIG. 2. Restriction map of plasmid pCZ143. The starting plasmid was pCZ100 (28). The construction of pCZ143 starting with pCZ100 has been described (41). The locations of the bGH gene and the kanamycin resistance gene (Kan^r) and the directions of transcription are indicated by arrows.

The synthetic cistron sequence was cloned between the *Xba* I and *Nde* I restriction sites in pCZ140 to create pCZ143 (Fig. 2). The bGH coding sequence present in pCZ143 contains the native codons for Met-[Ala]bGH, which is the methionyl derivative of the 191-amino acid protein that accounts for about 50% of growth hormone isolated from bovine pituitary (42); the remaining 50% lacks the alanine residue. These two forms arise presumably as a result of differential processing of the signal peptide (43). Because Met-[Ala]bGH with its native codons has been particularly difficult to express at high levels in *E. coli*, we chose this gene for our initial analysis of the two-cistron system.

Effect of Sequence Changes in the First Cistron on the Synthesis of Met-[Ala]bGH. As shown in Fig. 3 (lane 1), no detectable amounts of Met-[Ala]bGH were produced (as expected) by cultures harboring plasmid pCZ140. However, cultures harboring plasmid pCZ143 (lane 2) did not produce any measurable amounts of Met-[Ala]bGH either, indicating that the first cistron present in pCZ143 is not suitable for high-level expression of Met-[Ala]bGH.

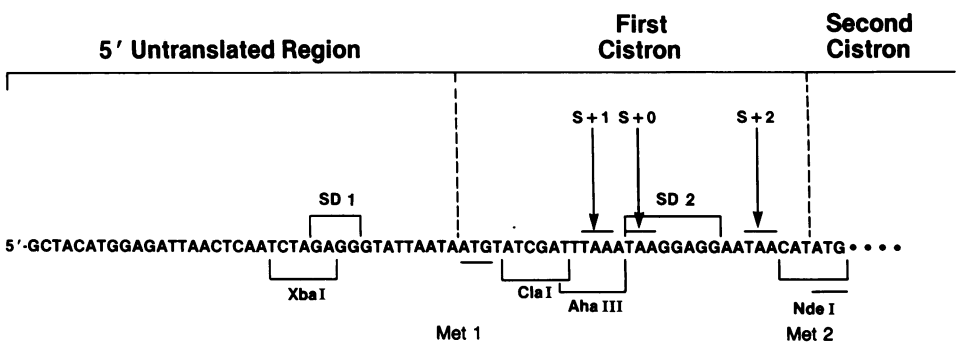

FIG. 1. DNA sequence corresponding to the 5′ end of the two-cistron mRNA. The 5′ untranslated region includes the SD 1 and extends to the ATG codon for Met 1. This sequence is identical to that of the *E. coli lpp* mRNA (40). The region 3′ to the ATG codon for Met 1 is the synthetic first cistron sequence. Restriction sites and the SD sequences are indicated by brackets; the three stop codons are indicated by arrows.

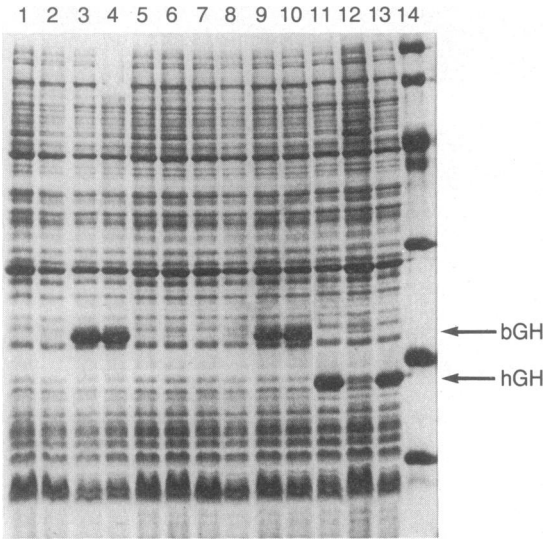

FIG. 3. Polyacrylamide gel analysis of gene expression. Lysates were prepared from 1-ml cultures of cells harboring the plasmids described in Fig. 4 after 6 hr of growth at 37°C. The lysates were loaded on a 12.5% polyacrylamide gel (1.5 mm × 18 cm × 18 cm) and run at 50 mA for about 3 hr. Lanes: 1, pCZ140; 2, pCZ143; 3, pCZ144; 4, pCZ145; 5, pCZ146; 6, pCZ147; 7, pCZ148; 8, pCZ180; 9, pCZ181; 10, pCZ182; 11, pCZ340; 12, pCZ343; 13, pCZ345; 14, molecular weight standards: phosphorylase B (92,500), bovine serum albumin (66,200), ovalbumin (45,000), carbonic anhydrase (31,000), soybean trypsin inhibitor (21,500), and lysozyme (14,400). The arrows indicate the positions of Met-[Ala]bGH and the 20-kDa derivative of hGH.

To investigate the lack of high-level expression of Met-[Ala]bGH with this particular first-cistron sequence, we introduced the series of base changes shown in Fig. 4. In pCZ143, the ATG (for Met 1) is in frame with the S+0 stop codon. In pCZ144 and pCZ145, the reading frame was shifted either by inserting two bases (G-C) at the *Cla* I restriction site or by deleting one base (thymine) from the *Aha* III restriction site in pCZ143. These changes placed the S+2 stop codon into frame with respect to the ATG codon (for Met 1). Measurements of Met-[Ala]bGH synthesis in cells harboring plasmids pCZ144 or pCZ145 revealed a dramatic increase in the amount of Met-[Ala]bGH produced compared with cells harboring pCZ143. Next, we shifted the reading frame to the S+1 stop codon by either inserting two bases (G-C) at the *Cla* I site in plasmid pCZ145 or by adding one base (cytosine) near the *Cla* I site in pCZ143. RV308 cells harboring these two plasmids (pCZ146 and pCZ147) produced undetectable amounts of Met-[Ala]bGH. These results suggested that the level of expression of the second cistron is very sensitive to changes in the first cistron that alter: (*i*) the length of the open reading frame in the first cistron, (*ii*) the position of the stop codon that terminates translation of the first cistron, (*iii*) the length of the intercistronic region, and (*iv*) the phasing of the stop codon in the first cistron relative to the start codon for the second cistron. To determine which of these variables is important, we first analyzed the effect of changing the phasing of the stop codon in the first cistron relative to the restart codon for the second cistron. When the stop and restart codons in pCZ143 were shifted into frame by inserting two bases (A-T) at the *Nde* I restriction site of pCZ143 (to create pCZ180), no detectable change in the low-level expression of Met-[Ala]bGH was observed. Shifting the stop and restart codons in pCZ145 out of frame and either two bases farther away from the restart codon (as in pCZ181) or two bases closer to the restart codon (as in pCZ182) decreases somewhat the expression of Met-[Ala]bGH relative to the level obtained with plasmids pCZ144 and pCZ145.

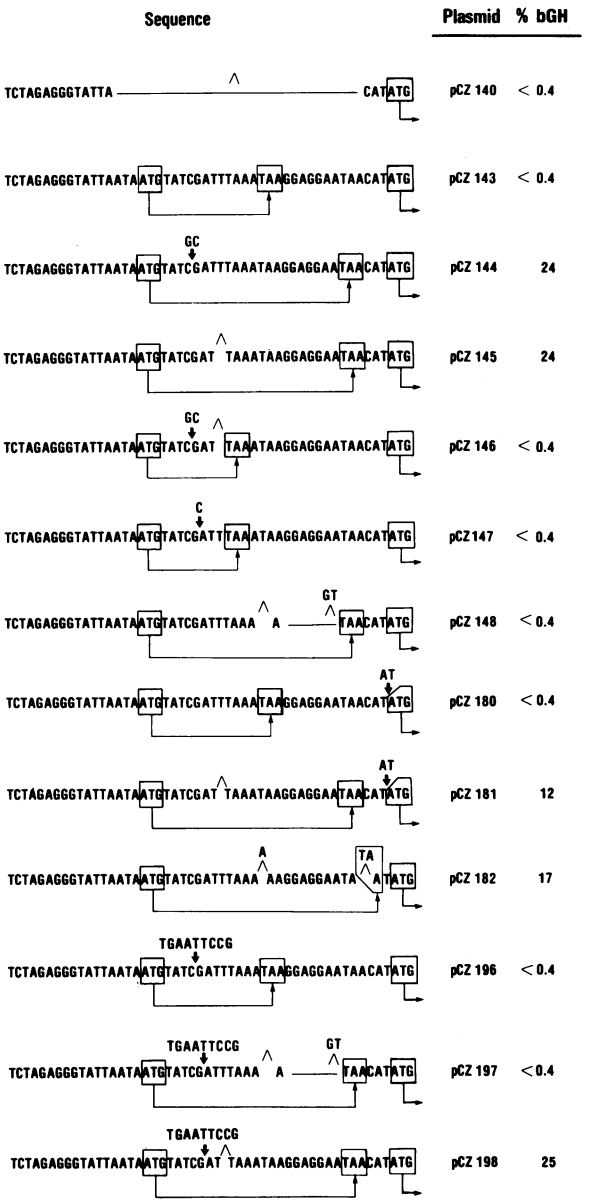

FIG. 4. First-cistron sequences. The only differences between the various two-cistron expression plasmids are localized within the region shown. This region extends from the *Xba* I site in the 5′ untranslated region of the two-cistron mRNA to the *Nde* I site at the beginning of the second cistron. Sequence changes are relative to the sequence in pCZ143. Some were introduced by "filling in" bases at restriction sites for *Cla* I (in pCZ144 and pCZ146) and *Nde* I (in pCZ180 and pCZ181). This was accomplished by cutting the starting plasmids with *Cla* I or *Nde* I, treating the linearized plasmids with the Klenow fragment of DNA polymerase I in the presence of all four deoxyribonucleotides, and reclosing the plasmid with T4 DNA ligase. Other changes were made by replacing the region containing the sequence change by synthetic DNA fragments. Specifically, the *Cla* I and *Nde* I restriction sites were used to construct pCZ145, pCZ147, pCZ148, and pCZ182. pCZ196, pCZ197, and pCZ198 were constructed by inserting a synthetic linker containing the bases shown, at the *Cla* I restriction site in pCZ143, pCZ148, and pCZ145, respectively. The arrows (↓) indicate the positions at which bases have been added, the carets (^) indicate deletions of bases. A caret with associated bases indicates the deletion of bases found in pCZ143 and the addition of new bases at that position. The predicted reading frame of the first cistron in each plasmid begins at the boxed-in ATG codon and terminates at the boxed-in TAA codon. The amounts of Met-[Ala]bGH measured in cells containing these plasmids are expressed as percent of total cell protein as determined by scanning Coomassie blue-stained polyacrylamide gels.

The addition of two bases (A-T) at the *Nde* I site in pCZ145 (to give pCZ181) not only changes the phasing of the stop and

restart codons but also expands the "window" region between SD 2 and the ATG codon for Met 2. Thus, if a SD sequence in the first cistron is required for Met-[Ala]bGH expression, a decrease in the amount of Met-[Ala]bGH produced might be expected as a result of changing the distance between the SD sequence and the restart codon. To examine this point further, we deleted SD 2 as shown in pCZ148 and pCZ197. The stop codons in pCZ148 and pCZ197 are in a favorable position relative to the restart codon, yet no detectable expression of Met-[Ala]bGH was observed. From these experiments we concluded that the SD sequence is required for expression of the second cistron. Further, the observed reduction in Met-[Ala]bGH expression with pCZ181 (where the window region was expanded by two bases) suggests that proper spacing between the SD sequence and the translational restart codon is important for optimal expression of the second cistron. The phasing of the stop codon and the restart codon for the second cistron appears to be relatively less important. These observations, however, do not adequately explain the large differences in the expression levels obtained between plasmids pCZ143 and pCZ144 or pCZ145. The open reading frame in the first cistron in pCZ143 contains 5 codons, whereas the open reading frame in pCZ144 and pCZ145 contains nine and eight codons, respectively. Thus, we questioned whether the number of codons that are translated in the first cistron are important for the efficient expression of the second cistron. To directly test this possibility, we inserted three codons into the *Cla* I restriction site of pCZ143 to create an open reading frame with eight codons (pCZ196). Expression of Met-[Ala]bGH from pCZ196 was undetectable. Therefore, we concluded that the position of the stop codon that terminates translation of the first cistron determines the efficiency of translation of the second cistron rather than the number of codons in the first cistron (i.e., five vs. eight or nine). Insertion of the same three codons into plasmid pCZ145 (to give pCZ198) did not lower the level of expression of Met-[Ala]bGH obtained with pCZ145.

To demonstrate that we were observing translational and not transcriptional effects, we measured the steady-state levels of the two-cistron mRNA in high and low producers of Met-[Ala]bGH by dot blot analysis. As shown in Fig. 5, intense hybridization with no more than 3-fold differences are seen with RNA transcribed from these plasmids. The highest and lowest amounts of hybridizing mRNA were found in cultures containing pCZ143 (a low producer of Met-[Ala]bGH) and pCZ140 (the plasmid without the first cistron), respectively. Intermediate levels of hybridizing mRNA were found in cultures that overproduce Met-[Ala]bGH. These measurements indicate that the two-cistron mRNAs are efficiently transcribed and confirmed our assumption that the large differences in Met-[Ala]bGH expression are due to differences in the translatability of these mRNAs. The 3-fold differences in the steady-state mRNA levels probably reflect differences in the mRNA turnover.

Expression of Other Mammalian Genes with the Two-Cistron System. From our studies, some rules have emerged (as outlined above) concerning the expression of Met-[Ala]bGH with a synthetic two-cistron system. In the next series of experiments, we examined whether our findings are applicable to the expression of other mammalian genes. We also wanted to determine whether the presence of a first cistron affects the expression level of genes that are not translationally blocked and therefore can be expressed without the use of a first cistron. An example of such a gene is the one that encodes the 20-kDa variant of hGH (44). The coding sequence for this 20-kDa methionyl hGH (Met-hGH) was inserted into plasmids pCZ140, pCZ143, and pCZ145 by replacing precisely the coding sequence for bGH. These new plasmids (pCZ340, pCZ343, and pCZ345, respectively) were transformed into RV308, and Met-hGH production was measured by $NaDodSO_4$/polyacrylamide gel analysis and scanning. The results are shown in Fig. 3 (lanes 11, 12, and 13). Lane 11 shows proteins from cell extracts containing plasmid pCZ340. The prominent band represents Met-hGH and constitutes 22% of the total cell protein. Note that pCZ340 was derived from and is identical to pCZ140 (except for the growth hormone gene) and thus has no first cistron sequence. We found that cells containing pCZ343 produce Met-hGH at less than 1% of total cell protein (Fig. 3, lane 12), and cells containing pCZ345 produce Met-hGH at more than 20% of total cell protein. Thus, placement of a "nonfunctional" cistron (as in pCZ343) in front of the Met-hGH coding sequence considerably reduces the amount of growth hormone produced. However, placement of a "functional" first cistron in front of the Met-hGH coding sequence (as in pCZ345) does not diminish the amount of protein produced, even though the cistron is not required for high-level expression. Similar results were obtained from studies with other mammalian genes including derivatives of tissue plasminogen activator, human protein C (B.S. and J. Hoskins, unpublished data) and human proinsulin.*

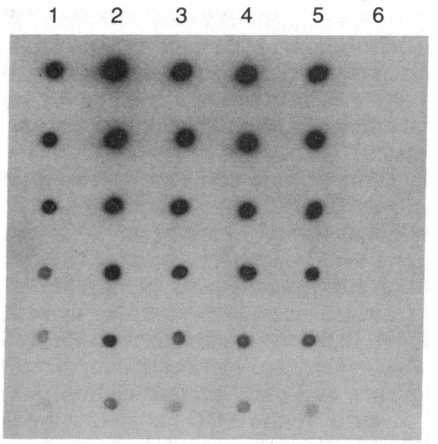

FIG. 5. RNA dot blot analysis of bGH mRNA in *E. coli*. Total cellular RNA (1 μg) was spotted in the top row of each lane on a sheet of nitrocellulose paper. Each row contains a serial dilution (1:3) of that RNA. The RNA was isolated from cells harboring the following plasmids: lane 1, pCZ140; lane 2, pCZ143; lane 3, pCZ144; lane 4, pCZ145; lane 5, pCZ182; and lane 6, pIMIA.

DISCUSSION

A typical *E. coli* ribosome binding site is about 40 nucleotides in length (45–47) and includes two highly conserved elements, the SD sequence and a properly spaced translation initiation codon (10–12). The sequences outside this region are more variable and are at least partially responsible for the vast differences seen in the rate of translation initiation with different mRNAs (11, 18, 19). The precise mechanism by which these sequences exert their effect is presently under investigation in several laboratories. Hui *et al.* (13) have argued for a direct involvement of the primary sequence in translation initiation, based on systematic changes that they introduced into the 5′ untranslated region of the mRNA. Studies by Stanssens *et al.* (17) suggest that the efficiency of translation initiation can also be affected indirectly by secondary structure formation between sequences in the coding region and the 5′ untranslated region. Because the coding sequences are different for different genes, the design of

*Shrote, J. S., Schoner, B. E., Belagaje, R. M. & Schoner, R. G., Annual Meeting of the American Society for Microbiology, March 23–28, 1986, Washington, DC, p. 134 (abstr.).

expression vectors must take these potential interactions into account if such vectors are to have general applicability. In view of this, we have explored the use of a two-cistron expression system as a general method to improve the translational efficiency of mRNAs and to achieve high-level gene expression.

Our initial assumption in designing the two-cistron expression system was that ribosome binding is favored when the SD sequence and the AUG are in an A+U-rich sequence context free of local secondary structure (10). Once bound, the translating ribosomes should be capable of disrupting secondary structures located downstream from the initiation site (25, 48, 49). In the two-cistron expression system, a short sequence (the first cistron) was inserted in front of the protein coding sequence (i.e., the second cistron) such that the sequences 5′ and 3′ to the AUG codon (for the first cistron) can be optimized for translation initiation without the constraints of preserving the coding information. Since the first cistron separates the 5′ untranslated sequence from the protein coding region, the opportunity for local secondary structure formation should be minimized.

We further assumed, based on translation of natural polycistronic mRNAs in *E. coli*, that ribosomes bound at the first cistron would initiate translation, proceed through the first cistron, and continue into the second cistron. With the synthetic two-cistron expression system, efficient readthrough translation into the second cistron was observed only when the second cistron had a SD sequence and when the stop codon that terminates translation of the first cistron lay 3′ to that SD sequence (SD 2). These observations are consistent with studies by Das and Yanofsky (50), who showed that at the *trpB–trpA* junction, relocation of the termination codon for *trpB* and alterations in the SD sequence for *trpA* reduce translation of *trpA*. Our data further suggest that the spacing between this SD 2 and the AUG codon for Met 2 can additionally affect the efficiency of readthrough translation into the second cistron.

Placement of the original first-cistron sequence in front of the coding sequence for Met-[Ala]bGH (as in pCZ143) resulted in very low expression of Met-[Ala]bGH, presumably because the ribosomes that are bound at the first cistron fail to continue translation into the second cistron. This may occur if the ribosomes terminate and reinitiate translation of the first cistron without dissociating from the mRNA or if the ribosomes stall in the first cistron. Since the first cistron is relatively short, the signals for translation initiation, elongation, and termination are within a span of less than 40 nucleotides, the average size of an *E. coli* ribosome binding site. In pCZ144 and pCZ145 (and in several of the other plasmids), where the stop codon in the first cistron lies 3′ to the SD 2 and close to the restart codon for the second cistron, the interaction between the 16S rRNA and the SD 2 may be strong enough to favor readthrough translation over reinitiation at the first cistron initiation site. Thus, in a two-cistron expression system, where the first cistron is short, the position of the stop codon relative to the SD sequence and the restart codon may become very important in determining the level of expression of the downstream cistron.

We thank Richard Jaskunas, Ignacio Larrinua, and John Wood for helpful discussions and for critically reading the manuscript, Stan Ly for excellent technical assistance, and Cheryl Alexander for preparing the manuscript. We also thank Richard Baltz and J. Paul Burnett for their continued support.

1. Masui, Y., Coleman, J. & Inouye, M. (1983) in *Experimental Manipulation of Gene Expression*, ed. Inouye, M. (Academic, New York), pp. 15–32.
2. Crowl, R., Seamans, C., Lomedico, P. & McAndrew, S. (1985) *Gene* **38**, 31–38.
3. Emerick, A. W., Bertolani, B. L., Ben-Bassat, A., White, T. J. & Konrad, M. W. (1984) *Bio/Technology* **2**, 165–168.
4. Ray, P. N. & Pearson, M. L. (1974) *J. Mol. Biol.* **85**, 163–175.
5. Ray, P. N. & Pearson, M. L. (1975) *Nature (London)* **253**, 647–650.
6. Kelley, R. L. & Yanofsky, C. (1982) *Proc. Natl. Acad. Sci. USA* **79**, 3120–3124.
7. Nagai, K. & Thogersen, H. C. (1984) *Nature (London)* **309**, 810–812.
8. Varadarajan, R., Szabo, A. & Boxer, S. G. (1985) *Proc. Natl. Acad. Sci. USA* **82**, 5681–5684.
9. Shine, J. & Dalgarno, L. (1975) *Nature (London)* **254**, 34–38.
10. Gold, L., Pribnow, D., Schneider, T., Shinedling, S., Singer, B. S. & Stormo, G. (1981) *Annu. Rev. Microbiol.* **35**, 365–403.
11. Stormo, G. D., Schneider, T. D. & Gold, L. M. (1982) *Nucleic Acids Res.* **10**, 2971–2996.
12. Kozak, M. (1983) *Microbiol. Rev.* **47**, 1–45.
13. Hui, A., Hayflick, J., Dinkelspiel, K. & de Boer, H. A. (1984) *EMBO J.* **3**, 623–629.
14. Shepard, M. G., Yelverton, E. & Goeddel, D. V. (1982) *DNA* **1**, 125–131.
15. de Boer, H. A., Hui, A., Comstock, L. J., Wong, E. & Vasser, M. (1983) *DNA* **2**, 231–235.
16. Whitehorn, E. A., Livak, K. J. & Petteway, S. R., Jr. (1985) *Gene* **36**, 375–379.
17. Stanssens, P., Remaut, E. & Fiers, W. (1985) *Gene* **36**, 211–223.
18. Roberts, T. M., Kacich, R. & Ptashne, M. (1979) *Proc. Natl. Acad. Sci. USA* **76**, 760–764.
19. Gold, L., Stormo, G. & Saunders, R. (1984) *Proc. Natl. Acad. Sci. USA* **81**, 7061–7065.
20. Gheysen, D., Iserentant, D., Derom, C. & Fiers, W. (1982) *Gene* **17**, 55–63.
21. Iserentant, D. & Fiers, W. (1980) *Gene* **9**, 1–12.
22. Schwartz, M., Roa, M. & Debarbouille, M. (1981) *Proc. Natl. Acad. Sci. USA* **78**, 2937–2941.
23. Hall, M. N., Gabay, J., Debarbouille, M. & Schwartz, M. (1982) *Nature (London)* **295**, 616–618.
24. Das, A., Urbanowski, J., Weissbach, H., Nestor, J. & Yanofsky, C. (1983) *Proc. Natl. Acad. Sci. USA* **80**, 2879–2883.
25. Berkhout, B. & van Duin, J. (1985) *Nucleic Acids Res.* **13**, 6955–6967.
26. George, H. J., L'Italien, J. J., Pilacinski, W. P., Glassman, D. L. & Krzyzek, R. A. (1985) *DNA* **4**, 273–281.
27. Seeburg, P. H., Sias, S., Adelman, J., de Boer, H. A., Hayflick, J., Jhurani, P., Goeddel, D. V. & Heyneker, H. L. (1983) *DNA* **2**, 37–45.
28. Schoner, B. E., Hsiung, H. M., Belagaje, R. M., Mayne, N. G. & Schoner, R. G. (1984) *Proc. Natl. Acad. Sci. USA* **81**, 5403–5407.
29. Maurer, R., Meyer, B. J. & Ptashne, M. (1980) *J. Mol. Biol.* **139**, 147–161.
30. Gait, M. J. & Sheppard, R. C. (1980) *Nucleic Acids Res.* **8**, 1081–1096.
31. Brown, E. L., Belagaje, R., Ryan, M. J. & Khorana, H. G. (1979) *Methods Enzymol.* **68**, 109–151.
32. Danner, D. B. (1982) *Anal. Biochem.* **125**, 139–142.
33. Wensink, P. C., Finnegan, D. J., Donelson, J. E. & Hogness, D. S. (1974) *Cell* **3**, 315–325.
34. Birnboim, H. C. & Doly, J. (1979) *Nucleic Acids Res.* **7**, 1513–1523.
35. Maxam, A. M. & Gilbert, W. (1980) *Methods Enzymol.* **65**, 499–560.
36. Laemmli, U. K. (1970) *Nature (London)* **227**, 680–685.
37. Young, F. S. & Furano, A. V. (1981) *Cell* **24**, 695–706.
38. Thomas, P. S. (1983) *Methods Enzymol.* **100**, 255–265.
39. Schoner, B. & Schoner, R. G. (1981) *Gene* **16**, 347–352.
40. Nakamura, K., Pirtle, R. M., Pirtle, I. L., Takeishi, K. & Inouye, M. (1980) *J. Biol. Chem.* **255**, 210–216.
41. Schoner, B. E., Belagaje, R. M. & Schoner, R. G. (1987) *Methods Enzymol.*, in press.
42. Li, C. H. & Ash, L. (1953) *J. Biol. Chem.* **203**, 419–424.
43. Wallis, M. & Davies, R. V. (1976) in *Growth Hormone and Related Peptides*, eds. Pecile, A. & Muller, E. E. (Elsevier, Holland), pp. 1–13.
44. Lewis, U. J., Bonewald, L. F. & Lewis, L. J. (1980) *Biochem. Biophys. Res. Commun.* **92**, 511–516.
45. Steitz, J. A. (1969) *Nature (London)* **224**, 957–964.
46. Hindley, J. & Staples, D. H. (1969) *Nature (London)* **224**, 964–967.
47. Gupta, S. L., Chen, J., Schaefer, L., Lengyel, P. & Weissman, S. M. (1970) *Biochem. Biophys. Res. Commun.* **39**, 883–888.
48. Kastelein, R. A., Berkhout, B. & van Duin, J. (1983) *Nature (London)* **305**, 741–743.
49. Baughman, G. & Nomura, M. (1983) *Cell* **34**, 979–988.
50. Das, A. & Yanofsky, C. (1984) *Nucleic Acids Res.* **12**, 4757–4768.