Proc. Natl. Acad. Sci. USA
Vol. 76, No. 10, pp. 4941–4945, October 1979
Biochemistry

leu operon of *Salmonella typhimurium* is controlled by an attenuation mechanism

(transcription termination/leader RNA/transcription coupled to translation)

ROBERT M. GEMMILL, SUSAN R. WESSLER, ELIZABETH B. KELLER, AND JOSEPH M. CALVO*

Section of Biochemistry, Molecular, and Cell Biology, Cornell University, Ithaca, New York 14853

Communicated by H. Edwin Umbarger, July 9, 1979

ABSTRACT **The nucleotide sequence of the control region of the *leu* operon of *Salmonella typhimurium* was determined. A prominent feature of this region is a signal for termination of transcription. *In vitro*, transcription does terminate at this site, yielding a leader RNA of about 160 nucleotides as a major product. This leader RNA is potentially translatable into a peptide containing 28 amino acids, 4 of which are adjacent leucine residues. Several regions of base complementarity exist within the leader, positioned such that pairing of one region precludes pairing of another. The position of the four leucine codons relative to two regions of base complementarity suggest a model for the regulation of the *leu* operon similar to that proposed by Yanofsky and coworkers for the *trp* operon. In addition, a third region of base complementarity was identified which, when incorporated into the model, explains why premature termination is the usual outcome when transcription is initiated *in vitro* by purified RNA polymerase.**

Genes specific to leucine biosynthesis are clustered on the chromosome of *Salmonella typhimurium* (1) and their expression is controlled by the intracellular concentration of leucine (2). Several classes of mutations affect expression of *leu* genes in a coordinate fashion. Promoter mutations (3) eliminate expression of all four genes whereas so-called operator mutations (4) result in constitutive expression of *leuA*, *-B*, *-C*, and *-D*. These mutations define the control region of the *leu* operon. In addition, mutations unlinked to *leu*, in *leuS* and *flrB*, lead to constitutive synthesis of *leu* and some *ilv* gene products (5).

The *leu* operon of *S. typhimurium* was cloned on plasmids pSC101 and pMB9 and sites at which restriction endonucleases cleave the operon were determined and correlated with a genetic map (unpublished data). In this paper we report the sequence of nucleotides in the *leu* operon control region. It is apparent from the sequence and from the results of *in vitro* transcription studies that the *leu* operon is controlled by an attenuation mechanism similar to that described by Yanofsky and his coworkers for the *trp* operon (6, 7).

METHODS

Sequencing Techniques. Conditions for plasmid DNA isolation (8), treatment with restriction endonucleases (9), and separation of DNA fragments by electrophoresis (9) were as cited. Fragments 1–6 identified by arrows in Fig. 1*B* were isolated as follows. The smaller of the two fragments produced by *Pst* I cleavage of *Eco*RI$_{5200}$, *Pst* I-B, was cleaved with *Hpa* II, the ends were labeled with ^{32}P by sequential treatment with bacterial alkaline phosphatase and polynucleotide kinase together with [γ-^{32}P]ATP (10), and the resulting mixture was treated sequentially with *Hin* II and *Hae* II. Acrylamide gel electrophoresis yielded fragment 1 (*Hin* II/*Hpa* II$_{126}$) and fragment 2 (*Hpa* II/*Hae* II$_{88}$), each labeled only at the end created by *Hpa* II. *Pst* I-B was cleaved with *Taq* I, ends were labeled by means of kinase, and, after treatment with *Hae* III, fragments 3 (*Hae* III/*Taq* I$_{320}$) and 4 (*Taq* I/*Hae* III$_{140}$) were isolated. Fragments 5 and 6 are identical to 3 and 4, respectively, except that labeling was effected at 3′ ends by means of reverse transcriptase (a gift from J. Beard) and [α-^{32}P]dCTP (11). DNA sequencing was performed by the method of Maxam and Gilbert (12) with pyridinium formate for the A+G cleavage reaction (A. Maxam, personal communication). All procedures using recombinant DNA were performed under P2 conditions in accordance with the National Institutes of Health guidelines.

***In Vitro* Transcriptions.** Conditions for *in vitro* transcription were modified slightly from those used by Lee and Yanofsky (6). Reaction mixtures (total volume, 100 μl) contained: Tris acetate (pH 7.9), 20 mM; KCl, 100 mM; magnesium acetate, 4 mM; EDTA, 100 μM; dithiothreitol, 100 μM; GTP, 125 μM; ATP, 125 μM; UTP, 75 μM; CTP, 75 μM; [α-^{32}P]UTP, 0.12 nmol [186 Ci/mmol (1 Ci = 3.7 × 10^{10} becquerels)]; RNA polymerase (either New England BioLabs or a gift from J. Alegre), 2.0 μg; and DNA, 2 μg (plasmid) or 0.5 μg (fragment). After incubation for 30 min at 37°C, RNA was isolated (6), fractionated by electrophoresis on 0.4-mm-thick 8% acrylamide gels containing 8 M urea (13), and visualized by autoradiography.

RESULTS

Identification of DNA Fragments Carrying the Control Region of the *leu* Operon. An *Eco*RI-generated fragment of *Salmonella* DNA cloned on plasmid pMB9 (yielding plasmid pCV12) carries *leuABC* and part of *leuD* together with a functional *leu* promoter and control region (unpublished data). A restriction endonuclease map and the genetic map of this fragment are shown in Fig. 1 *A* and *D*, respectively. Correlation between the two maps was achieved by analyzing the operon from strains having deletions within *leu* (detailed results to be published elsewhere). From a genetic analysis, one end of deletion *leu-4168* (Fig. 1*C*) lies between known promoter mutations and operator mutations (unpublished results). Deletion *leu-5058*, on the other hand, extends through the promoter/operator region and just enters *leuA* (it recombines with 270 of 272 mutations in *leuA*) (14). Based upon hybridization of *Hin* II- and *Pst* I-cleaved total *Salmonella* DNA to labeled *leu* DNA by procedures developed by Southern (15), both *leu-4168* and *leu-5058* end within fragment *Hin* II-C and to the left of the *Pst* I site (Fig. 1*A*). A more precise determination of the end point of *leu-4168* was obtained by cloning the operon from the strain and comparing endonuclease-generated frag-

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.

* To whom reprint requests should be addressed.

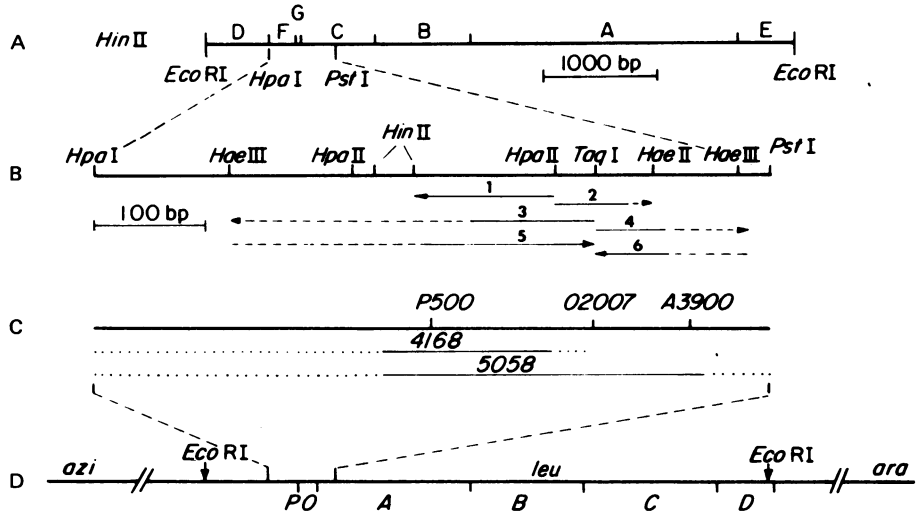

FIG. 1. DNA fragments from the *leu* control region chosen for sequence determination. (*A*) *Hin* II map of the DNA fragment generated by *Eco*RI treatment of plasmid pCV12. *Hin* II fragments are designated A (largest) through G (smallest). bp, Base pairs. (*B*) Detailed restriction map of region bounded by *Hpa* I and *Pst* I sites. Arrows denote the fragments that were sequenced and the extent of sequencing (solid portion). The direction of the arrow indicates whether the relevant sequence was of the coding strand (arrow points leftward) or the noncoding strand (points rightward). (*C*) Genetic map of the *leu* control region expanded to the same scale as in *B*. The locations of sites *leuP500*, *leu02007*, and *leuA3900* are only approximate. Dots indicate the uncertainty in the end points of deletions 4168 and 5058. (*D*) Genetic map of the *leu* operon.

ments from it with those generated from the wild-type operon. From this analysis, the *leu-4168* deletion was found to terminate between the *Hpa* II and *Hae* II sites of fragment *Hin* II-C shown in Fig. 1*B*.

Sequence of Nucleotides in the *leu* Control Region. The sequence of DNA fragments identified by arrows in Fig. 1*B* was determined by the method of Maxam and Gilbert (12). The sequence of 216 bases between the *Hae* II site in fragment *Hin* II-C and the closest *Hin* II site is shown in Fig. 2. The first base pair of the *Hin* II site is arbitrarily designated position 1. A prominent feature of this region is a sequence of 7 consecutive T residues between positions 191 and 197; to the left of this sequence and overlapping it is an 11-base-pair palindromic sequence composed predominantly of G and C residues. These palindromic and T_7 sequences together may comprise a recognition element for termination of transcription (16).

Transcription Initiated *In Vitro* at the *leu* Promoter Yields a Short RNA Transcript as a Major Product. Plasmids pCV12 and pCV23 are identical except that the latter has promoter mutation *leu-500* (unpublished data). Strains carrying *leu-500* produce only barely detectable levels of *leu* enzymes (1). DNA from these plasmids served as templates for *in vitro* transcrip-

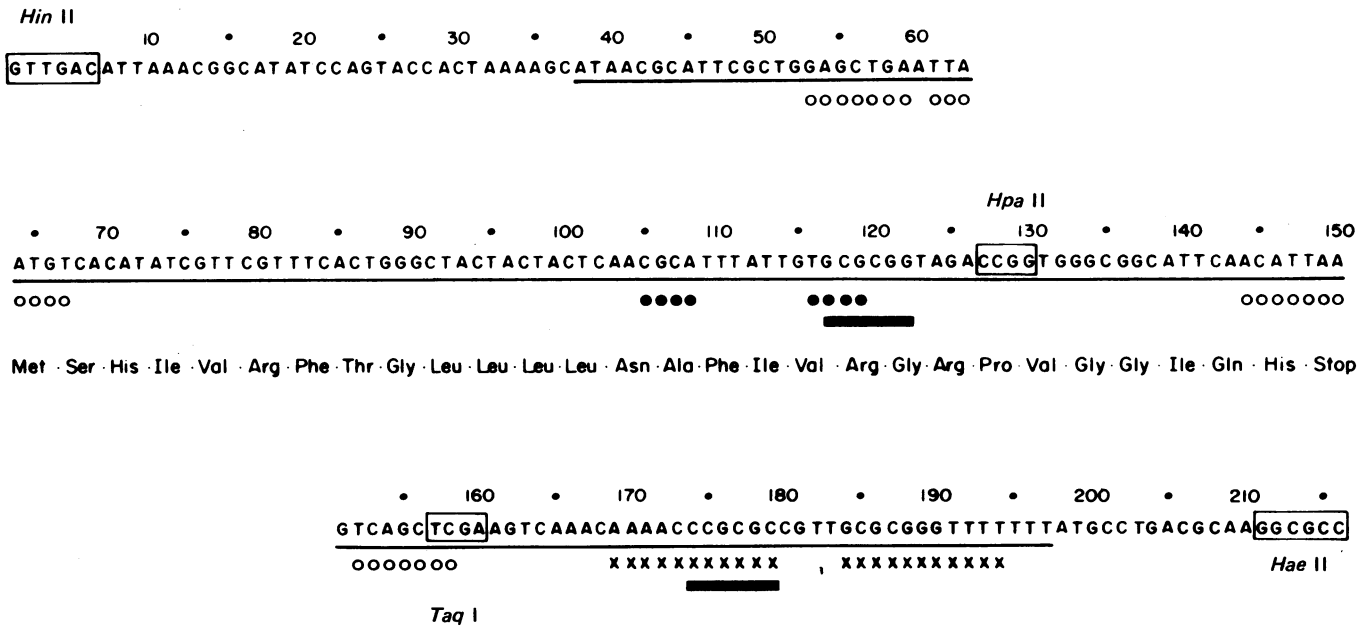

FIG. 2. Nucleotide sequence of the noncoding strand of the *leu* control region. Continuous line under the sequence represents the corresponding leader RNA, the exact boundaries of which have not yet been determined. The sequence of a putative 28-amino acid peptide specified by the leader is indicated. The boxed regions are recognition sites for the indicated restriction endonucleases. Three regions of potential secondary structure within the putative leader RNA are indicated: the terminator (X's, 169–179 pairs with 184–194), the terminator preemptor (black bars, 117–122 pairs with 174–179), and the terminator protector (open circles, 53–59 and 61–67 pair with 144–150 and 152–158; black circles, 105–108 pairs with 116–119).

tion experiments using purified RNA polymerase, and the products of these reactions were analyzed by acrylamide gel electrophoresis (Fig. 3, compare lanes A and B). A prominent transcript of about 165 nucleotides was evident when pCV12 DNA served as template. Synthesis of this transcript was initiated at *leuP* because mutation *leu-500* prevented its formation. The following evidence indicates that this RNA species was transcribed from just that region of DNA that was sequenced. (*i*) It was the major product synthesized *in vitro* when the *Hae* III fragment shown in Fig. 1*B* was used as template (Fig. 3, compare lanes C and D). (*ii*) Cutting the *Hae* III fragment with *Hae* II did not abolish its synthesis whereas cutting with *Hin* II or separately with *Hpa* II did (data not shown). (*iii*) The relevant RNA hybridized to the *Hpa* II_{180} and the *Hpa* II/*Pst* I_{190} fragments defined in Fig. 1*B* but not to the fragment extending from the left *Eco*RI site to the *Hpa* II site in *Hin* II-F (Fig. 1*A*) (data not shown).

These results demonstrate that the relevant RNA was transcribed from a region lying between the *Hin* II and *Hae* II sites in Fig. 1*B*, exactly the region that was sequenced. With the *Hae* III fragment as template, a readthrough transcript of about 250 nucleotides was expected whose synthesis originated at *leuP* and terminated at the right end of the *Hae* III fragment. Several minor species of about that size were visible on the autoradiograms but their synthesis was not abolished by the *leu-500* promoter mutation. Lee and Yanofsky (6) observed that the amount of *trp* readthrough RNA synthesized *in vitro* was dramatically increased when ITP replaced GTP. Under conditions of *in vitro* synthesis similar to theirs, the most prominent product was comparable in size to the putative readthrough RNA (Fig. 3, lane E). There was also a minor product corresponding in size to the major species observed when reactions contained GTP. Neither species was formed when the template contained the *leu-500* mutation (Fig. 3, lane F). These results suggest that *leu* RNA synthesis is initiated about 250 base pairs from the right end of the *Hae* III fragment (Fig. 1*B*).

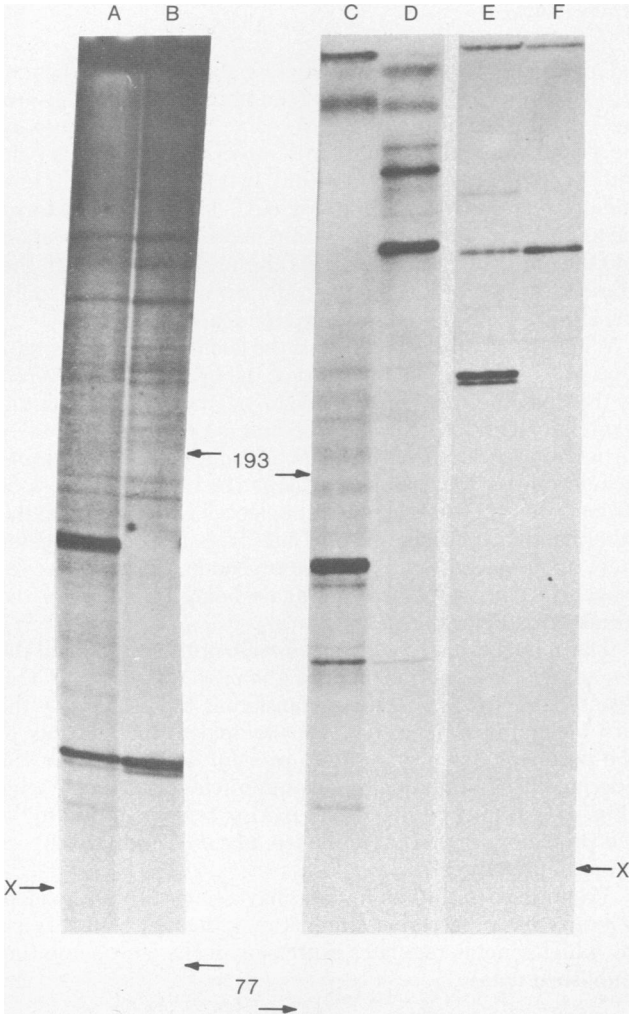

FIG. 3. *In vitro* transcription from *leu*-containing templates. The position of the xylene cyanol marker is denoted by an X. RNA molecules (4S and 6S) synthesized from phage λ in the same experiment were used as markers. Their position and approximate size (nucleotides) are indicated (16). Lanes: A, pCV12; B, pCV23; C and D, fragment *Hae* III_{460} from pCV12 and pCV23, respectively; E and F, same as C and D except that ITP replaced GTP.

DISCUSSION

The DNA molecules that were sequenced were known from genetic and restriction endonuclease analysis to contain the *leu* promoter and control region. A striking feature of the derived sequence is a signal for termination of transcription (16), hereafter referred to as the *leu* attenuator. The presence of a functional attenuator was confirmed by the observation of a 165-nucleotide RNA molecule as a primary product of *in vitro* transcription. The following evidence indicates that the synthesis of this RNA transcript was initiated at the *leu* promoter and terminated at the *leu* attenuator. (*i*) A *leu* promoter mutation abolished *in vitro* synthesis of the transcript. (*ii*) *In vitro* synthesis programmed by various templates and hybridization studies demonstrated that the RNA molecule was transcribed from just the region of DNA that was sequenced. (*iii*) The major product of transcription from the *Hae* III fragment in the presence of ITP was about 250 nucleotides long. Its synthesis was initiated at *leuP* and most likely terminated at the right end of the fragment. This places the point of transcription initiation about 250 base pairs from the right end of the *Hae* III fragment, near position 35 in Fig. 2. Between positions 1 and 30 there is a potential region of interaction with RNA polymerase (17). (*iv*) In the presence of GTP, synthesis of the major RNA product terminated at a point about 165 base pairs from *leuP*—that is, near position 195 (Fig. 2). This corresponds to the position of the proposed attentuation site. The strand shown in Fig. 2 is demonstrated to be the noncoding strand on the basis of the following results. The ATG that signals initiation of translation of α-IPM synthase (*leuA* gene product) is 41 bases downstream from the attenuator site (position 238 using the numbering system in Fig. 2). This result was deduced from additional nucleotide sequences (data not shown) together with the sequence of 10 amino acids at the NH_2 terminus of α-IPM synthase (unpublished data).

The leader transcript was the predominant product of transcription when a *Hae* III fragment was used as substrate. A putative readthrough transcript was not observed, indicating that termination of transcription was extremely efficient under the conditions of these experiments. Presumably, termination of transcription is not so efficient *in vivo*, at least not under conditions of leucine limitation in which the amount of *leu* mRNA is increased as much as 20-fold relative to conditions of leucine sufficiency (ref. 18; unpublished data). A mechanism by which termination of transcription might be suppressed *in vivo* has been suggested by Zurawski *et al.* (7). This mechanism focuses on certain secondary structures within the *trp* and *phe* leader RNAs that affect the frequency of termination at the respective attenuators and on the relationship of translation of the leader RNAs to transcription termination. The *leu* leader contains a single AUG codon centered at position 65, together

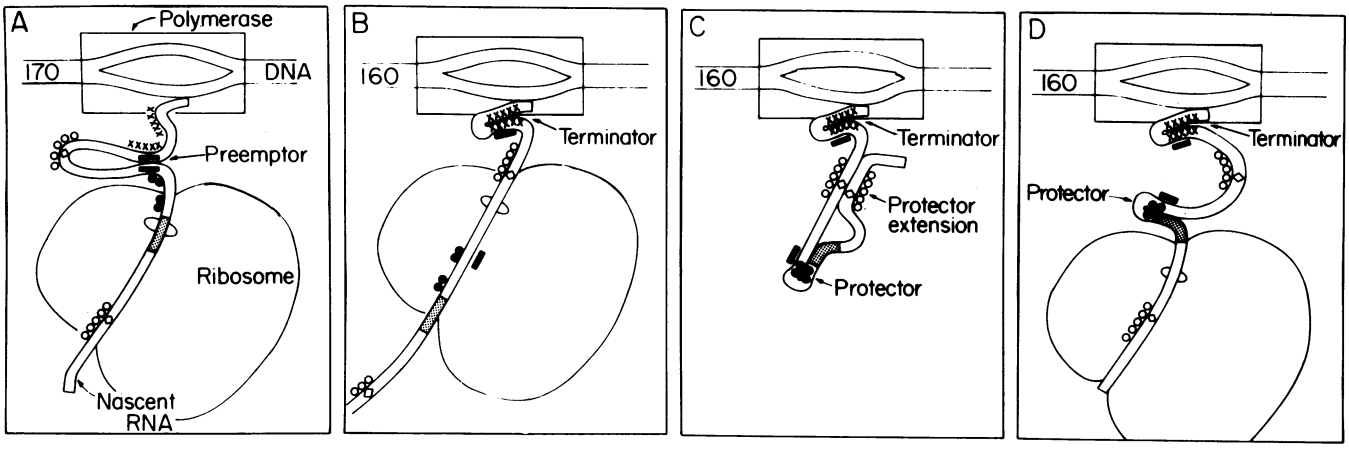

FIG. 4. Diagrams of events postulated to occur during coupled transcription and translation of the *leu* leader. DNA, RNA, RNA polymerase, and ribosomes are drawn approximately to scale but the shapes are schematic. ●, ○, Protector region; ■, preemptor region; X, terminator region. The number of transcribed base pairs is indicated just to the left of the RNA polymerase. The number of bases that cannot form a secondary structure because of an interaction with the ribosome was arbitrarily chosen as 40. The A site on the ribosome is denoted by an oval. (*A*) Readthrough situation. With ribosome stalled at leucine codons because of a leucine limitation, pairing of the preemptor prevents pairing of the terminator. (*B* and *C*) Termination situation. Under conditions of leucine excess, a ribosome has approached the stop codon (*B*) and subsequently fallen off (*C*). Continued polymerase movement in either case allows pairing of the terminator regions and results in transcription termination. The RNA structure depicted in *C* is also expected to form when transcription occurs in the absence of translation. (*D*) Termination situation resulting from ribosome being stalled at one of the codons preceding the leucine codons.

with a purine-rich region including a G-G-A-G sequence centered at position 54 that might constitute a ribosome binding site (19). Chain-terminating codons are centered at positions 124 (UAG) and 149 (UAA), and the latter is in the same translational reading frame as the AUG. The *leu* leader could code for a peptide having 28 amino acids, 4 of which are adjacent leucines (Fig. 2). This region of the *leu* leader capable of coding for a peptide may be compared with that for *trp* (14 amino acids, 2 adjacent tryptophans) (6), *his* (16 amino acids, 7 adjacent histidines) (20, 21), *phe* (14 amino acids, 7 phenylalanines) (22), and *thr* (21 amino acids, 8 threonines, and 4 isoleucines) (23).

Regions of secondary structure that might form within the *leu* leader were deduced by an analysis developed by one of us (unpublished). Three regions of strong base complementarity exist within the leader, positioned such that pairing of one region precludes pairing of another. The strongest of these regions (−23 kcal) (24, 25), is called the terminator (6, 20–23) because it is thought to signal transcription termination upon forming a stem and loop structure (Fig. 2, xxx). An alternative region, indicated in Fig. 2 by solid bars, is termed the preemptor because, upon pairing, it preempts formation of the terminator. Similar terminator and preemptor regions have been identified for the *trp* (6), *phe* (22), *his* (20, 21), and *thr* (23) operons. We also identified a third strong pairing region which we call the *protector* because, upon pairing, it precludes formation of the preemptor, thus protecting the terminator (Fig. 2, circles). The protector is made up of two parts—a weaker region near the four leucine codons (the protector itself, indicated by solid circles), and a stronger region which we view as an extension of the protector (open circles). The following analysis, derived from a consideration of events occurring during transcription and translation, suggests that, under most conditions, terminator regions will pair, resulting in transcription termination. This analysis is similar but not identical to that developed by Lee and Yanofsky (6).

When about 75 base pairs of the *leu* leader have been transcribed, a ribosome may initiate translation and, moving at the same rate as the polymerase (26, 27) will follow the polymerase as the latter moves along the template. If leucyl-tRNA^{Leu} is limiting, the ribosome will pause or be arrested at the leucine codons and lag behind the polymerase. At the point shown in Fig. 4*A*, protector regions have been transcribed but they are prevented from pairing by the ribosome. However, as soon as the second half of the preemptor is transcribed, it is free to pair with its complementary region and, in so doing, prevents terminator regions from pairing (Fig. 4*A*). The inability to form the termination stem and loop results in continued transcription into the *leu* structural genes. This circumstance, in which the ribosome pauses at the leucine codons, we believe is the only one allowing substantial transcription readthrough.

When the ribosome moves past the leucine condons (conditions of leucine sufficiency), part of the preemptor is covered by the ribosome and is unavailable to preempt formation of the terminator (Fig. 4*B*). If the ribosome pauses near the UAA termination codon for even a fraction of a second, that is sufficient time for the termination loop to be formed (RNA elongation rate at 37°C is about 45 nucleotides/sec) (27). On the other hand, if the ribosome separates from the RNA immediately upon reaching the termination codon, the protector regions will pair, with the result again being formation of the termination loop (Fig. 4*C*).

This model explains why termination of transcription at the *leu* attenuator is so efficient during *in vitro* transcription. This is so because in the absence of translation, the protector is the first region to pair (Fig. 4*C*). The physiological significance of the protector region may be to prevent derepression of the operon during limitation for an amino acid other than leucine (Fig. 4*D*). It may be noted that pairing regions similar to the *leu* protector are present in the *trp*, *phe*, *his*, and *thr* (23) operons (our analysis).

Preliminary results on mutants having constitutive levels of *leu* enzymes suggest that attenuation is the major mechanism by which leucine regulates expression of the *leu* operon (unpublished data).

Note Added in Proof. The unpublished data cited in the *Introduction* and *Results* are reported in ref. 28. Several investigators have independently developed a model for control of transcription termination that includes as a basic feature a pairing region within the leader RNA that is analogous to the protector region described here (C. Yanofsky; J. Roth; personal communication).

We thank V. Vogt and J. Roberts for helpful discussions. This research was supported by grants from the National Science Foundation (BMS74-21987 to J.M.C.) and the National Institutes of Health (GM10791 to E.B.K. and AI14340 to J.M.C.). R.M.G. and S.R.W. were supported by National Institutes of Health Training Grant GM07273.

1. Margolin, P. (1963) *Genetics* **48,** 441–457.
2. Calvo, J. M., Freundlich, M. & Umbarger, H. E. (1969) *J. Bacteriol.* **97,** 1272–1282.
3. Dubnau, E., Lenny, A. B. & Margolin, P. (1973) *Mol. Gen. Genet.* **126,** 191–200.
4. Calvo, J. M., Margolin, P. & Umbarger, H. E. (1969) *Genetics* **61,** 777–787.
5. Friedberg, D., Mikulka, T. W., Jones, J. & Calvo, J. M. (1974) *J. Bacteriol.* **118,** 942–951.
6. Lee, F. & Yanofsky, C. (1977) *Proc. Natl. Acad. Sci. USA* **74,** 4365–4369.
7. Zurawski, G., Elseviers, D., Stauffer, G. V. & Yanofsky, C. (1978) *Proc. Natl. Acad. Sci. USA* **75,** 5988–5992.
8. Kupersztoch-Portnoy, Y. M., Lovett, M. A. & Helinski, D. R. (1974) *Biochemistry* **13,** 5484–5490.
9. Wu, R., Jay, E. & Roychoudhury, R. (1976) in *Methods in Cancer Research*, ed. Busch, H. (Academic, New York), Vol. 12, pp. 87–176.
10. Lillehaug, J. R. & Kleppe, K. (1975) *Biochemistry* **14,** 1225–1229.
11. Bahl, C., Wu, R., Strawinsky, J. & Narang, S. (1977) *Proc. Natl. Acad. Sci. USA* **74,** 966–970.
12. Maxam, A. M. & Gilbert, W. (1977) *Proc. Natl. Acad. Sci. USA* **74,** 560–564.
13. Maniatis, T., Jeffrey, A. & van deSande, H. (1975) *Biochemistry* **14,** 3787–3794.
14. Calvo, J. M. & Worden, H. E. (1970) *Genetics* **64,** 199–214.
15. Southern, E. M. (1975) *J. Mol. Biol.* **98,** 503–517.
16. Adhya, S. & Gottesman, M. (1978) *Annu. Rev. Biochem.* **47,** 967–996.
17. McConnell, D. J. (1979) *Nucleic Acid Res.* **6,** 525–544.
18. Davis, M. G. & Calvo, J. M. (1977) *J. Bacteriol.* **131,** 997–1007.
19. Steitz, J. A. & Jakes, K. (1975) *Proc. Natl. Acad. Sci. USA* **72,** 4734–4738.
20. DiNocera, P. P., Blasi, F., DiLauro, R., Frunzio, R. & Bruni, C. B. (1978) *Proc. Natl. Acad. Sci. USA* **75,** 4276–4280.
21. Barnes, W. M. (1978) *Proc. Natl. Acad. Sci. USA* **75,** 4281–4285.
22. Zurawski, G., Brown, K., Killingly, D. & Yanofsky, C. (1978) *Proc. Natl. Acad. Sci. USA* **75,** 4271–4275.
23. Gardner, J. F. (1979) *Proc. Natl. Acad. Sci. USA* **76,** 1706–1710.
24. Tinoco, I., Jr., Borer, P. N., Dengler, B., Levine, M. D., Uhlenbeck, O. C., Crothers, D. M. & Gralla, J. (1973) *Nature (London) New Biol.* **246,** 40–41.
25. Borer, P. N., Dengler, B., Tinoco, I., Jr. & Uhlenbeck, O. C. (1974) *J. Mol. Biol.* **86,** 843–853.
26. Lacroute, F. & Stent, G. S. (1968) *J. Mol. Biol.* **35,** 165–173.
27. Rose, J. K., Mosteller, R. D. & Yanofsky, C. (1970) *J. Mol. Biol.* **51,** 541–550.
28. Hertzberg, K. M., Gemmill, R. M., Jones, J. & Calvo, J. M. (1979) *Gene*, in press.