Gene Expression, Vol. 8, pp. 141–149, 1999
Printed in the USA.

1052-2166/99 $20.00 + .00

The T7 Concatemer Junction Sequence Interferes With Expression From a Downstream T7 Promoter In Vivo

BOHDAN HARVEY,* MALGORZATA KORUS,† AND EMANUEL GOLDMAN*†[1]

Graduate School of Biomedical Sciences, †New Jersey Medical School, Department of Microbiology & Molecular Genetics, University of Medicine & Dentistry of New Jersey, Newark, NJ

A recently described new signal for transcription termination in vitro by T7 RNA polymerase has now been tested in vivo. This signal, identified during transcription of the cloned human preproparathyroid hormone (PTH) gene, is also found in the phage T7 genome, at the concatemer junction (CJ). We introduced the 17-bp concatemer junction sequence at the ends of a test gene and control gene (both derived from T7 gene *9*) in a T7 vector previously used to study effects of rare codons on expression. The CJ elements replaced the original vector's RNase III processing sites, and a new T7 promoter was also introduced to drive the downstream (control) gene. We assayed for test and control gene mRNA and protein by direct labeling with [^{32}P]phosphate and [^{35}S]methionine. The altered vector with CJ sequences (pCT1.1) expressed the upstream test gene, but showed poor expression of the downstream control gene. No discrete T7 mRNA bands could be discerned by direct labeling with ^{32}P. A precursor vector with only the control gene in single copy expressed the protein much better, suggesting that the inhibition of control gene expression in pCT1.1 was a result of the upstream CJ element at the 3′ end of the test gene. RT-PCR experiments were consistent with readthrough and possibly pausing at CJ. An RNA folding program predicts a highly stable secondary structure between the upstream CJ element and the control gene's translation start signals. These data support an interpretation that the CJ element is ineffective as a T7 transcription terminator in vivo in this vector, and that structure of the readthrough transcript blocks ribosome access to the downstream translation start. The readthrough transcripts are also likely to be less stable than properly terminated or processed T7 mRNA, because levels of test protein expression in pCT1.1 were reduced compared to original vector, and basal expression was negligible, while the original codon test vector shows substantial basal expression.

Concatemer junction T7 promoter Preproparathyroid hormone

A new class of signal for transcription termination in vitro by T7 RNA polymerase has recently been characterized (8,10,11), based upon the finding that a sequence in the cloned human preproparathyroid hormone (PTH) led to transcription termination by T7 RNA polymerase in vitro. A 7-bp core sequence (ATCTGTT in the sense strand) appears to be responsible, and this same core sequence is also found in the phage T7 genome, at the concatemer junction (CJ) (2). Although correlations to properties of mutant T7 RNA polymerases have been noted (10,11,17,18), this new termination motif has not yet, to our knowledge, been specifically tested in vivo. Further, the T7 concatemer junction sequence has been suggested to cause the T7 RNA polymerase to pause rather than terminate in vivo (10,18), and an unpublished personal communication from D. Mead, as cited in He et al. (8), indicates that this sequence facilitates transcription termination on linear rather than supercoiled templates.

We were interested in investigating expression from T7 vectors in which the encoded mRNAs were less stable than messages transcribed from currently available T7 constructs, whose mRNAs possess 3′

[1]Address correspondence to Emanuel Goldman, Department of Microbiology & Molecular Genetics, University of Medicine & Dentistry of New Jersey, 185 South Orange Avenue, Newark, NJ 07103. Tel: (973) 972-4367; Fax: (973) 972-3644; E-mail: egoldman@umdnj.edu

stem–loop structures. To this end, we introduced 17 bp from the T7 CJ sequence (nt 160–176 in the T7 genome, TATCTGTTACAGTCTCC) at the ends of two genes (test and control genes, both derived from T7 gene *9*) in a T7 vector previously developed to study effects of rare codons on expression (15). The CJ elements replaced the original vector's RNase III processing sites, and on the assumption that the CJ element would facilitate transcription termination, a new T7 promoter was also introduced to drive the downstream (control) gene. Unexpectedly, we have found that the CJ element at the end of the upstream (test) gene significantly depresses expression from the downstream control gene. However, precursor plasmids with CJ at the 3′ ends of the inserts, containing only single copies of either control gene or test gene, express their respective proteins equally well. We assayed for test and control gene mRNA and protein by direct labeling with [^{32}P]phosphate and [^{35}S]methionine, which can resolve both mRNA and protein in a single gel electrophoresis in this system (3). However, we were unable to detect discrete T7 mRNA bands in the constructs with CJ elements. More sensitive RT-PCR experiments suggested readthrough and possibly pausing at CJ. An RNA folding program predicts that the control gene translation start signals may be occluded in secondary structure with the CJ transcript. These data support a model in which the CJ element is ineffective as a transcription terminator in vivo in this system, and that the readthrough transcript from the upstream gene hinders translation starts from the downstream gene. The readthrough transcript is also likely to be less stable than properly terminated or processed T7 transcripts, because levels of test protein expression were reduced compared to original vector, and basal expression was negligible, while the original codon test vector shows substantial basal expression.

MATERIALS AND METHODS

Construction of pCT1.1

The vector into which we introduced CJ sequences was based upon vector pCT1 (15), previously used to study effects of rare codons on expression (4,7,15). With the exception of the CJ element, all the components of the vector were assembled, as described in Rosenberg et al. (15), from pET vectors and plasmids carrying elements of T7 DNA, generously provided by Alan Rosenberg (Brookhaven National Laboratory). The CJ element was synthesized by the NJ Medical School Molecular Resource Facility as two complementary single-stranded oligonucleotides, with an *Nhe*I site at the 5′ end (in the sense strand) and a *Bgl*II site at the 3′ end. The oligos were annealed and inserted into precursor vectors precut with the same enzymes. Colony PCR was used to screen transformants for the presence of CJ sequences. The final vector was placed in *E. coli* strain Xac for maintenance, and inserted into strain BL27(DE3)pLysS (15) for expression experiments. The plasmid, called pCT1.1, was sequenced from the maintenance strain by the NJ Medical School Molecular Resource Facility, and all junctions, sites, and much of the remaining sequence were confirmed. Elements and design of the test and control genes of pCT1.1 are shown in Fig. 1.

Precursor vectors pTG and pCG contained only the T7 promoter, either the test gene (pTG) or control gene (pCG), followed by the CJ sequence (Fig. 1).

Measurements of mRNA and Protein Expression

Tests were carried out in a T7 expression system, in which the host cell contains the gene for T7 RNA polymerase integrated into the chromosome, under control of the *lac*UV5 promoter. Upon addition of IPTG, the host RNA polymerase transcribes the T7 RNA polymerase gene, and in turn, the T7 RNA polymerase transcribes any target genes in the cell under control of T7 promoters, such as the test and control genes on our plasmids. Because *lac* repression is not absolute, there is basal level transcription of the T7 RNA polymerase at all times; in order to minimize this basal activity of the T7 RNA polymerase (so that the target genes are repressed), a compatible plasmid is present in the cell, pLysS, which makes a moderate amount of T7 lysozyme, an inhibitor of T7 RNA polymerase. Only upon induction is there sufficient excess of the T7 RNA polymerase to significantly amplify transcription of the T7 target genes. T7 lysozyme has also been noted to potentiate transcription termination in vitro by the PTH/CJ class of termination signals (10).

Labeling and detection of mRNA and protein in vivo followed the procedures described by Gao and Goldman (3). Briefly, cells were grown at 37°C in 3-[*N*-morpholino]propanesulfonic acid (MOPS) minimal medium (a low-phosphate medium) with 50 μg/ml ampicillin and 25 μg/ml chloramphenicol to mid-log (2×10^8 cells/ml) and induced with 1 mM IPTG. Ten minutes after induction, ~50 μCi/ml [^{32}P]phosphoric acid (Amersham) was added to the cultures. Some cultures also had 100 μg/ml rifampicin added with the ^{32}P (which shuts off host transcription while allowing T7 transcription to continue). Thirty minutes after induction, 40 mM (nonradioactive) sodium phosphate, pH 7.4, was added as a chase. Aliquots (50 μl) were labeled at the indicated times with 20

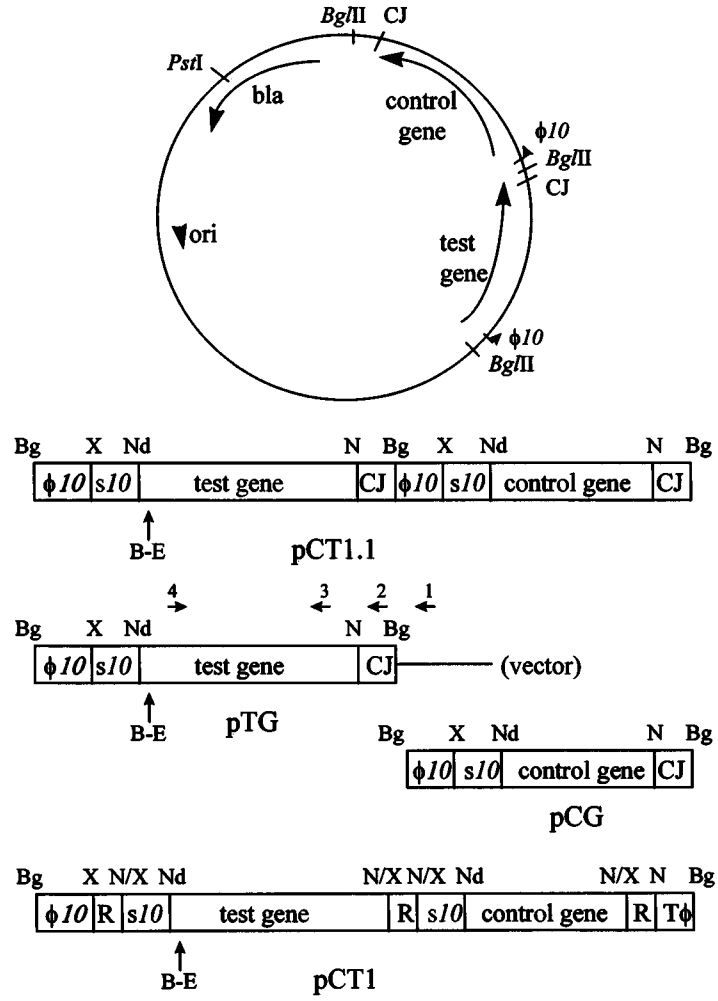

FIG. 1. Arrangement of elements in vector pCT1.1. *ϕ10* represents a promoter for T7 RNA polymerase, *s10* is a strong translational start site, and CJ is a 17-bp sequence from the T7 concatemer junction. The test and control genes are derived from T7 gene *9*, with a 45-codon in-frame deletion from the test gene to derive the control gene. A *Bam*HI-*Eco*RI site in the test gene allows insertion of nucleotides after codon 13. pTG and pCG are the single gene copy precursor plasmids used to make pCT1.1. Shown above pTG are the relative locations and orientations of four primers used for RT-PCR (see Materials and Methods). Also shown is the arrangement of elements in parental vector pCT1 (15). R represents an RNase III cleavage site, and Tϕ is a transcriptional terminator for T7 RNA polymerase. *bla* represents the plasmid β-lactamase gene, and *ori* represents the plasmid's origin of replication. Restriction sites: Bg, *Bgl*II; X, *Xba*I; Nd, *Nde*I; N, *Nhe*I; N/X, *Nhe*I-*Xba*I fusion; B-E, *Bam*HI-*Eco*RI cloning site.

μCi/ml, 50 Ci/mmol [^{35}S]methionine (Amersham) for 2 min. Samples were stopped with 25 μl 3× MOPS cracking buffer [3× MOPS cracking buffer contains 60 mM MOPS, 24 mM sodium acetate, 3% SDS, 6 mM EDTA, 30% glycerol, and 0.15% each of bromophenol blue and xylene cyanole FF, and is stored at room temperature; 3% (v/v) β-mercaptoethanol is added to complete the reagent immediately before use], boiled in capped tubes for 2 min, and subjected to electrophoresis on sodium-dodecyl-sulfate polyacrylamide gels (SDS-PAGE) in which the 5% top gel was of approximately equal length to the 10% bottom gel, so that RNA species (up to ~2 kB) could be resolved in the top and protein species in the bottom, as described (3). The dried gel was subjected to autoradiography.

Assay for RNA by Reverse Transcriptase-Polymerase Chain Reaction (RT-PCR)

Cells containing pTG were grown and induced for 1 h as described in the previous section (with the exception that rifampicin was added 30 min after induction), pelleted, and concentrated in a small volume of 1× MOPS modified cracking buffer [20 mM MOPS, 0.2 M sodium acetate (pH 5.5), 1% SDS, 2 mM EDTA, 1% β-mercaptoethanol]. Samples were boiled for 2 min in capped tubes, frozen in a dry ice-metha-

nol bath, thawed, and extracted with phenol-chloroform. The aqueous layer was precipitated and washed with 70% ethanol, dried in a speed-vac, and the pellet dissolved in sterile water. The sample was treated with 40 μg/ml DNase (RNase-free; Worthington) for 30 min at 37°C in 40 mM Tris-HCl, pH 7.5, 6 mM $MgCl_2$, 2 mM $CaCl_2$. Sodium acetate, pH 5.5, was added to 0.3 M and samples were again extracted with phenol-$CHCl_3$ and precipitated in ethanol, as above. The pellet was dissolved in water and stored at −80°C.

RT-PCR primer locations and orientations are indicated in Fig. 1, above pTG. Reverse transcription used either primer 1 (GGGTTCCGCGCACATTT CCC), to measure readthrough transcription into the vector, or primer 2 (GAGACTGTAACAGATA), which is complementary to CJ. Primer 1 is located about 100 nt downstream of CJ, and about 100 nt upstream of the translation start of the vector's β-lactamase gene. PCR amplification of the cDNA used primers 3 (ATCATCTCGCTACGGTCCGC) and 4 (TCCGCTGTGATGTCTGGTGG), both of which are located about 30 codons from the respective ends of the test gene. RT and PCR reactions were carried out using reagents from Gibco BRL, following the manufacturer's protocols.

RESULTS AND DISCUSSION

Inhibition of pCT1.1 Control Gene Expression In Vivo

Upon completion of construction of pCT1.1, and confirmation of the sequence, the plasmid was placed in an expression host, and assayed for production of test and control mRNAs and proteins, in parallel to cells containing the parental plasmid, pCT1. The differences between the plasmids are that the three RNase III processing sites of pCT1 (flanking and between the test and control genes) are absent in pCT1.1, with the RNase III sites at the 3′ ends of the test and control genes substituted by CJ elements. The T7 transcription terminator (Tϕ) at the 3′ end of the pCT1 transcription unit is also absent. Finally, just downstream of the CJ element at the 3′ end of the test gene, an additional T7 promoter (with a *Bgl*II site at the 5′ end) has been introduced in pCT1.1 (see Fig. 1 for comparison of the two vectors).

Cells were grown and the T7 system induced as described in Materials and Methods. Samples were subjected to SDS-PAGE under conditions that allow detection of both mRNA and protein (3). Uninduced pCT1 (lane 1) and pCT1.1 (lane 5) showed little detectable test and control mRNAs after 20-min labeling, and low levels of test and control proteins in a 2-min pulse (Fig. 2). Following induction for varying lengths of time, discrete mRNA bands for test and control messages were evident in pCT1 (lanes 2–4) but not in pCT1.1 (lanes 6–8). In these samples, ^{32}P was added 10 min after induction, and an excess of nonradioactive phosphate was added 30 min after induction to chase the ^{32}P label. Lanes 2 and 6 show incorporation after the initial 30-min labeling, while the other lanes were chased for 15 (lanes 3 and 7) or 30 (lanes 4 and 8) min (i.e., 45 or 60 min after induction, respectively). Some of the ^{32}P label appeared to chase into a lower molecular weight band in pCT1 (lanes 3 and 4) and also in pCT1.1 (lane 8). Samples were also labeled for 2 min with [^{35}S]methionine at 30 (lanes 2 and 6), 45 (lanes 3 and 7), and 60 (lanes 4 and 8) min after induction. While, as expected, there was good synthesis of both test and control proteins in pCT1 (lanes 2–4), only the test protein appeared to be synthesized in substantial quantity in pCT1.1 (lanes 6–8).

Portions of the cultures had rifampicin added at the same time they were labeled with ^{32}P (Fig. 2, lanes 9–12 for pCT1 and lanes 13–16 for pCT1.1); otherwise, these samples were treated identically to those without rifampicin. After a 20-min rifampicin treatment, pCT1 showed substantial synthesis of test and control proteins even without induction (lane 9) whereas, without induction, pCT1.1 showed only barely detectable synthesis of test protein (lane 13). Thus, even the basal level of expression of test protein was significantly reduced by placing CJ at the 3′ end of its message. Upon induction, samples in the presence of rifampicin showed a similar pattern of RNA synthesis as seen in the samples without rifampicin, except that the lower molecular weight chased band showed up a bit earlier in pCT1.1 (lane 15). This band most likely represents a stable degradation product of T7 mRNA, because it appears after chase with nonradioactive phosphate in untreated and rifampicin-treated samples of both pCT1 and pCT1.1. The slightly higher molecular weight induced RNA band seen in pCT1.1 (lanes 6–8) is apparently not T7 specific, because it is not synthesized in the presence of rifampicin (lanes 14–16). This band could reflect host synthesis, because pCT1.1 does not appear to effectively shut off the host compared to standard T7 expression vectors. The induced protein expression in the rifampicin-treated samples clearly shows that pCT1 synthesizes both test and control proteins in roughly equal amounts (lanes 10–12), while pCT1.1 synthesizes test protein but much less control protein (lanes 14–16).

These samples were also run on a 1.5% formaldehyde-agarose gel (6); however, no discrete higher molecular weight bands except near the origin could

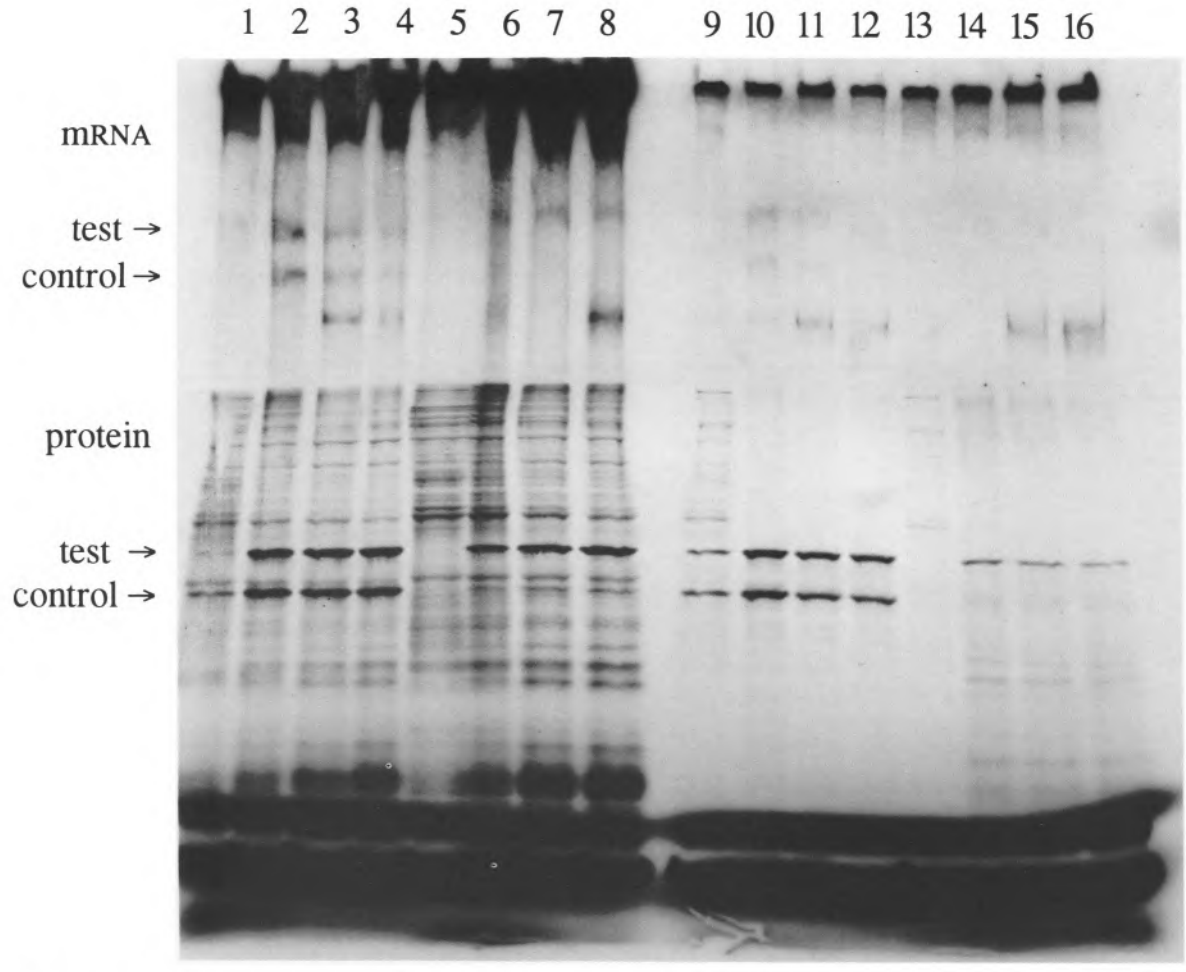

FIG. 2. Test and control RNA and protein synthesis in cells containing pCT1.1 and pCT1. Cultures were grown, labeled, and subjected to SDS-PAGE as described in Materials and Methods. pCT1: lanes 1–4 and 9–12; pCT1.1: lanes 5–8 and 13–16. Uninduced cultures were labeled for 20 min with [^{32}P]phosphate and 2 min with [^{35}S]methionine in the absence (lanes 1 and 5) or presence (lanes 9 and 13) of 100 μg/ml rifampicin (added at the same time as the ^{32}P label). All the other cultures were induced with IPTG at time 0, labeled with [^{32}P]phosphate at 10 min in the absence (lanes 2–4 and 6–8) or presence (lanes 10–12 and 14–16) of rifampicin (added at the same time as the ^{32}P label), and chased with sodium phosphate at 30 min. Samples (50 μl) were labeled for 2 min with [^{35}S]methionine at 30 min (lanes 2, 6, 10, 14), 45 min (lanes 3, 7, 11, 15), or 60 min (lanes 4, 8, 12, 16) postinduction. (Note that the 45- and 60-min samples were chased with nonradioactive phosphate starting at 30 min postinduction.) The dried gel was exposed to X-ray film for approximately 2 days.

be discerned by autoradiography of the dried gel (not shown). These results are consistent with the absence of specific transcription termination at the CJ elements, in which case the T7 RNA polymerase is expected to continue around the plasmid many times (13,16), presumably terminating randomly. The putative giant mRNAs, however, are probably less stable than properly terminated (or processed) T7 messages, because both basal and induced expression of test protein are lower in pCT1.1 compared to pCT1 (e.g., compare lanes 13–16 with lanes 9–12). Quantitation of the protein bands from lanes 9–16 in a phosphorimager (not shown) indicates approximately fourfold lower expression of test protein in pCT1.1 compared to a parallel culture of pCT1. By contrast, RNase III processed T7 mRNAs, such as those in pCT1, are known to be highly stable (14,15).

Control Gene Expresses Much Better in Single Copy on the Plasmid

The low expression of control protein seen in Fig. 2 could have been either a result of introduction of the upstream CJ sequence, or because the control gene as cloned was intrinsically defective in expression (e.g., the CJ sequence at the 3′ end of the control gene might somehow interfere with its expression). To distinguish between these possibilities, we tested expression of the precursor plasmids used to generate pCT1.1.

Plasmids containing a T7 promoter, either test gene (pTG) or control gene (pCG), and the CJ sequence (see Fig. 1) were inserted into the expression host, and similar experiments as those described for pCT1.1 were carried out. In Fig. 3, lane 1 is a marker of expression with pCT1, and lane 10 is a marker of expression with pCT1.1. Lanes 2–5 show expression with pTG (test gene only), while lanes 6–9 show expression with pCG (control gene only). Uninduced samples (lanes 2 and 6) showed little expression of test or control protein, while induced samples showed good expression of either test (lanes 3–5) or control (lanes 7–9) proteins. The same experiment was carried out in the presence of rifampicin, which again showed synthesis of either test (lanes 12–14) or control (lanes 16–18) proteins. As was the case with pCT1.1, there were no discrete mRNA bands in any of these samples except for marker pCT1 (lane 1). Thus, poor expression of the control gene in pCT1.1 is not due to an intrinsic difficulty with the gene as cloned in this vector, but is due to the configuration of the gene downstream of the CJ element at the 3′ end of the test gene.

Evidence for Readthrough and Possible Pausing at CJ Element

We used reverse transcriptase to copy RNA from induced cells containing the test gene vector (pTG), followed by PCR amplification, with the primers diagramed in Fig. 1. Starting the reverse transcription from primer 1 downstream of CJ, we were able to obtain a product that was amplified by PCR to the expected size with primers 3 and 4 (Fig. 4, lane 2). This indicates that the T7 RNA polymerase did in fact read past the CJ sequence. However, from the same RNA preparation, there was even greater amplification of the signal (Fig. 4, lane 3) when the reverse transcription reaction started from primer 2, complementary to CJ, suggesting a greater abundance of mRNA containing CJ compared to mRNA containing sequences from further downstream. RNA from cells without pTG did not amplify (Fig. 4, lanes 4 and 5), and the PCR products in lanes 2 and 3 are not the result of contaminating DNA (lane 9). These results could reflect paused transcription at CJ followed by readthrough. A caveat to this suggestion is that the

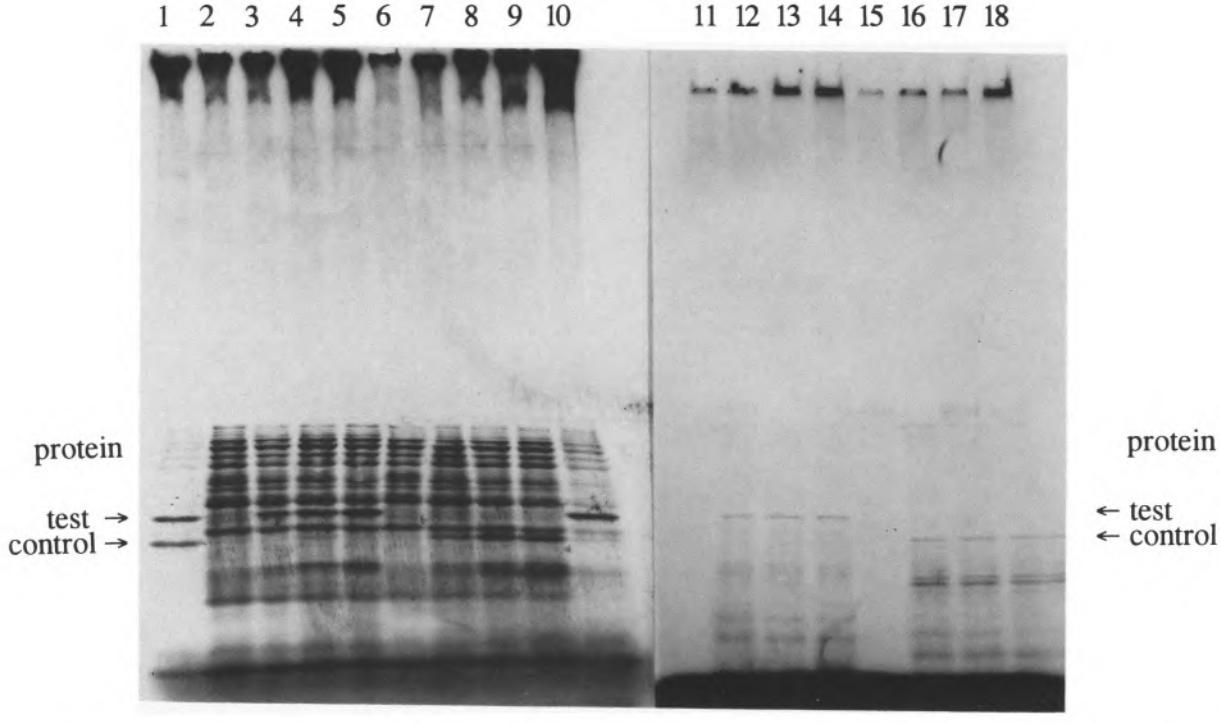

FIG. 3. Test and control RNA and protein synthesis in cells containing precursor vectors. Cells containing pTG or pCG (precursor vectors in the construction of pCT1.1; see Fig. 1) were grown, labeled, and subjected to SDS-PAGE as described in Materials and Methods. Lane 1 is a pCT1 marker (composed of a mixture from samples 2–4 in Fig. 2) and lane 10 is a pCT1.1 marker (composed of a mixture from samples 6–8 in Fig. 2). pTG: lanes 2–5 and 11–14; pCG: lanes 6–9 and 15–18. Uninduced cultures were labeled for 20 min with [^{32}P]phosphate and 2 min with [^{35}S]methionine in the absence (lanes 2 and 6) or presence (lanes 11 and 15) of 100 μg/ml rifampicin (added at the same time as the ^{32}P label). All the other cultures were induced with IPTG at time 0, labeled with [^{32}P]phosphate at 10 min in the absence (lanes 3–5 and 7–9) or presence (lanes 12–14 and 16–18) of rifampicin (added at the same time as the ^{32}P label), and chased with sodium phosphate at 30 min. Samples (50 μl) were labeled for 2 min with [^{35}S]methionine at 30 min (lanes 3, 7, 12, 16), 45 min (lanes 4, 8, 13, 17), or 60 min (lanes 5, 9, 14, 18) postinduction. The dried gels were exposed to X-ray film for approximately 2 days.

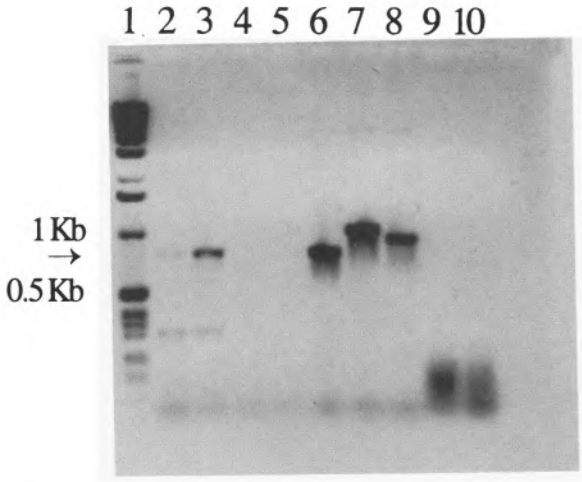

FIG. 4. RT-PCR of RNA from cells with and without pTG. Cells were grown, induced for 1 h (with rifampicin added 30 min postinduction), and RNA was extracted and subjected to reverse transcription (RT) and PCR amplification as described in Materials and Methods. See Fig. 1 for location and orientation of primers. Lane 1: 1 kb DNA ladder (Gibco BRL); lane 2: pTG RNA, RT reaction with primer 1; lane 3: pTG RNA, RT reaction with primer 2; lane 4: BL27(DE3)pLysS RNA, RT reaction with primer 1; lane 5: BL27(DE3)pLysS RNA, RT reaction with primer 2; lane 6: pTG plasmid DNA, PCR reaction with primers 3 and 4; lane 7: pTG plasmid DNA, PCR reaction with primers 1 and 4; lane 8: pTG plasmid DNA, PCR reaction with primers 2 and 4; lane 9: pTG RNA, no RT reaction; lane 10: BL27(DE3)pLysS RNA, no RT reaction. PCR reactions for lanes 2–5, 9, and 10 all used primers 3 and 4. Arrow: PCR amplified fragment from primers 3 and 4. Samples were subjected to electrophoresis on a 1% agarose gel, stained with ethidium bromide, and photographed under ultraviolet light.

efficiency of priming the reverse transcription reaction with primer 1 could have been significantly lower than that of primer 2, resulting in a greater quantity of cDNA template from primer 2. However, we note that primer 1 had a much higher G-C content (13 out of 20 nt) compared to primer 2 (6 out of 16 nt); therefore, one might expect primer 1 to be the more efficient primer. Indeed, primer 1 was somewhat more efficient than primer 2 in directly priming PCR reactions with primer 4 on plasmid DNA (compare lanes 7 and 8 in Fig. 4).

A Folded RNA Structure Suggests Possible Translational Inhibition of Control Gene

Having established that there is nothing intrinsically wrong with the control gene construct, the question remains as to why its expression is inhibited in pCT1.1. After all, the control gene in this construct has a perfectly good T7 promoter to drive its expression. Three obvious possibilities could explain these data. The first is that failure of CJ to act as a transcription terminator allows continued readthrough by elongating T7 RNA polymerase, which prevents access for an initiating polymerase to find the downstream promoter for the control gene. However, to our knowledge there is no precedent of promoter interference by transcriptional readthrough; further, if this occurred, we would expect it would have already been noted in studies of T7 morphogenesis.

A second possibility is that if the CJ element is acting as a pause site for T7 RNA polymerase, the paused enzyme blocks access for another polymerase molecule to initiate at the downstream promoter, which is only about 20 nt away. Arguing against this explanation is that we did not see an mRNA band for the test gene in our gels, suggesting that this putative pause site is not particularly effective. Evidence suggesting pausing comes from a more sensitive RT-PCR assay.

The third possibility is that the readthrough transcript from the upstream test gene through the control gene has a superstructure that occludes the control gene translation start site, so that ribosomes cannot easily initiate translation. There is ample precedent for this kind of mechanism regulating translation [e.g., reviewed in Gold (5)].

The DNA sequence of the intergenic region from the TAA translation stop of the upstream control gene through the first four codons of the downstream control gene is shown in Fig. 5 (panel I). Using an RNA folding program on the transcript of this primary sequence, starting 9 bases upstream of CJ, through the intergenic region (containing ϕ*10* and s*10*), and ending 9 bases downstream of the control gene AUG, a highly stable structure can be derived (Fig. 5, panel II) in which both the control gene AUG and its ribosome binding site are tied up in a structure by base pairing to CJ. Maintaining the same sequence except for reversing the CJ segment does not yield a structure that would obscure the AUG and Shine-Dalgarno (not shown). Thus, it appears plausible that RNA structure of the readthrough transcript is responsible for at least part of the low-level expression of the control gene. However, that still does not explain why expression from the downstream promoter does not provide higher yields of control protein, because we know that when that promoter does not have an upstream T7 gene and/or CJ sequence, expression is unhindered. The answer may have to do with the putative giant transcripts generated from this promoter. Only the first copy of the coding sequence in the giant message would be translatable, while all subsequent readthrough copies would have the upstream inhibitory sequences (Fig. 5, panel III). Thus, readthrough around the plasmid will juxtapose the inhibitory CJ sequence (in the transcript) upstream of the control gene translation start. This mechanism would

I.

*Nhe*I |-------CJ--------| *Bgl*II (22881) |--------ϕ*10*-------1[st] nt--| *Xba*I |----------s*10*----------SD------| *Nde*I (22969)

taa*gctagc*tatctgttacagtctcc*agatct*cgatcccgCgaaatTaatacgactcactataGggagAccacaacggtttccc*tctaga*ataatattttgtttaactttaagaaggagatata*catatg*gctgaatct

II.

```
        10         20         30          40         50         60         70
UAAGC|      C   UACA      AGA   CG --UCCC          AUACGACUCA   U          AC
     UAGCUAU UGU     GUCUCC   UCU  A      GCGAAAUUA           CUA AGGGAGACC
     GUCGGUA ACA     UAGAGG   AGA  U      UGUUUUAAU           GAU UCCCUUUGG  A
UCUAA^      U   --UA      --A   AU UCAAUU          -------AAA   C          CA
      130         120         110        100                90         80
```

III.

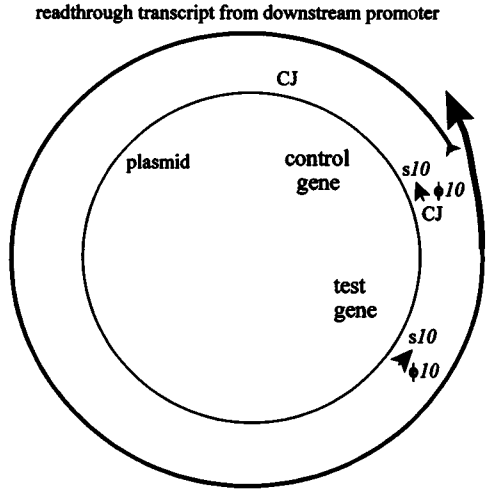

FIG. 5. (I) DNA sequence of the intergenic region between test and control genes in pCT1.1. This region encompasses the CJ element (in bold), a T7 promoter element (ϕ*10*) and s*10* translation start signal, from *Nhe*I (*gctagc*) and *Bgl*II (*agatct*) sites through a naturally occurring *Nde*I (*catatg*) site. The uppercase C denotes the first nucleotide from the T7 genome used in this promoter element, nt 22881. The uppercase T denotes the start of the core 23-bp T7 promoter (corresponding to nt 47–69 in panel II), the uppercase G denotes the first transcribed nt from this promoter, and the uppercase A denotes the last nt of the core T7 promoter. The *Xba*I site (*tctaga*) at nt 22925–22930 is naturally occurring in the T7 genome, as is the entire s*10* element, from *Xba*I through *Nde*I, nt 22925–22969 (corresponding to nt 85–129 in panel II). The Shine-Dalgarno (SD) (ggag) and the atg start codon (last 3 nt of the *Nde*I site) for the control gene are underlined, as is the taa stop codon of the upstream test gene. (II) Potential folded RNA structure between CJ sequence and control gene translation start. 138 nt starting 9 nt 5′ to CJ (***UAUCUGUUACAGUCUCC***) (in bold italics) through 9 nt 3′ of translation start for control gene. Both the Shine-Dalgarno (GGAG) and start codon (AUG) are underlined. Initial energy = −38.9 kcal. RNA folded by mfold version 3.0 [(12) http://www.ibc.wustl.edu/~zuker/rna/form1.cgi]. (III) Proposed readthrough transcript from downstream promoter (in front of control gene), which goes around the plasmid, illustrating how CJ is located upstream of control gene translation start in the readthrough transcript.

not affect test gene expression, because there are no inhibitory sequences upstream of its translation start. To test this hypothesis, it will be necessary to construct a derivative of pCT1.1 with the T7 terminator placed at the 3′ end of the control gene to prevent further readthrough (but retaining the CJ sequence between the genes).

Could Inhibition of Downstream Expression by CJ Play a Role in T7 Morphogenesis?

The natural location of the CJ sequence in the 39937 nt linear T7 genome is immediately 3′ to the 160-bp terminal repeat at the mature left end [in this work, we constructed the CJ sequence starting from the last nt ("T") of the terminal repeat]. In the concatemer, a single copy of the terminal repeat joins the right and left ends of adjacent T7 genomes. During T7 morphogenesis, the concatemer is cut at the 3′ end of the terminal repeat to form the mature right end of the genome, while the terminal repeat is duplicated to form the mature left end [(1), and references therein]. The last open reading frame upstream of the terminal repeat is the nonessential gene *19.5* (nt 39389–39538), encoding a 49-amino acid sequence that has been proposed to facilitate degradation of host DNA (9). Downstream of the terminal repeat, the next rightward promoter is T7 ϕOL (nt 388–410), which serves as a secondary T7 origin of replication, followed by three *E. coli* promoters responsible for transcription of class I (early) T7 genes (2). The first gene at the left end of the T7 genome is gene *0.3*,

whose coding sequence begins at nt 925, downstream of an RNase III processing site at nt 890. Because of this processing, a transcript of the CJ sequence, which is ~750 nt upstream of the translation start, could not have a *cis*-acting effect on expression of gene *0.3* mRNA; perhaps the T7 genome evolved to include the RNase III cut site at nt 890, which otherwise has no obvious function, to prevent interactions with transcripts of upstream sequences, such as CJ. Because of the relative isolation of the CJ element from any open reading frames in its natural setting, it is not evident how or if the property of inhibition of downstream expression reported here might be involved in T7 morphogenesis.

ACKNOWLEDGMENTS

We are indebted to Alan Rosenberg, Brookhaven National Laboratory, for providing us precursor plasmids, which enormously facilitated the construction of pCT1.1, and also for helpful advice and discussions. We also acknowledge Bill Studier (Brookhaven) for insightful comments, and John Dunn (Brookhaven) and Bill McAllister (SUNY-Brooklyn) for critical reading of the manuscript. We gratefully acknowledge Marylynn Snyder, a graduate student in this department, for assisting in the vector construction. This research was supported by the National Science Foundation (grant MCB-9513127).

REFERENCES

1. Chung, Y. B.; Hinkle, D. C. Bacteriophage T7 DNA packaging. II. Analysis of the DNA sequences required for packaging using a plasmid transduction assay. J. Mol. Biol. 216:927–938; 1990.
2. Dunn, J. J.; Studier, F. W. Complete nucleotide sequence of bacteriophage T7 DNA and the locations of T7 genetic elements. J. Mol. Biol. 166:477–535; 1983.
3. Gao, W.; Goldman, E. (1997) Use of SDS-polyacrylamide gel electrophoresis to resolve mRNA and its protein product in one gel. FASEB J. 11:1153–1156; 1997.
4. Gao, W.; Tyagi, S.; Kramer, F. R.; Goldman, E. Messenger RNA release from ribosomes during 5′-translational blockage by consecutive low-usage arginine but not leucine codons in *Escherichia coli.* Mol. Microbiol. 25:707–716; 1997. [Erratum 27:669; 1998.]
5. Gold, L. Posttranscriptional regulatory mechanisms in *Escherichia coli.* Annu. Rev. Biochem. 57:199–233; 1988.
6. Goldman, E.; Zubay, G. Investigating codon usage in *E. coli.* Methods Mol. Genet. 3:159–178; 1994.
7. Goldman, E.; Rosenberg, A.; Zubay, G.; Studier, F. W. Consecutive low-usage leucine codons block translation only when near the 5′ end of a message in *Escherichia coli.* J. Mol. Biol. 245:467–473; 1995.
8. He, B.; Kukarin, A.; Temiakov, D.; Chin-Bow, S. T.; Lyakhov, D. L.; Rong, M.; Durbin, R. K.; McAllister, W. T. Characterization of an unusual, sequence-specific termination signal for T7 RNA polymerase. J. Biol. Chem. 273:18802–18811; 1998.
9. Kim, S. H.; Chung, Y. B. Isolation of a mutant bacteriophage T7 deleted in nonessential genetic elements, gene 19.5 and m. Virology 216:20–25; 1996.
10. Lyakhov, D. L.; He, B.; Zhang, X.; Studier, F. W.; Dunn, J. J.; McAllister, W. T. Mutant bacteriophage T7 RNA polymerases with altered termination properties. J. Mol. Biol. 269:28–40; 1997.
11. Lyakhov, D. L.; He, B.; Zhang, X.; Studier, F. W.; Dunn, J. J.; McAllister, W. T. Pausing and termination by bacteriophage T7 RNA polymerase. J. Mol. Biol. 280:201–213; 1998.
12. Mathews, D. H.; Sabina, J.; Zuker, M.; Turner, D. H. Expanded sequence dependence of thermodynamic parameters provides robust prediction of RNA secondary structure. J. Mol. Biol. 288:911–940; 1999.
13. McAllister, W. T.; Morris, C.; Rosenberg, A. H.; Studier, F. W. Utilization of bacteriophage T7 late promoters in recombinant plasmids during infection. J. Mol. Biol. 153:527–544; 1981.
14. Panayotatos, N.; Truong, K. Cleavage within an RNase III site can control mRNA stability and protein synthesis in vivo. Nucleic Acids Res. 13:2227–2240; 1985.
15. Rosenberg, A. H.; Goldman, E.; Dunn, J. J.; Studier, F. W.; Zubay, G. Effects of consecutive AGG codons on translation in *Escherichia coli* demonstrated with a versatile codon test system. J. Bacteriol. 175:716–722; 1993.
16. Studier, F. W.; Moffatt, B. A. Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. J. Mol. Biol. 189:113–130; 1986.
17. Zhang, X.; Studier, F. W. Isolation of transcriptionally active mutants of T7 RNA polymerase that do not support phage growth. J. Mol. Biol. 250:156–168; 1995.
18. Zhang, X.; Studier, F. W. Mechanism of inhibition of bacteriophage T7 RNA polymerase by T7 lysozyme. J. Mol. Biol. 269:10–27; 1997.