JOURNAL OF BACTERIOLOGY, Jan. 1991, p. 28–36
0021-9193/91/010028-09$02.00/0
Vol. 173, No. 1

Transcription and Decay of the *lac* Messenger: Role of an Intergenic Terminator

GEORGE J. MURAKAWA,† CATHERINE KWAN,‡ JOHN YAMASHITA,§ AND DONALD P. NIERLICH*

Department of Microbiology and Molecular Genetics, and Molecular Biology Institute, University of California, Los Angeles, California 90024-1489

Received 23 April 1990/Accepted 28 September 1990

Prior work has indicated that the polycistronic *lacZYA* mRNA of *Escherichia coli* is cleaved during decay at approximately intergenic sites (L. W. Lim and D. Kennell, J. Mol. Biol. 135: 369–390, 1979). In this work, we characterized the products by using probes specific for the different cistrons. This analysis indicated that six *lac* mRNA species are present in the following order of decreasing abundance: *lacZ*, *-A*, *-ZYA*, *-ZY*, *-YA*, and *-Y*. Very little *lacYA* and *lacY* mRNAs were present, whereas in cells induced to steady state, there was 10 times more *lacZ* than *lacZYA* mRNA. The *lacZ* mRNA appeared as a discrete species extending to a site in the *lacZ-Y* intergenic space (ca. residue 3150). This site is just distal to a potential rho-independent termination sequence. We examined the function of this sequence to determine whether it contributes to the distribution of the mRNAs. Although the termination sequence was shown to function in vitro, when it was recloned into an expression vector, no termination was seen in vivo. Moreover, direct examination of the kinetics of *lac* messenger synthesis revealed that after initiation, most transcription continued to the end of the operon. We conclude that during normal growth, the operon is transcribed in its entirety and that the individual *lac* mRNAs are formed by cleavage. These results confirm earlier work implying that the *lac* operon is transcribed in its entirety but are in conflict with several recent reports suggesting that internal termination occurs. Our findings indicate that the natural polarity of the operon (*lacZ* is expressed sixfold more strongly than *lacA*) is based on posttranslational effects and not on polarity of transcription.

There have been extensive studies of the transcription and decay of the *lac* operon mRNA (reviewed in reference 18). According to current understanding, the *lacZYA* mRNA is transcribed polycistronically and then is cleaved at approximately intergenic sites. These cleavages, near the 5′ ends of the *lacY* and *lacA* transcripts and at an additional site near the 5′ end of the *lacZ* transcript, inactivate the distal mRNAs. The further chemical decay of these mRNAs then occurs as a net 5′-to-3′ process. In this process, endonucleolytic cleavages occur on the mRNA that is exposed as ribosomes run off inactivated messages. The resulting fragments are then removed by the 3′-to-5′ exonucleases, RNase II, polynucleotide phosphorylase and possibly other enzymes (1, 14, 18). The endonucleolytic cleavage sites on the *lacZ* and *lacY* species have been characterized but the enzyme(s) responsible for the endonucleolytic cleavages has not been identified (7, 34). It appears, however, that RNase III is not involved in *lac* mRNA decay, although it does inactivate a few other mRNA species (2, 18, 27a). A newly described endonuclease has some properties inferred for the enzyme that cleaves the *lac* transcripts (6).

The cleavage of the *lacZYA* mRNA was first deduced from the observation that the *lacA* mRNA decays about twice as fast as *lacZ* mRNA (19). Subsequently, the cleavage sites were inferred from the size distribution of the mRNA detected by using specific *lacZ* and *lacYA* probes (18, 21). That these cleavages might inactivate the messages was proposed by Lim and Kennell because they found that smaller *lac* mRNAs did not form initiation complexes in vitro with ribosomes as did *lacZYA* mRNA (21). Endonucleolytic cleavages of the polycistronic *gal*, *trp*, and *malEFG* mRNAs have also been demonstrated, and different segments decay at different rates (14, 18).

To get a clearer view of the fragmentation of the *lac* mRNA in the course of decay, we carried out a Northern (RNA) blot analysis of the *lac* mRNA, using probes specific for the individual cistrons. In the course of this analysis, we characterized the *lacZ* mRNA and found that there is a potential rho-independent terminator located in the *lacZ-Y* intergenic space at a point close to the end of *lacZ* messenger. This finding opened the question as to whether termination might contribute to the natural polarity of *lac*; the *lacZ* gene is expressed at approximately six times the rate of *lacA* (3, 17, 40). This question has been previously addressed, with the conclusion that the operon is transcribed in its entirety (9, 20). Nonetheless, other, more recent work, described in Discussion, suggests that termination occurs within *lac*; this finding led us also to analyze the role of the termination sequence in the formation of the *lacZ* transcript.

MATERIALS AND METHODS

Materials. [α-^{32}P]dATP and [α-^{32}P]CTP were purchased from Amersham; [γ-^{32}P]ATP was from ICN. All solutions were prepared in double-distilled H_2O treated with diethyl pyrocarbonate (Sigma). Restriction enzymes and DNA polymerase (Klenow fragment) were obtained from New England BioLabs, International Biotechnologies, Inc., Promega Biotec, and Bethesda Research Laboratories, Inc.; they were used as recommended by the suppliers.

Bacterial strains and plasmids. *Escherichia coli* W3110 was

* Corresponding author.
† Present address: Harvard Medical School, Boston, MA 02115.
‡ Permanent address: Mount St. Mary's College, Los Angeles, CA 90049.
§ Present address: George Washington University Medical School, Washington, DC 20037.

TABLE 1. Plasmids used

Plasmid	Vector	Insert	*lac* region[a]	Relevant features	Source[b]
		*Hae*III-*Hae*III	−140 to 63	*lac* promoter region, *Eco*RI linkers added	7
		*Hae*III-*Hae*III	−140 to 63	*lac* L8/UV5 promoter region, *Eco*RI linkers added	7
pGEM1				T7/SP6 transcription vector	4
pGEM2				T7/SP6 transcription vector	4
pGM8			3057 to >6000	Derived by *Eco*RI cleavage of pMC1396 and religation	2, 3
pGM201	pGEM2	*Bgl*II-*Bgl*II	Phage lambda 35711 to 38103	Lambda p_R in *Bam*HI site of pGEM2	3
pGM240	pGM201	*Eco*RI-*Fsp*I	3057 to 3380	Lambda p_R upstream of *lacZ-Y* intergenic region in *Eco*RI site	3
pGM241	pGM201	*Fsp*I-*Eco*RI	3380 to 3057	Same as pGM240 but *lacZ-Y* sequence in reverse orientation	3
pGM401	pKO6	*Eco*RI-*Eco*RI from *lac* promoter plasmid	−140 to 63	*lac* promoter upstream of *galK* upstream *Eco*RI site removed	3
pGM402	pGM401	*Eco*RI-*Hae*III	3057 to 3127	*lac'Z*-no intergenic hairpin (see Fig. 3)	3
pGM403	pGM401	*Eco*RI-*Rsa*I	3057 to 3168	*lac'Z*-intergenic hairpin (see Fig. 3)	3
pGM404	pGM401	*Eco*RI-*Fsp*I	3057 to 3380	*lac'Z*-intergenic hairpin-*lacY'* (see Fig. 3)	3
pGM451	pKO6	*Eco*RI-*Eco*RI from *lac* L8/UV5 promoter plasmid	−140 to 63	*lac* L8/UV5 promoter upstream of *galK*, upstream *Eco*RI site removed	3
pGM452	pGM451	*Eco*RI-*Hae*III	3057 to 3127	*lac'Z*-no intergenic hairpin	3
pGM453	pGM451	*Eco*RI-*Rsa*I	3057 to 3168	*lac'Z*-intergenic hairpin	3
pGM454	pGM451	*Eco*RI-*Fsp*I	3057 to 3380	*lac'Z*-intergenic hairpin-*lacY'*	3
pGM820	pGEM2	*Msp*I-*Hae*III	−19 to 63	5'-*Z* probe (Fig. 1), SP6 RNA polymerase	3, 5
pGM830	pGEM1	*Rsa*I-*Fsp*I	3168 to 3380	5'-*Y* probe (Fig. 1), T7 RNA polymerase	3
pGM840	pGEM1	*Hpa*II-*Hpa*II	4550 to 5043	*A* probe (Fig. 1), T7 RNA polymerase	3
pKO6				GalK expression vector	6
pMC1396			63 to >6000	Starting material for pGM8	1

[a] Defined such that the first nucleotide (5'-AATT--) of the *lac* transcript is +1, as shown in Fig. 1.
[b] 1, Casadaban et al. (8); 2, Hediger et al. (13); 3, this report; 4, Promega Biotec; 5, Murakawa and Nierlich (27a); 6, McKenney et al. (23); 7, Stefano and Gralla (33).

used for the preparation of RNA. Cultures were grown in YT medium (24) at 37°C with vigorous shaking. The plasmids used are listed in Table 1.

The plasmids used for preparing RNA probes for the Northern blots were derived from pGEM1 and pGEM2 (Promega Biotec), which allow for the specific transcription in vitro of the cloned sequences. Plasmid pGM820 contains the first 63 nucleotides of the *lac* operon message (26, 27). After linearization of the plasmid with *Hin*dIII, the probe was synthesized with SP6 RNA polymerase as described below. Plasmid pGM830 was constructed by purifying the appropriate *lacY* fragment from pGM404 and inserting it into pGEM1, using standard procedures (22). This plasmid was digested with *Hin*dIII prior to transcription with T7 RNA polymerase. Plasmid pGM840 was constructed by purifying an *Eco*RI-*Hin*dIII (Table 1) from an M13mp18 clone used previously in sequencing the *lacA* gene (13). To prepare the *lacA*-specific (*A*) probe, pGM840 was digested with *Hin*dIII and transcribed with T7 polymerase. All constructions were verified by dideoxy sequencing.

Plasmid pGM8 contains the distal 53 nucleotides of *lacZ* and all of *lacYA*. Plasmid pGM8 was used in the original sequence determination of the *lacA* gene, verifying the presence of the intact *lacY* and *lacA* genes on the plasmid (13). Plasmid pGM8 was used for S1 nuclease mapping of the 3' end of the *lacZ* transcript and was the starting material for

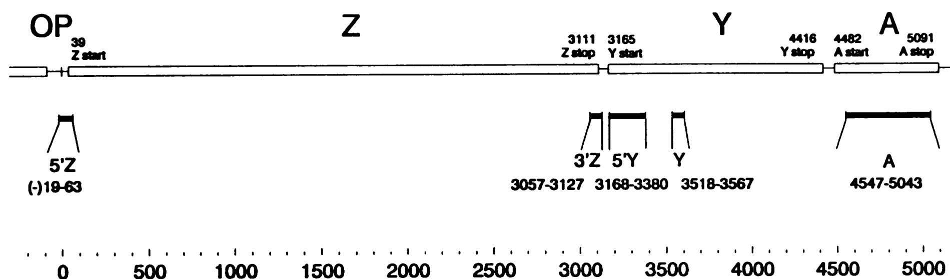

FIG. 1. Map of the *lac* operon showing probes used in Northern blot hybridization. Data for the *lac* genes are from the GenBank data base.

construction of the various plasmids spanning the 3′ *lacZ*, *lacY*, and *lacA* regions.

Construction of pKO-derived plasmids for in vivo determination of *lacZY* termination efficiency. Plasmid pKO6, a *galK* expression vector lacking an upstream promoter, was used to construct six *lac-galK* plasmids (23). Two 203-bp *Eco*RI fragments containing either the wild-type or mutant L8/UV5 *lac* promoter (33; Table 1) were inserted into pKO6 at the *Eco*RI site. The upstream *Eco*RI site was removed by partially digesting each plasmid with *Eco*RI, end filling with the Klenow fragment of DNA polymerase, and religating (22). The resulting plasmids, pGM401 and pGM451, contained the wild-type and L8/UV5 promoters, respectively, with a unique downstream *Eco*RI site. Three separate fragments containing part or all of the *lacZ-Y* intercistronic region (Table 1) were isolated and inserted into these plasmids. The fragments were purified from acrylamide gels of digests of pGM8 as indicated. *Eco*RI linkers were added to the blunt ends, and the fragments were digested with *Eco*RI and inserted into pGM401 and pGM451. Galactokinase assays were performed as described previously (23). Plasmid copy numbers were determined by assaying for β-lactamase expression (31). There was less than 15% difference in the plasmid copy numbers.

In vitro transcription. Plasmid pGM240 was the template for in vitro transcription of the *lacZ-Y* intergenic region. A 2,392-bp fragment bearing the lambda p_R promoter was isolated after digestion of lambda DNA with *Bgl*II. It was inserted into pGEM2 cleaved with *Bam*HI to give pGM201. This plasmid was subsequently cleaved at an adjacent *Eco*RI site, and the *lac Eco*RI-*Fsp*I (*Eco*RI) fragment from pGM404 was introduced in both orientations to give pGM240 and pGM241. Transcription was carried out essentially as described by Schmidt and Chamberlin (32). *E. coli* RNA polymerase holoenzyme was prepared by published methods (12); rho was prepared as described by Mott et al. (25). Size standards for gel electrophoresis were obtained by transcription of plasmids; 102 bases, pGM202 cut with *Eco*RI; 150 bases, pGM201 cut with *Pvu*II; and 530 bases, pGM202 cut with *Sph*I.

Northern blot hybridizations. Cultures were grown in YT medium to mid-log phase (optical density at 540 nm of 0.6), and *lac* expression was induced with isopropyl-β-D-thiogalactoside (IPTG; 5×10^{-4} M). Cyclic AMP (5.0 mM) was added 1 min prior to induction. Where indicated, rifampin was added to 300 μg/ml from a 100-mg/ml stock in dimethyl sulfoxide. Portions (10 ml) were taken at the indicated times and poured onto 2 ml of frozen M9 salts (24) containing 60 mM sodium azide and 1.2 mg of chloramphenicol per ml. After centrifugation for 5 min at 6,800 × *g* the cells were resuspended in 1 ml of cold lysis buffer (10 mM Tris hydrochloride [pH 7.5], 10 mM KCl, 5 mM $MgCl_2$, 200 μg of chloramphenicol per ml, 10 mM sodium azide), and 10 mg of lysozyme was added. The suspension was held on ice for 5 min and then frozen in dry ice-ethanol. Sodium dodecyl sulfate (SDS; 150 μl of a 10% [wt/vol] stock solution) and sodium acetate (35 μl, 3.0 M, pH 5.0) were added, and the cells were lysed at 65°C. An equal volume of hot, redistilled phenol (equilibrated with 100 mM sodium acetate, pH 5.0) was added, and the samples were extracted (22). Routinely four to five phenol extractions were performed until no visible material was seen at the interphase. The samples were then extracted twice with chloroform-isoamyl alcohol (24:1). Finally, the RNA was precipitated with ethanol (2.5 ml) after addition of 100 μl of 3.0 M sodium acetate (pH 5.0).

RNA samples (10 μg) were separated by electrophoresis on horizontal 1% agarose gels containing 1.2 M formaldehyde (22), which were then stained with acridine orange and photographed. The RNA was transferred to a nylon-based membrane (BioDyne; ICN) by capillary action, using 20× SSC (3 M NaCl, 0.3 M sodium citrate, pH 7.0). For probing, the blots were incubated at 45°C in prehybridization buffer (50% formamide, 6× SSC, 5× Denhardt solution [0.1% polyvinylpyrrolidone, 0.1% bovine serum albumin, 0.1% Ficoll], 0.1% SDS, 1 mg of denatured salmon sperm DNA per ml, 50 mM $NaPO_4$, pH 6.5 [35]). After at least 4 h, approximately 10^7 cpm of probe (see below) was added to the hybridization bag. Hybridization was carried out for 12 to 16 h at 37 to 65°C, optimized for the different RNA probes. The filters were washed, dried, and autoradiographed, using intensifying screens and Kodak XAR-5 film. Densitometric scannings were performed by using an Optronics drum scanner.

The RNA probes were prepared by transcription of pGEM plasmids with T7 or SP6 RNA polymerase, using [α-^{32}P]CTP (10, 27). The plasmids (Fig. 1 and Table 1) were linearized with the restriction enzymes indicated. After transcription, the products were separated by electrophoresis on a 5% acrylamide gel containing 7 M urea, and the desired band was excised and added directly to the hybridization mixture. In addition, several DNA oligomers (40- and 50-mers) were synthesized on a Dupont Coder 300 synthesizer and end labeled with [γ-^{32}P]ATP and T4 polynucleotide kinase for use as probes (22).

S1 nuclease analysis. RNA was annealed to a labeled DNA probe and then digested with S1 nuclease as described by Berk and Sharp (4). The products were separated on an 8% acrylamide–7 M urea gel. The probe was prepared by digestion of pGM8 with *Eco*RI, and the staggered ends (at position 3058) were filled with DNA polymerase (Klenow fragment) and [α-^{32}P]dATP.

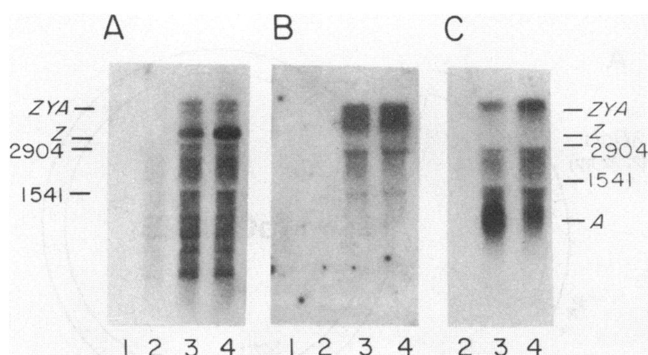

FIG. 2. Northern blot analysis of *lac* mRNA. Probes used: (A) 5'-*Z*; (B) 5'-*Y*; (C) *A* (see Fig. 1). Lanes: 1, uninduced; 2 to 4, 1, 5, and 15 min, respectively, following induction.

RESULTS

Characterization of cellular *lac* transcripts. To determine the identities and relative abundances of *lac* operon messenger species, we used Northern blot hybridizations with RNA probes. For this purpose, the *lac* fragments shown in Fig. 1 were cloned into vectors carrying promoters from phage T7 and SP6 so that the inserts could be transcribed in vitro. In addition, several 40- and 50-mer DNA oligomers were synthesized to use as probes. Blots of *lacZ*-, *lacY*-, and *lacA*-containing messengers can be seen in Fig. 2. The rRNAs (1,541 and 2,904 bases) in the samples were used for estimating the sizes of the mRNAs as well as providing an internal control. They saturate the binding sites on the nylon membrane at their positions and create "shadows" in the *lac* mRNA pattern.

Figure 2A shows the results obtained by using the 5' *lacZ*-specific (5'-*Z*) probe. There is little difference between samples from cultures that were induced for 5 or 15 min, indicating that a steady state of *lacZ* transcription had nearly been reached within 5 min of induction. The predominant band seen at steady state is that of the approximately 3,100-nucleotide band corresponding to the *lacZ* mRNA; there is less of the 5,200-nucleotide full-length *lacZYA* mRNA. The full-length band is consistent with our finding that the major 3' RNA endpoint at the end of the operon lies about 100 bases beyond the *lacA* gene (13). Just below the *lacZYA* band can be seen a *lacZY* species. It can be identified by size (ca. 4,500 bases) and failure to hybridize with the *lacA* probe (Fig. 2C). The *lacZY* and *lacZYA* species can also be seen, however faintly, with *lacY* probes (Fig. 2B); they are obscured by the background of growing and decaying chains (see below).

The relatively distinct bands of *lacZ*, -*ZY*, and -*ZYA* species in Fig. 2A are superimposed on a significant background of mRNA fragments. These fragments presumably represent mRNAs that were growing or decaying or mRNAs that were degraded in the course of RNA purification. That purification degraded a negligible fraction of the molecules is shown by the fact that the band patterns were gene specific. For example, when the *lacZYA*-*ZY* region was probed with a *lacA* probe (see below and Fig. 2C), only the *lacZYA* band was detected.

In fact, on the blot shown in Fig. 2A, the mRNA fragments detected with the 5'-*Z* probe that run ahead of the different distinct bands are largely growing transcripts because the 5' end of the *lacZ* mRNA is removed first during decay (5), as demonstrated below. The significant amount of *lacZ* fragments compared with the amount of *lacZ* mRNA is consistent with the fact that the time of synthesis of *lacZ* mRNA (about 80 s for 3,100 bases) is about the same as its half-life (105 s [17]).

To quantitate the relative amounts of the mRNA species, densitometric scans of the autoradiograms were carried out. Integration of the peaks from Fig. 2A indicates that there is approximately 10-fold more *lacZ* messenger than *lacZYA* messenger. Thus, from the results of the Northern blots, we conclude that the majority of the *lac* mRNA molecules are terminated or cleaved in the *lacZ*-*Y* intercistronic region. Cleavage of the *lacZYA* mRNA to produce monocistronic *lacZ*-containing species has been demonstrated previously (18, 21).

When an identically prepared blot was probed with the *lacY* probe, species varying from about 3,300 bases to full-length *lacZYA* mRNA were seen to predominate (Fig. 2B, lanes 3 and 4). The 1-min sample did not hybridize with the *lacY* probe and *lacY* transcripts (greater than 3,200 bases in length) could only be detected after 1.5 min of induction (not shown). Allowing 10 s for induction lag, this is consistent with a rate of transcription of about 40 nucleotides per s. As expected, the 3,100-base band identified as *lacZ* mRNA did not appear. However, a similarly strong and distinct *lacZ* band was detected when gels were probed with the adjacent 3'-*Z* probe (Fig. 1) (27a). From this finding, we conclude that the 3' endpoint of the *lacZ* message lies close to or in the *lacZ*-*Y* intergenic space, consistent with our S1 mapping of the 3' end (see below).

If the *lacZYA* transcript is cleaved to yield a *lacZ* species, one would also expect to find *lacYA* and possibly *lacY* messenger present. A small amount of *lacYA* message could be detected with the *Y* probe (Fig. 2B) and also with the *A* probe (Fig. 2C). It ran between the two rRNA markers, in accord with its predicted length of 2,000 bases. Remarkably, there was only the smallest amount of *lacY* mRNA species detectable. *lacY* mRNA, about 1,300 bases, should migrate ahead of the 1,541-base 16S rRNA. To confirm this finding, we used a second *lacY*-hybridizing probe, a synthetic deoxyoligomer extending from 3518 to 3567, and obtained an identical result (results not shown). In addition, to exclude the possibility that a band was obscured by the 16S rRNA, a series of gels, probed with the *Y* probe, was run in which the agarose concentration and conditions of denaturation were altered (34a).

Figure 2C shows an identically prepared gel that was probed with an *A* probe. An abundant monocistronic *lacA* mRNA is clearly visible in addition to the polycistronic species. Electrophoresis of samples on 1.2% gels indicated that there were two *lacA* species at this position; they were approximately 600 to 700 bases in length, which suggests that they represent the *lacA* transcript with and without the 100-base extension to the predominant *lac* end that we mapped previously (13).

Figure 2 shows the distributions of the different *lac* mRNA species. Present are *lacZYA*, -*ZY*, -*YA*, -*Z*,-*Y*, and -*A* messages. All except the *lacY* and -*YA* species formed distinct, strong bands.

The *lacZ* mRNA. To characterize the *lacZ* message, we mapped the 5' and 3' ends. The 5' end of the predominant *lac* species in the cell has been previously determined to coincide with the start points (residues A1 and A2) of *lac* transcription in vitro, and we have confirmed this by primer extension analysis (5; 37a). The 3' end was determined by S1 nuclease mapping. The results (Fig. 3) demonstrate a single, strong terminus at approximately residue 3150. (A preliminary report has been made [28].)

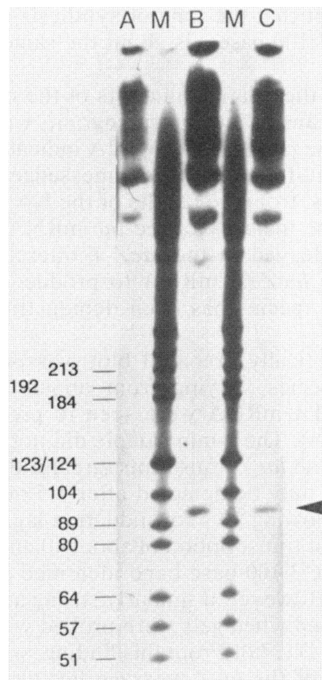

FIG. 3. S1 nuclease mapping of the 3′ end of the *lacZ* transcript. Lanes: A, RNA from uninduced cells; B and C, RNA from induced cells treated for 30 and 60 min, respectively, with S1 nuclease; M, size markers (*Hae*III digest of nick-translated pBR322; positions indicated in nucleotides on the left). The protected band is indicated by an arrow.

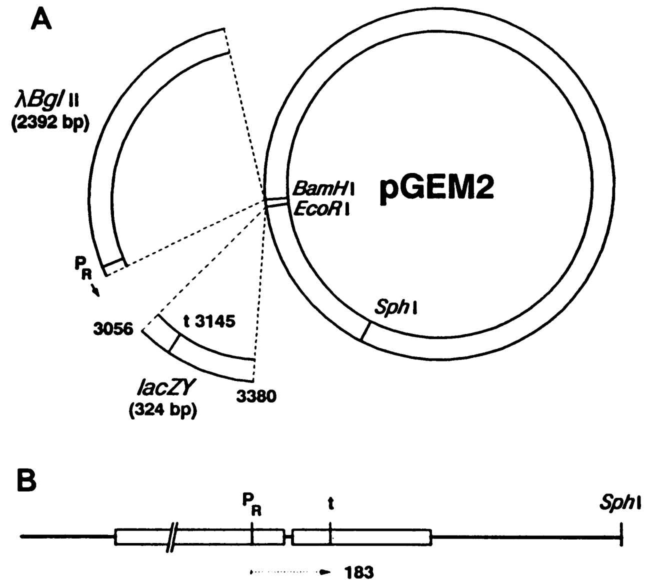

FIG. 5. Template for in vitro transcription. (A) Construction of pGM240; (B) transcripts formed on an *Sph*I-cleaved template.

In vitro and in vivo assessment of transcriptional termination. The *lacZ* endpoint (Fig. 3) is located approximately 8 bases beyond the base of the stem-loop sequence that is a potential transcriptional terminator (Fig. 4). Rho-independent termination sequences are characterized by a G+C-rich dyad followed by a sequence rich in U's in the RNA transcript (38). This finding opens the possibility that transcriptional termination contributes to the formation of the *lacZ* species. Alternatively, such stem-loop sequences have also been shown to protect mRNAs from attack by 3′ exonucleases (Discussion). Thus, it is possible that the 3′ end of the *lacZ* mRNA is the consequence of the cleavage of the mRNA distal to it.

To determine whether the intergenic sequence can function as a terminator, a fragment containing this region was subcloned behind the lambda p_R promoter (Fig. 5) and transcribed in vitro with *E. coli* RNA polymerase (Fig. 6). On a superhelical template, termination occurred at a single site, close to that observed in vivo (ca. residue 183 in Fig. 5). No termination was observed with a template possessing an inverted intergenic sequence. On a template linearized by *Sph*I cleavage, termination at the intergenic site was 5 to 10% of the 764-base runoff transcript, but there was also significant random termination (not shown). Addition of rho factor did not measurably alter the results with either template.

To assess the extent of termination in vivo, three fragments that span part or all of the *lacZ-Y* intercistronic region were purified and inserted into a derivative of pKO6 as described in Materials and Methods. Plasmid pKO6 functions as an expression vector in which promoters and terminators affect expression of a *galK* gene located downstream of a convenient cloning site (23). Both the wild-type and mutant L8/UV5 *lac* promoters were introduced upstream of the *galK* gene in combination with the various

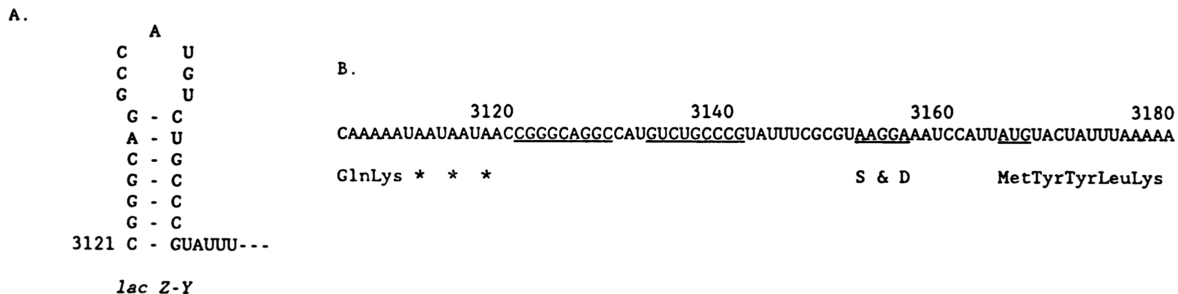

FIG. 4. Transcriptional terminator and sequence of the *lacZ-Y* intergenic region. (A) Secondary-structure diagram. The structure possesses a calculated ΔG of −17.3 kcal (ca. −72.4 kJ)/mol (36). (B) Sequence of the intergenic region. Shown are the three consecutive stop codons of *lacZ* (***), the *lacY* Shine-Dalgarno region and initiation codon (S & D and underscoring), and hairpin sequence (underscored). The end of the *lacZ* transcript mapped in Fig. 3 is at ca. 3150.

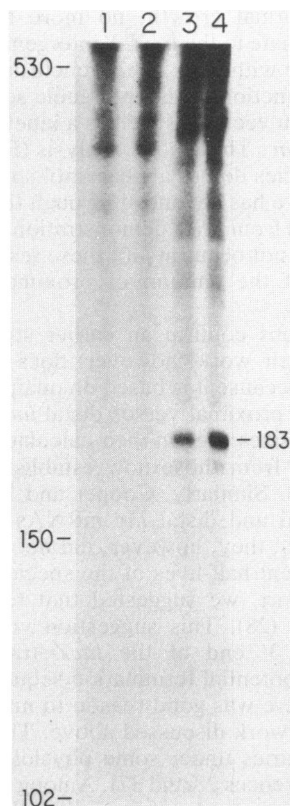

FIG. 6. Transcription in vitro of the intergenic region. Plasmids pGM240 (lanes 3 and 4) and pGM241 (lanes 1 and 2) in supercoiled form were transcribed, and the phenol-purified products run on an 8% acylamide-7 M urea gel. Positions of RNA size markers (in nucleotides) are indicated on the left. The site of termination (about residue 183) corresponds to residue 3150 of Fig. 2.

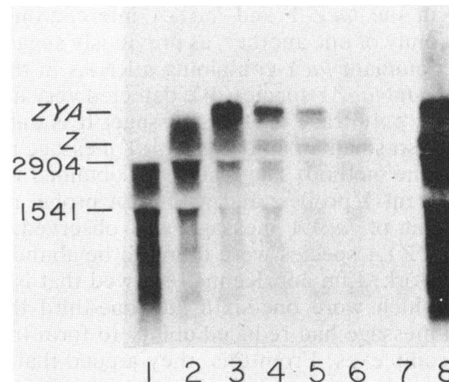

FIG. 7. Kinetics of *lac* transcription. A culture of *E. coli* W3110 was induced with cyclic AMP and IPTG and treated 1 min later with rifampin; samples were removed at various times. Lanes: 1 to 6, 1 to 6 min, respectively, after induction; 7, 10 min after induction; 8, 10 min after induction, no rifampin added.

lacZ-Y fragments. No evidence was found that the *lacZ-Y* region contains an active terminator (data not shown). That is, *galK* expression was approximately the same whether the full *lacZ-Y* intergenic sequence (on an *Eco*RI-*Rsa*I or *Eco*RI-*Fsp*I fragment) was present upstream of the *galK* gene or the fragment ending short of the terminator (*Eco*RI-*Hae*III fragment or no insert) was present.

Kinetics of *lac* mRNA accumulation in vivo. To confirm the result discussed above and to demonstrate directly the formation of the *lacZ* transcript by cleavage, the kinetics of accumulation in vivo of the 3,150-nucleotide *lacZ* mRNA were determined. Cells were induced with IPTG, and 1 min later rifampin was added to prevent further transcriptional initiation. Cells were harvested at that time and at timed intervals thereafter. The RNA samples were then analyzed by Northern blot analysis, using the 5′-*Z* probe. After the addition of rifampin, the nascent molecules grew in length (Fig. 7); the small fragments (<3,100 bases) disappeared, consistent with their being largely composed of growing rather than decaying chains, as mentioned above. Most important, the wave of growing transcripts largely passed through the *lacZ*-length region (Fig. 7, lane 2) to form full-length *lacZYA* molecules (lane 3). We estimate that not more than 10% of the transcription complexes terminate in the *lacZ-Y* intergenic region. During the subsequent minutes (lanes 4 to 6), the *lacZ* message persisted while the *lacZYA* message rapidly disappeared. This finding underlies the observation that in steady state there is more *lacZ* message than *lacZYA* message (Fig. 7, lane 4 versus lane 8; Fig, 2A, lane 4).

To determine whether termination might be observed under other conditions, we examined the *lac* transcripts of *E. coli* strains carrying RNA polymerase mutations that result in enhanced or diminished termination by rho-independent terminators (11). When RNAs from these strains were analyzed on Northern blots, no significant difference in the relative abundance of proximal and distal transcripts was seen with either the 5′-*Z* or 5′-*Y* probe (data not shown).

DISCUSSION

We have determined the distribution in the cell of the several species of *lac* operon transcripts and confirmed that the individual cistronic messengers are formed by cleavage of the polycistronic transcript. In addition, we examined in some detail a rho-independent terminator located in the *lacZ-Y* intergenic region. We conclude that it does not function significantly in normally growing cells.

Transcription and cleavage of the *lac* mRNA. In prior work, Lim and Kennell were able to determine that the *lacZYA* transcript was cleaved to form a distinct *lacZ* mRNA species and several other species, as deduced from their sizes and from hybridization with a *lacZYA*, -*Z*, or -*YA* probe (21). Distinct *lacY* or *lacA* species were not identified. In the work reported here, the greater resolution of Northern blot hybridization along with the use of specific probes allowed us to extend their findings. We are able to detect *lacZ*, -*Y*, -*A*, -*ZY*, -*YA*, and -*ZYA* species. Although it is difficult to make precise quantitative comparisons between the blots examined with different probes, the relative abundance of the species was found to be *lacZ* ≥ *lacA* > *lacZYA* = *lacZY* > *lacYA* > *lacY*. The virtual absence of discrete *lacY* and *lacYA* messages indicates that following their formation by cleavage from the *lacZ* transcript, the *lacY*-containing species are rapidly degraded.

Measurement of the time course of appearance of the individual *lac* transcripts clearly shows that they are cleaved from the *lacZYA* transcript. However, it appears that transcription of the full-length species need not be complete for cleavage to occur; some cistronic *lacZ* mRNA can be detected prior to 2 min. The appearance of *lacZY* and *lacYA* species, with the monocistronic species, implies that the

cleavages in the *lacZ-Y* and *lacY-A* intercistronic regions occur randomly of one another, as previously suggested (34).

The predominant *lacY*-containing mRNAs in the cell are the *lacZYA* and *lacZY* species. We detected very little *lacYA* message and only trace of *lacY* message. It is unlikely that detection of so small an amount of *lacY* message represents a failure of the method. The result was obtained with either of two different *Y* probes; and with an *A* probe, a similarly small amount of *lacYA* message was observed, although *lacA* and *lacZYA* species were found to be abundant.

In prior work, Lim and Kennell showed that isolated *lac* messages which were one-sixth and one-third the size of full-length message had reduced ability to form translation-initiation complexes. From this, they argued that the intergenic cleavages inactivate the distal cistronic transcripts (21). Our observations indicate that after initial cleavage in the *lacZ-Y* intergenic region, secondary cleavages occur very rapidly. Moreover, while it seems likely that the cleavage that separates the *lacZ* and *lacY* transcripts concomitantly inactivates the *lacY* mRNA, it is also possible that a secondary event is required. We are currently analyzing the inactivation of the *lacA* mRNA, since the greater abundance of the mRNA promises to make this question easier to examine. The presence of a strong *lacA* band on the blots suggests that this species may be translationally active. On the other hand, there is evidence that the *lacY* and *lacA* messengers in the cell decay at the same rate, each at twice the rate of the *lacZ* mRNA (7). This finding suggests that the *lacA* species that we observe is also inactive but not subject to so rapid a chemical decay.

Recently, Cannistraro et al. mapped the ends of the *lac* messenger species in the *lacZ-Y* region (7). They found that the 5′ ends of molecules representing *lacY* transcripts were distributed over a substantial number of discrete sites varying in abundance. This distribution of 5′ cleavage sites is consistent with a model in which decay of the *lacY* mRNA either follows quickly after a preferential cut close to the *lacZ-Y* region or is the result of a process in which cleavage of the *lacY* messenger can occur throughout *lacY*. Either model is consistent with our finding that *lacY* and *lacYA* mRNAs are in low abundance.

In the same work, Cannistraro et al. (7) also mapped the 3′ ends in the *lacZ-Y* region by using S1 analysis in much the way that we have. In contrast to our finding of a single, strong band for the 3′ end of *lacZ* mRNA at approximately position 3150, they found numerous bands in the *lacZ-Y* region, including a minor band at about 3150 (89 in the numbering of Fig. 1 of reference 7). In addition, in the context of our earlier report that there was a potential termination sequence at this site (28), they also examined indirectly whether termination occurred at this site relative to the other sites. They concluded that termination did occur at this site but that it was of minor importance in forming the *lacZ* mRNA because this site was minor relative to the many cleavage sites. We agree with this conclusion. However, it is not altogether clear why they observe so many different 3′ ends on the *lacZ* messenger. Nonetheless, the two observations are not necessarily mutually exclusive. It is possible that cleavage initially occurs distal to 3150 and that 3′ exonucleases shorten the *lacZ* mRNA to ca. 3150 (see below). Thus, in situations in which different times of induction or different strains were used, the ends obtained might vary. The discrete *lacZ* band that we observe on Northern blots (e.g., Fig. 2A) indicates that the *lacZ* species is fairly well defined.

Transcriptional termination in *lac*. We conclude that under conditions of normal growth, no more than 10% of the transcripts terminate in the *lacZ-Y* intergenic space or at any other distinct site within *lac*. This conclusion is based on an analysis of the function of the intergenic sequence recloned into an expression vector, as well as a kinetic analysis of the transcription of *lac*. The kinetic analysis (Fig. 7) shows that the cistronic species do not appear until some time after the transcription wave has advanced through the *lacZ-Y* region. This result differs from prior demonstrations (see below) that termination does not occur in *lac*; these results are based on measurements of the amount of proximal and distal *lac* mRNAs.

Our observations confirm an earlier study by Lim and Kennell (20). Their work, however, does not entirely rule out termination because it is based on quantitative estimates of the amount of proximal versus distal *lac* messenger, and the relative gene sizes used in their calculations were almost twofold different from those now established by DNA sequencing (Fig. 1). Similarly, Cooper and Magasanik found that the proximal and distal *lac* mRNAs were present in equal amounts (9); they, however, did not take into consideration the different half-lives of the species (7).

In a prior report, we suggested that termination might occur within *lac* (28). This suggestion was based on our mapping of the 3′ end of the *lacZ* transcript and our recognition of a potential termination sequence in the intergenic region. There was good reason to make this proposal despite the prior work discussed above. The sixfold natural polarity of *lac* varies under some physiological conditions (reviewed in references 28 and 37). Among these are a series of reports by Ullmann and colleagues that the ratio of *lacA* to *lacZ* expression changes during catabolite repression and that this response is relieved in rho mutants (37).

In addition, Holben and Morgan (15) reported that *lacA* expression is increased when *lac* is transcribed from a fusion with the *rrnC* promoter and leader region; the *rrn* leader possesses antitermination sequences (15, 16). Utilization of the *rrnC* promoter was shown to increase production of transacetylase relative to β-galactosidase by 9.6-fold (15). Because of the antitermination activity of the *rrn* promoters is directed toward rho-dependent events, it seems possible that their results are a consequence of induced polarity stemming from the particular constructions that they used. Their experiments were carried out with recombinant strains in which *rrnC* was joined to a *trp-lac* fusion. One previously observed feature of such *lac* fusions is the inefficient expression of the *lacZ* gene and the relative overexpression of the *lacA* genes (30, 39).

Although it remains to be determined whether the weak terminator in the *lacZ-Y* intergenic region functions more strongly under some specific physiological conditions, we conclude that the natural polarity of *lac* is not due to termination of transcription within *lac* and concur with Lim and Kennell as to their analysis (21). They attribute the natural polarity of *lac* first to an approximately twofold faster decay of *lacA* mRNA than *lacZ* mRNA (7, 21) and second to the weak translational efficiency of *lacA* mRNA inferred to exist on the basis of the relative abundances of mRNAs and translational products (17, 20). In keeping with this view, it can now be noted that *lacA* possesses an unusual initiation codon, UUG, and uses relatively rare codons (13). Reddy et al. have shown that altering the natural UUG start codon of adenylate cyclase to AUG increases its expression three- to sixfold (29).

Role of the *lacZ-Y* sequence. What then is the role of the intercistronic *lacZ-Y* stem-loop? One possibility is that it is

an RNA cleavage site, as suggested by a slightly different folding of this same sequence to resemble an RNase III cleavage site (7). A second possibility is that it stabilizes the *lacZ* mRNA against the activities of 3' exonucleases, an inference for which we now have direct evidence (19a). Hairpin structures located at the 3' end of several bacterial transcripts have been shown to stabilize the upstream mRNA (14). For example, the stem-loop contained in the *trpAt* terminator appears to prevent exonucleolytic decay of the *trpA* messenger (25). *trpAt* encodes a structure that strongly resembles the *lacZ* structure (Fig. 4A; Fig 2 of reference 28) and similarly has been shown to be a weak terminator in vitro and in vivo. Also, another strong stem-loop structure is present just distal to the *lacA* gene (13). Stem-loop structures are also formed by *rep* (repetitive extragenic palindromic) sequences (14), one of which is present in the *lacY-A* intergenic region (13). Such sequences are found in the extragenic regions of many operons and have been directly shown to stabilize upstream transcripts (14). Thus, strong stem-loop sequences are located distal to each of the three *lac* genes. Their function appears to prevent 3' exonuclease degradation of the upstream coding regions.

ACKNOWLEDGMENTS

We thank T. Thai, H.-Y. Wei, and R. Stewart, who helped at various stages of this project. We thank K. McKenney, T. Platt, and C. Yanofsky for plasmids and strains, J. Boroweic and J. Gralla for RNA polymerase, and M. Schmidt for helpful suggestions.

This research was supported by Public Health Service grant GM 37126 from the National Institutes of Health. G.M. was supported by National Research Service award GM 07104 from the National Institutes of Health.

REFERENCES

1. **Arraiano, C. M., S. D. Yancey, and S. R. Kushner.** 1988. Stabilization of discrete breakdown products in *ams pnp rnb* multiple mutants in *Escherichia coli* K-12. J. Bacteriol. **170:** 4625–33.
2. **Bardwell, J. C. A., P. Regnier, S.-M. Chen, Y. Nakamura, M. Grunberg-Manago, and D. L. Court.** 1989. Autoregulation of RNaseIII operon by mRNA processing. EMBO J. **8:**3401–3407.
3. **Beckwith, J.** 1987. The lactose operon, p. 1444–1452. *In* F. C. Neidhardt, J. L. Ingraham, K. B. Low, B. Magasanik, M. Schaechter, and H. E. Umbarger (ed.), *Escherichia coli* and *Salmonella typhimurium*: cellular and molecular biology. American Society for Microbiology, Washington, D.C.
4. **Berk, A. J., and P. A. Sharp.** 1977. Sizing and mapping of the early adenovirus mRNAs by gel electrophoresis of S1 endonuclease-digested hybrids. Cell **12:**721–732.
5. **Cannistraro, V. J., and D. Kennell.** 1985. The 5' ends of *Escherichia coli lac* mRNA. J. Mol. Biol. **182:**241–248.
6. **Cannistraro, V. J., and D. Kennell.** 1989. Purification and characterization of ribonuclease M and mRNA degradation in *Escherichia coli*. Eur. J. Biochem. **181:**363–370.
7. **Cannistraro, V. J., M. N. Subbarao, and D. Kennell.** 1986. Specific endonucleolytic cleavage sites for decay of *Escherichia coli* mRNA. J. Mol. Biol. **192:**257–274.
8. **Casadaban, M. J., J. Chou, and S. N. Cohen.** 1980. In vitro gene fusions that join an enzymatically active β-galactosidase segment to amino-terminal fragments of exogenous proteins: *Escherichia coli* plasmid vectors for the detection and cloning of translational initiation signals. J. Bacteriol. **143:**971–980.
9. **Cooper, T. G., and B. Magasanik.** 1974. Transcription of the *lac* operon. J. Biol. Chem. **249:**6556–6561.
10. **Davanloo, P., A. H. Rosenberg, J. J. Dunn, and F. W. Studier.** 1984. Cloning and expression of the gene for bacteriophage T7 RNA polymerase. Proc. Natl. Acad. Sci. USA **81:**2035–2039.
11. **Fisher, R. F., and C. Yanofsky.** 1983. Mutations of the β subunit of RNA polymerase alter both transcription pausing and transcription termination in the *trp* operon leader region in vitro. J. Biol. Chem. **258:**8146–8150.
12. **Gonzalez, N., J. Wiggs, and M. J. Chamberlain.** 1977. A simple procedure for resolution of *Escherichia coli* RNA polymerase holoenyzme from core polymerase. Arch. Biochem. Biophys. **182:**404–408.
13. **Hediger, M. A., D. F. Johnson, D. P. Nierlich, and I. Zabin.** 1985. DNA sequence of the *lactose* operon: the *lacA* gene and the transcriptional termination region. Proc. Natl. Acad. Sci. USA **82:**6414–6418.
14. **Higgins, C. F., R. S. McLaren, and S. F. Newbury.** 1988. Repetitive extragenic palindromic sequences, mRNA stability and gene expression: evolution by gene conversion?—a review. Gene **72:**3–14.
15. **Holben, W. E., and E. A. Morgan.** 1984. Antitermination of transcription from an *Escherichia coli* ribosomal RNA promoter. Proc. Natl. Acad. Sci. USA **81:**6789–6793.
16. **Holben, W. E., S. M. Prasad, and E. A. Morgan.** 1985. Antitermination by both the promoter and the leader regions of an *Escherichia coli* ribosomal RNA operon. Proc. Natl. Acad. Sci. USA **82:**5073–5077.
17. **Kennell, D., and H. Riezman.** 1977. Transcription and translation initiation frequencies of the *Escherichia coli lac* operon. J. Mol. Biol. **114:**1–21.
18. **Kennell, D. E.** 1986. The instability of messenger RNA in bacteria, p. 101–142. *In* W. S. Resnikoff and L. Gold (ed.), Maximizing gene expression. Butterworths, Stoneham, Mass.
19. **Kepes, A.** 1967. Sequential transcription and translation in the lactose operon of *E. coli*. Biochim. Biophys. Acta **138:**107–123.

19a. **Kwan, C., and D. P. Nierlich.** 1990. Abstr. Annu. Meet. Am. Soc. Microbiol. 1990, H-89, p. 169.

20. **Lim, L. W., and D. Kennell.** 1974. Evidence against transcription termination within the *E. coli lac* operon. Mol. Gen. Genet. **133:**367–371.
21. **Lim, L. W., and D. Kennell.** 1979. Models for decay of *Escherichia coli* messenger RNA and evidence for inactivating cleavages between its messages. (Appendix: Mathematical models for decay of linear macromolecules.) J. Mol. Biol. **135:**360–390.
22. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
23. **McKenney, K., H. Shimatake, D. Court, U. Schmeissner, C. Brady, and M. Rosenberg.** 1981. A system to study promoter and terminator signals recognized by *E. coli* RNA polymerase, p. 383–415. *In* J. C. Chirikjian and T. S. Papas (ed.), Gene amplification and analysis, vol. II. Elsevier North-Holland, New York.
24. **Miller, J. H.** 1972. Experiments in molecular genetics. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
25. **Mott, J. E., J. L. Galloway, and T. Platt.** 1985. Maturation of *Escherichia coli* tryptophan operon mRNA: evidence for 3' exonucleolytic processing after rho-dependent termination. EMBO J. **4:**1887–1891.
26. **Murakawa, G. J., C.-H. B. Chen, M. D. Kuwabara, D. P.Nierlich, and D. S. Sigman.** 1989. Scission of RNA by 1,10-phenanthroline-copper ion: preference for single-stranded loops. Nucleic Acids Res. **17:**5361–5375.
27. **Murakawa, G. J., and D. P. Nierlich.** 1989. Mapping the *lacZ* mRNA ribosome binding site by RNA footprinting. Biochemistry **28:**8067–8072.

27a. **Murakawa, G. J., and D. P. Nierlich.** Unpublished data.

28. **Nierlich, D. P., C. Kwan, G. J. Murakawa, P. Mahoney, A. W. Ung, and D. Caprioglio.** 1985. Intercistronic sites in the *lac* operon, p. 185–193. *In* M. Schaechter, F. C. Neidhardt, and J. L. Ingraham (ed.), The Molecular biology of bacterial growth. Jones and Bartlett, Boston.
29. **Reddy, P., A. Peterkofsky, and K. McKenney.** 1985. Translation efficiency of the *Escherichia coli* adenylate cyclase gene: mutating the UUG initiation codon to GUG or AUG results in increased gene expression. Proc. Natl. Acad. Sci. USA **82:** 5656–5660.
30. **Reznikoff, W. S., C. A. Michels, T. G. Cooper, A. E. Silverstone, and B. Magasanik.** 1974. Inhibition of *lacZ* gene translation in

trp-lac fusion strains, J. Bacteriol. **117**:1231–1239.

31. **Ross, G., and C. O'Callaghan.** 1975. β-Lactamase assays. Methods Enzymol. **43**:69–85.
32. **Schmidt, M. C., and M. J. Chamberlin.** 1987. *nusA* protein of *Escherichia coli* is an efficient transcription termination factor for certain termination sites. J. Mol. Biol. **195**:809–818.
33. **Stefano, J. E., and J. D. Gralla.** 1980. Kinetic investigation of the mechanism of RNA polymerase binding to mutant *lac* promoters. J. Biol. Chem. **255**:10423–10430.
34. **Subbarao, M. N., and D. Kennell.** 1985. Evidence for endonucleolytic cleavage in decay of *lacZ* and *lacI* mRNAs. J. Bacteriol. **170**:2860–2865.

34a. **Thai, T., and D. P. Nierlich.** Unpublished data.

35. **Thomas, P. S.** 1980. Hybridization of denatured RNA and small DNA fragments transferred to nitrocellulose. Proc. Natl. Acad. Sci. USA **77**:5201–5205.
36. **Tinoco, I., P. N. Borer, B. Dengler, M. D. Levine, O. C. Uhlenbeck, D. M. Crothers, and J. Gralla.** 1973. Improved estimation of secondary structure in ribonucleic acids. Nature (London) New Biol. **246**:40–41.
37. **Ullmann, E. Joseph, and A. Danchin.** 1979. Cyclic AMP as a modulator of polarity in polycistronic transcriptional units. Proc. Natl. Acad. Sci. USA **76**:3194–3197.

37a. **Wei, H.-Y.** 1990. M.S. thesis. University of California, Los Angeles.

38. **Yager, T. D., and P. H. von Hippel.** 1987. Transcript elongation and termination in *Escherichia coli*, p. 1241–1275. *In* F. C. Neidhardt, J. L. Ingraham, K. B. Low, B. Magasanik, M. Schaechter, and H. E. Umbarger (ed.), *Escherichia coli* and *Salmonella typhimurium*: cellular and molecular biology. American Society for Microbiology, Washington, D.C.
39. **Yu, X.-M., L. M. Munson, and W. S. Resnikoff.** 1984. Molecular cloning and sequence analysis of *trp-lac* fusion deletions. J. Mol. Biol. **172**:355–362.
40. **Zabin, I., and A. V. Fowler.** 1970. β-Galactosidase and thiogalactoside transacetylase, p. 27–47. *In* J. R. Beckwith and D. Zipser (ed.), The lactose operon. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.