

In vitro synthesis of the tryptophan operon leader peptides of *Escherichia coli*, *Serratia marcescens*, and *Salmonella typhimurium*

(attenuation/leader peptide synthesis/translation inhibition)

ANATHBANDHU DAS*, JOSEPH URBANOWSKI†, HERBERT WEISSBACH†, JOHN NESTOR‡, AND CHARLES YANOFSKY*

*Department of Biological Sciences, Stanford University, Stanford, California 94305; †Department of Biochemistry, Roche Institute of Molecular Biology, Roche Research Center, Nutley, New Jersey 07110; and ‡Institute of BioOrganic Chemistry, Syntex Research, Palo Alto, California 94304

Contributed by Charles Yanofsky, February 14, 1983

ABSTRACT We used an *in vitro* DNA-dependent protein-synthesizing system to demonstrate *de novo* synthesis of the leader peptide specified by the tryptophan (*trp*) operons of several bacterial species. Peptide synthesis was directed by self-ligated short restriction fragments containing the *trp* promoter and leader regions. Synthesis of leader peptides was established by demonstrating that they were labeled *in vitro* only by those amino acids predicted to be present in the peptides. Leader peptide synthesis was abolished by the addition of the *Escherichia coli* *trp* repressor. The *E. coli* *trp* leader peptide was found to be extremely labile *in vitro*; it had a half-life of 3–4 min. In a highly purified DNA-dependent peptide-synthesizing system, synthesis of the di- and tripeptides predicted from the *Salmonella typhimurium* *trp* operon leader sequence, fMet-Ala and fMet-Ala-Ala, also was observed. Using this dipeptide synthesis system, we demonstrated that translation initiation at the ribosome binding site used for *trp* leader peptide synthesis was reduced 10-fold when the transcript contained a segment complementary to the ribosome binding site.

Many of the amino acid biosynthetic operons of bacteria are regulated by attenuation. Each of these operons has a transcribed leader region containing a short peptide-encoding segment rich in codons for the regulatory amino acid(s) (1–10). The presumptive leader peptides specified by these coding regions have been sought but have eluded detection. Nevertheless, gene fusion studies have demonstrated that the ribosome binding site of the *trp* operon leader region is an efficient site for translation initiation (11, 12). We considered two explanations for our inability to detect the *trp* leader peptide. First, because synthesis of the peptide but not its survival was presumably of regulatory significance, we thought that the peptide might be rapidly degraded both *in vivo* and *in vitro* (11, 13). Second, scrutiny of the sequences of the *trp* leader transcripts of seven enteric bacterial species revealed that the distal segment of each leader transcript was complementary to the leader peptide ribosome binding site and conceivably could pair with that site and prevent multiple rounds of synthesis of the *trp* leader peptide (refs. 14–16; unpublished observations). The approach that was successful in the initial detection of the *trp* leader peptide, described herein, was suggested by Roberto Kolter of our laboratory.

We performed *in vitro* coupled transcription/translation analyses, using as template a short restriction fragment containing the *trp* promoter-operator and leader peptide-encoding region and lacking the distal portion of the *trp* leader region that is complementary to the leader peptide ribosome binding

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.

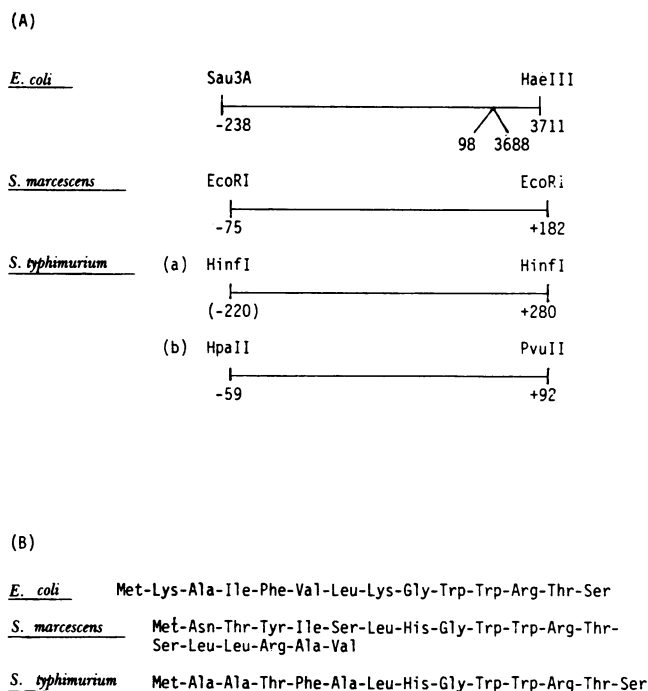


FIG. 1. (A) Physical maps of the *trp* operon fragments used (not to scale). (B) Amino acid sequences of the *trp* leader peptides of *E. coli*, *S. marcescens*, and *S. typhimurium*. Nucleotide numbering is from the site of transcription initiation (residue +1). Negative numbers indicate the number of base pairs preceding the transcription start point. Numbers within parentheses indicate that the exact site is not known. The *E. coli* template has a leader-*trpC* deletion which fuses nucleotides 98 and 3,688 of the *trp* operon (25). This deletion removes the leader termination (attenuator) region (nucleotides 114–140) of the operon. The *S. marcescens* *trp* operon segment has the wild-type promoter-operator-attenuator region. The *S. typhimurium* fragment a also has the wild-type *trp* promoter-operator-attenuator region. *S. typhimurium* fragment b, an internal fragment of fragment a, does not contain the attenuator region (nucleotides 117–142) or the distal region (nucleotides 102–140) that is complementary to the Shine-Dalgarno region (26) and start codon presumably used in *trp* leader peptide synthesis (see also Fig. 4).

site. This restriction fragment was self-ligated, and the circular molecules that formed were used as template. We reasoned that with this template the pattern of polypeptide synthesis should be simple because a single coding region would be directing *in vitro* peptide synthesis. In addition, the molar concentration of the template could be increased appreciably. Furthermore, the absence of the complementary RNA region would permit multiple rounds of translation of the leader transcript.

In this paper we describe application of the above approach

in the detection of the synthesis of the leader peptides of the *trp* operons of *Escherichia coli*, *Serratia marcescens*, and *Salmonella typhimurium*. We also describe experiments with a highly defined system for di- and tripeptide synthesis (17–19) that demonstrate that our templates direct efficient synthesis of the expected NH₂-terminal peptides. We also show that dipeptide synthesis is reduced when the template contains the DNA region corresponding to the RNA segment that is complementary to the leader peptide ribosome binding site.

MATERIALS AND METHODS

Materials. L-[³⁵S]Methionine (150 Ci/mmol; 1 Ci = 3.7 × 10¹⁰ Bq), L-[³H]proline (29.5 Ci/mmol), and L-[³H]leucine (60 Ci/mmol) were from New England Nuclear. L-[³H]Tryptophan (25 Ci/mmol), L-[³H]alanine (75 Ci/mmol), L-[³H]arginine (17 Ci/mmol), L-[³H]histidine (51 Ci/mmol), and L-[³H]lysine (33 Ci/mmol) were from Amersham. Restriction endonucleases were from Bethesda Research Laboratories and New England BioLabs. *N*-Formyl-L-methionine and *N*-formyl-L-methionyl-L-alanine were from Sigma.

Plasmids. DNA containing the *E. coli trp* promoter-operator and leader peptide-encoding region was isolated as a 360-base-pair (bp) *Sau3A-Hae* III restriction fragment from plasmid pGM22. This plasmid has a leader-*trpC* deletion that removes the sequences immediately beyond the peptide-encoding region. The fragment, after filling in, was cloned into the *Sma* I site of plasmid pUC8 (20) (the resulting plasmid was designated pAD2). The *S. typhimurium trp* template that lacks the complementary region was isolated as a 151-bp *Hpa* II-*Pvu* II fragment. This fragment was obtained by digestion of a 500-bp *Hinf*I fragment of plasmid pKB5 (21). Both *Hinf*I₅₀₀ and *Hpa* II-*Pvu* II₁₅₁ fragments, after filling in and ligating with *Hind*III linkers, were cloned into the *Hind*III site of pUC8 (designated pAD6 and pAD7, respectively). The *S. marcescens trpPOL* region (containing the leader peptide-encoding region and the attenuator) was isolated as a 257-bp *Eco*RI fragment (14). Plasmid pBN37, containing this *Eco*RI fragment, was kindly provided by B. Nichols (Department of Biological Sciences, Univ. of Illinois, Chicago). Plasmid DNA was prepared as described (22).

Fragments containing the *trpPOL* regions of *E. coli*, *S. mar-*

cens, and *S. typhimurium* were isolated by digestion of the above plasmids with the restriction enzymes *Eco*RI/*Bam*HI, *Eco*RI, and *Hind*III, respectively. Where indicated, these linear fragments were circularized by ligating with T4 DNA ligase.

In Vitro Protein Synthesis. Cell-free (S-30) extracts were prepared from *E. coli* strain W3110 *trpR tnaA2 ΔtrpEA2 rna*. Reaction conditions were as described (22) except that there was a preincubation period of 15 min at 37°C after which a radioactive amino acid containing 20 μCi of ³H or 2 μCi of [³⁵S]-methionine was added, and incubation was continued for an additional 5 min. Aliquots (10 μl) were analyzed on NaDodSO₄/polyacrylamide gels (15%) prepared as described by Laemmli (23). Gels were run overnight at 40–50 V. Gels were fixed for 4 hr in 25% methanol/10% acetic acid, washed in water for about 1 hr, soaked in 1 M sodium salicylate for 30 min (24), dried, and then fluorographed for 2 days at –80°C.

To analyze the stability of the *E. coli trp* leader peptide, translation products were labeled uniformly for 40 min with [³H]-tryptophan. Unlabeled tryptophan (7 μl of 20 mM, resulting in a 200-fold dilution of specific radioactivity) was then added and incubation was continued at 37°C. Aliquots (10 μl) were removed at the indicated times (see Fig. 3) and analyzed on NaDodSO₄/polyacrylamide gels as described above.

DNA-dependent di- and tripeptide synthesis was performed as described (17–19). The reaction mixture contained DNA template, purified RNA polymerase, initiation factors, elongation factor(s), ribosomes, charged tRNAs, salt, and other components. Analysis of the dipeptide fMet-Ala used an ethyl acetate extraction procedure (19), whereas TLC was used to assay for the tripeptide fMet-Ala-Ala (17, 18). Linear DNA fragments were used as templates in the purified peptide-synthesizing system. Circular, ligated fragments were used as templates in the S-30 system.

RNA synthesis was measured as described (22).

RESULTS

The *trp* operon fragments of *E. coli*, *S. marcescens*, and *S. typhimurium* used in this study are described in Fig. 1 along with the predicted amino acid sequences for the corresponding *trp* leader peptides. The *E. coli* and *S. marcescens trp* operon frag-

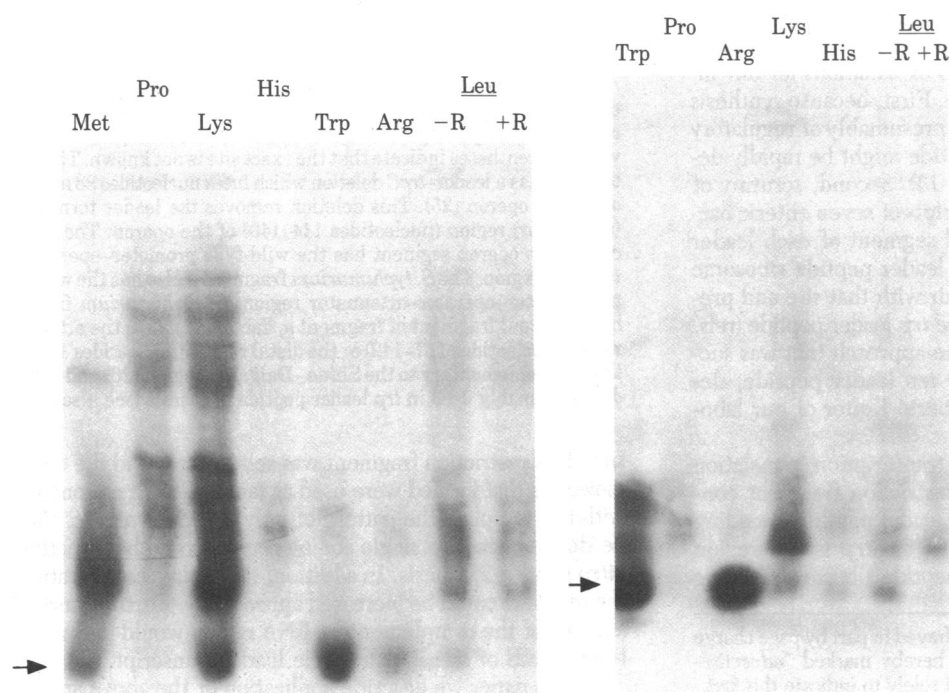


FIG. 2. Analysis of *in vitro* translation products. Translation products were analyzed as described. The template DNAs (60 ng) used were *E. coli* (A) and *S. marcescens* (B) *trp* operon segments. The radioactive amino acid added is indicated above each line. Lanes +R contained in addition 0.2 μg of purified *E. coli trp* repressor. The leader peptide band is marked by an arrow.

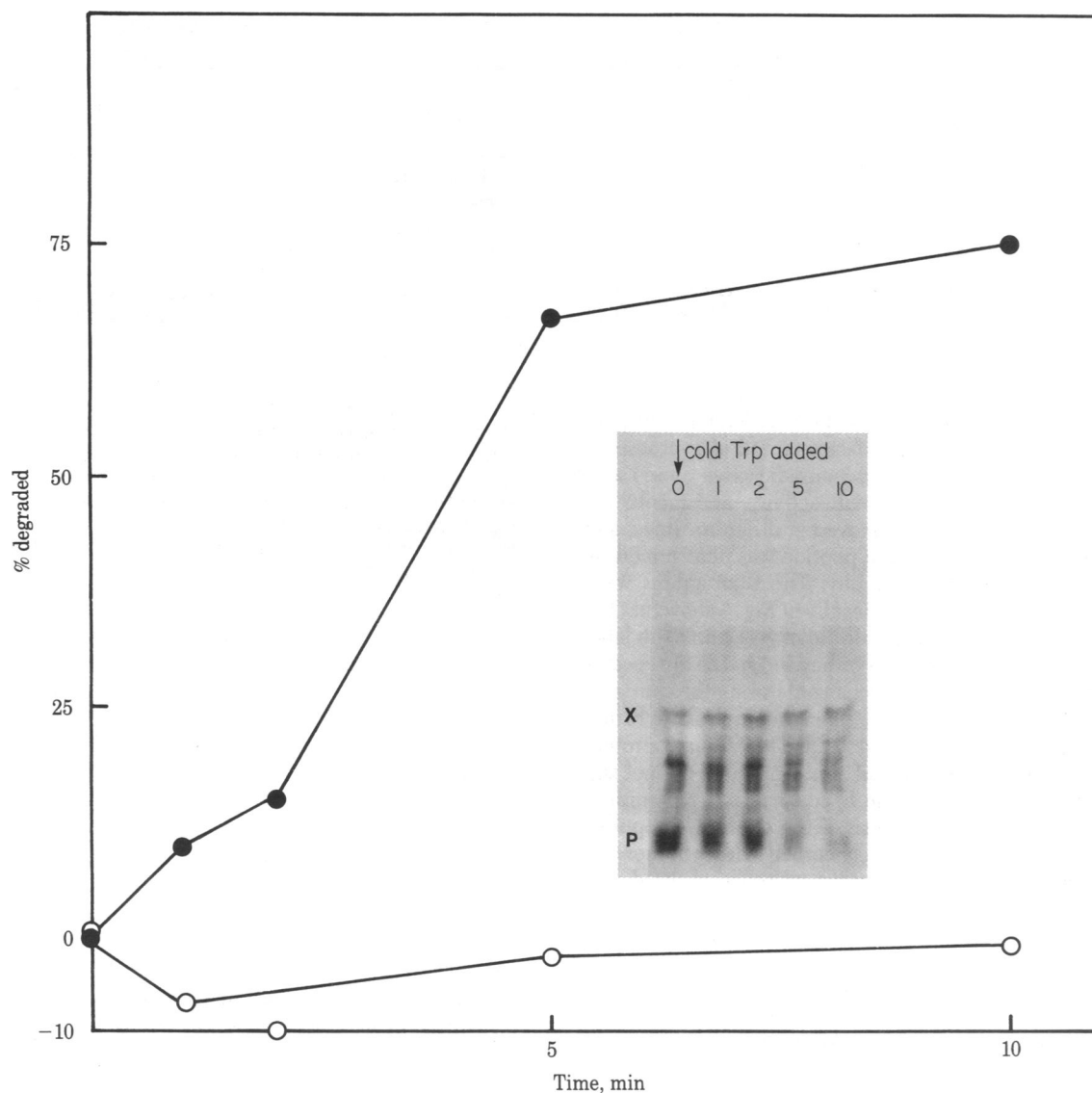


FIG. 3. Stability of the *E. coli trp* leader peptide. Peptide synthesis was performed in a total volume of 50 μ l, and 20 μ Ci of [3 H]tryptophan was used as the labeling amino acid. After 40 min of incubation at 37°C, 7 μ l of 20 mM tryptophan was added, and 10- μ l aliquots were removed at the indicated times. Samples were analyzed by NaDodSO₄/polyacrylamide gel electrophoresis and fluorography. Fluorograms (*Inset*) were scanned to quantify the bands. ●—●, *trp* leader peptide (band P); ○—○, band X. The zero time was taken as 100%.

ments were circularized and then used as templates in an *in vitro* DNA-dependent protein-synthesizing system. Analysis of the labeled translation products by NaDodSO₄/polyacrylamide gel electrophoresis revealed the synthesis of a major low molecular weight peptide that migrated similarly to synthetic peptides 15–20 residues in length and the insulin A and B chains (data not shown). The amino acid composition of the major peptide detected was determined by adding different radiolabeled amino acids to the reaction mixture. The *E. coli* peptide was labeled by methionine, lysine, tryptophan, arginine, and leucine but not by histidine or proline (Fig. 2A). Similarly, the *S. marcescens* peptide was labeled by tryptophan, arginine, histidine, and leucine but not by proline or lysine (Fig. 2B). The *S. marcescens* peptide also was labeled by methionine (data not shown). These labeling patterns were exactly those expected from the predicted sequences of the two leader peptides (Figs. 1 and 2). The labeled band in the lysine lane with the *S. marcescens* template was shown to be due to endogenous synthesis by the extract (data not shown). However, with most labeling amino acids, there was no endogenous radioactive product that

migrated similarly to the major small peptide. With both the *E. coli* and *S. marcescens* templates, peptide synthesis was abolished by the addition of *E. coli trp* repressor (Fig. 2, lanes -R and +R). These findings establish that *trp* leader peptide synthesis proceeds *in vitro*.

Initiation of synthesis of the *S. typhimurium trp* leader peptide was studied by measuring di- and tripeptide synthesis in a defined amino acid-incorporating system (17–19). Plasmid pAD7 and its *Hind*III fragment, both of which contain the coding region for the *S. typhimurium trp* leader peptide, direct the synthesis of the peptides fMet-Ala and fMet-Ala-Ala, the products predicted from the DNA sequence (Fig. 1 and Table 1). Di- and tripeptide synthesis was dependent on template DNA and fMet-tRNA^{Met} addition (results not shown). The identity of the dipeptide was verified by comparing migration with synthetic fMet-Ala on thin-layer chromatograms. The composition of the di- and tripeptide was determined in a dual-labeling experiment with [35 S]Met-tRNA^{Met} and [3 H]Ala-tRNA^{Ala} (17). It was demonstrated that the ratios of 3 H to 35 S in the dipeptide (fMet-Ala) and tripeptide (fMet-Ala-Ala) isolated on thin-layer chro-

Table 1. Template-directed di- and tripeptide synthesis

Template DNA	Amino acid content of synthesized peptides, pmol					
	fMet-Ala			fMet-Ala-Ala		
	Ala	Met	Ala/Met	Ala	Met	Ala/Met
Plasmid pAD7	4.7	3.5	1.3	1.5	0.7	2.1
pAD7-HindIII fragment	4.1	3.1	1.3	1.6	0.8	2.0

Di- and tripeptide synthesis was performed as described in refs. 17 and 19; 1 μ g of pAD7 DNA and 30 ng of pAD7-HindIII fragment DNA were used as templates.

matograms were 1.3 and 2.0, respectively (Table 1).

We also examined the stability of the *E. coli trp* leader peptide in an S-30 extract (Fig. 3). After a period of synthesis of labeled *trp* leader peptide, excess nonradioactive amino acid was added to lower the specific radioactivity, and incubation was continued. Aliquots were removed at different times, and the radioactivity remaining in the peptide was determined by gel electrophoresis and fluorography. The fluorograms were scanned to quantify the peptide. Results in Fig. 3 show that the *E. coli trp* leader peptide has a half-life of 3–4 min. The label in band X, an unknown product synthesized in the absence of any added template, remained unchanged during the course of incubation. In these experiments, phenylmethylsulfonyl fluoride, a serine protease inhibitor, was added to minimize protein degradation. However, omission of this compound did not significantly alter the half-life of the *E. coli trp* leader peptide.

Having established that synthesis of the *trp* leader peptides proceeds *in vitro*, we next were concerned with whether the yield of peptide was influenced by the presence of the distal, complementary segment of the transcript. As mentioned in the introduction, one striking structural feature of the *trp* operon leader transcripts of several bacterial species is that a distal segment can base pair with the Shine-Dalgarno sequence and start codon of the leader peptide-encoding region (Fig. 4) (refs. 14–16; unpublished data). If indeed such pairing does occur, it would inhibit ribosome binding and prevent multiple rounds of translation. To examine the possible effect of such pairing on translation initiation efficiency, we constructed plasmids that contain the *S. typhimurium trp* promoter-operator and leader peptide-encoding region that either do (plasmid pAD6) or do not (plasmid pAD7) contain the DNA segment corresponding to the distal, complementary RNA region (see Figs. 1 and 4). Appropriate restriction fragments were used as templates in

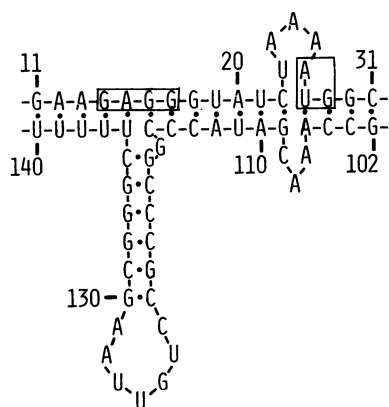


FIG. 4. Postulated pairing of the ribosome binding site of the *trp* leader transcript with a distal, complementary segment of the *S. typhimurium* transcript. The Shine-Dalgarno sequence and the start codon are boxed.

Table 2. Effect of the distal, complementary RNA segment on initiation of synthesis of the *S. typhimurium trp* leader peptide

Template DNA	mRNA, units	fMet-Ala synthesized, pmol	Peptide per unit message $\times 100$
pAD6-HindIII fragment	76	0.83	1.1
pAD7-HindIII fragment	65	7.4	11.2

Data presented above are the average of two experiments. Procedures were as described elsewhere (17, 19). [α - 32 P]UTP and [3 H]Ala-tRNA^{Ala} were used to label the RNA template and dipeptide, respectively. The incubation conditions for RNA synthesis were similar to those used for dipeptide synthesis except that [α - 32 P]UTP was used in addition to [3 H]Ala-tRNA^{Ala}. After incubation at 37°C for 60 min, a 5- μ l aliquot was removed for RNA analysis by electrophoresis with 10% polyacrylamide/urea gels (22). The RNA band, localized by autoradiography, was excised from the gels and assayed by Cerenkov radiation. The remaining 30 μ l of each incubation mixture was assayed for dipeptide synthesis by extraction of the labeled dipeptide into ethyl acetate (19). A unit of mRNA is defined as cpm per nucleotide in the mRNA synthesized *in vitro*.

the DNA-dependent dipeptide synthesis system. Under conditions where approximately equal amounts of transcript were produced, the DNA template that lacks the complementary region (pAD7-HindIII fragment) directed the synthesis of 10-fold more dipeptide than did the wild-type template (pAD6-HindIII fragment) (Table 2).

DISCUSSION

Expression of the *E. coli trp* operon is regulated by both repression and attenuation (for reviews, see refs. 27–29). Whereas repression regulates gene expression by controlling transcription initiation, attenuation acts by controlling transcription termination at a site within the leader region of the operon (2). *trp* leader RNA can form alternate secondary structures that are thought to regulate attenuation. According to the current model (30, 31), two mutually exclusive secondary structures, termed terminator and antiterminator, determine whether RNA polymerase terminates transcription at the attenuator or continues transcription into the structural genes of the operon. Formation of one or the other of the alternative regulatory secondary structures is believed to be determined by the position of the ribosome engaged in translation of the short coding region located early in *trp* leader RNA. In the amino acid biosynthetic operons of the bacterial species that have been studied to date, the peptides specified by the short coding regions are 14–30 residues in length and are rich in the amino acid(s) that controls the operon. One essential feature of the above model is a requirement for efficient synthesis of the short peptide encoded in the leader region of the respective operons. Although gene fusion analyses had demonstrated that the *trp* leader ribosome binding site was functional (11, 12), a leader peptide specified by an amino acid biosynthetic operon had never been detected previously.

Using an *in vitro* DNA-dependent protein synthesis system, we demonstrated the synthesis of the leader peptide specified by the *trp* operons of *E. coli*, *S. marcescens*, and *S. typhimurium*. Peptide synthesis was abolished by addition of purified *E. coli trp* repressor (Fig. 2), demonstrating that the peptide was translated from transcripts originating at the *trp* promoter. The authenticity of the peptides was established by demonstrating that both the *E. coli* and *S. marcescens trp* leader peptides were selectively labeled only by those amino acids predicted from the DNA sequence. The *E. coli trp* leader peptide

was shown to be labile *in vitro* and to have a half-life of 3–4 min. This lability undoubtedly contributed to our inability to detect the peptide in previous studies.

The *S. typhimurium trp* operon segment directed the synthesis of the predicted NH₂-terminal di- and tripeptides, fMet-Ala and fMet-Ala-Ala, respectively, in a defined DNA-dependent peptide-synthesizing system (Table 1). In this system, both plasmid DNA and linear *trp* operon fragments served equally well as template. However, in S-30 extracts, we failed to detect synthesis of the leader peptide when plasmid DNA was used. This may have been due to the lower molar concentration of plasmid template used relative to that of the ligated fragment that was used successfully.

To determine whether the RNA segment complementary to the wild-type leader ribosome binding site can block translation initiation, we assayed dipeptide synthesis in reaction mixtures primed with DNA templates containing or lacking the downstream, complementary segment. Under conditions where approximately equal amounts of RNA were synthesized from the templates, the template lacking the complementary sequence yielded 10-fold more dipeptide than did the wild-type template (Table 2). Therefore, it seems likely that pairing of the distal, complementary segment of the transcript with the ribosome binding site used for *trp* leader peptide synthesis can inhibit translation initiation. Because analogous base pairing can occur in *trp* leader transcripts of other bacterial species, it appears that the *trp* leader ribosome binding site can be blocked in these species as well. Because the distal, complementary segment of the wild-type transcript presumably is synthesized *in vivo* only after at least one round of translation of the peptide-encoding region, the role of this inhibition of translation initiation might be to prevent further leader peptide synthesis once the decision-making ribosome has transversed the peptide-encoding region. This possibility seems plausible in the context of our current view of attenuation in amino acid biosynthetic operons and is consistent with the lability of the leader peptide and the difficulty we experienced in detecting the peptide with wild-type templates.

We thank Richard Kelley for providing us with purified *E. coli trp* repressor, William Brusilow for scanning the fluorograms, and our colleagues for many helpful suggestions and critical reading of the manuscript. One of the authors (A.D.) is a Career Investigator fellow of the American Heart Association. A second (C.Y.) is a Career Investigator of the American Heart Association. These studies were supported by grants from the U.S. Public Health Service (GM 09738), the National

Science Foundation (PCM 8208866), and the American Heart Association (69C-15).

1. Kasai, T. (1974) *Nature (London)* **249**, 523–527.
2. Bertrand, K., Korn, L., Lee, F., Platt, T., Squires, C. L., Squires, C. & Yanofsky, C. (1975) *Science* **189**, 22–26.
3. Lee, F. & Yanofsky, C. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 4365–4369.
4. Zurawski, G., Brown, K., Killingly, D. & Yanofsky, C. (1978) *Proc. Natl. Acad. Sci. USA* **75**, 4271–4275.
5. DiNocera, P. P., Blasi, F., DiLauro, R., Frunzio, R. & Bruni, C. B. (1978) *Proc. Natl. Acad. Sci. USA* **75**, 4276–4280.
6. Barnes, W. M. (1978) *Proc. Natl. Acad. Sci. USA* **75**, 4281–4285.
7. Gardner, J. F. (1979) *Proc. Natl. Acad. Sci. USA* **76**, 1706–1710.
8. Gemmill, R. M., Wessler, S. R., Keller, E. B. & Calvo, J. M. (1979) *Proc. Natl. Acad. Sci. USA* **76**, 4941–4945.
9. Nargang, F. E., Subrahmanyam, C. S. & Umbarger, H. E. (1980) *Proc. Natl. Acad. Sci. USA* **77**, 1823–1827.
10. Lawther, R. P. & Hatfield, G. W. (1980) *Proc. Natl. Acad. Sci. USA* **77**, 1862–1866.
11. Miozzari, G. F. & Yanofsky, C. (1978) *J. Bacteriol.* **133**, 1457–1466.
12. Schmeissner, U., Ganem, D. & Miller, J. H. (1977) *J. Mol. Biol.* **109**, 303–326.
13. Zurawski, G., Elseviers, D., Stauffer, G. V. & Yanofsky, C. (1978) *Proc. Natl. Acad. Sci. USA* **75**, 5988–5992.
14. Miozzari, G. & Yanofsky, C. (1978) *Nature (London)* **276**, 684–689.
15. Blumenberg, M. & Yanofsky, C. (1982) *J. Bacteriol.* **152**, 49–56.
16. Blumenberg, M. & Yanofsky, C. (1982) *J. Bacteriol.* **152**, 57–62.
17. Robakis, N., Meza-Basso, L., Brot, N. & Weissbach, H. (1981) *Proc. Natl. Acad. Sci. USA* **78**, 4261–4264.
18. Cenatiempo, Y., Robakis, N., Reid, B. R., Weissbach, H. & Brot, N. (1982) *Arch. Biochem. Biophys.* **218**, 572–578.
19. Peacock, S., Cenatiempo, Y., Robakis, N., Brot, N. & Weissbach, H. (1982) *Proc. Natl. Acad. Sci. USA* **79**, 4609–4612.
20. Vieira, J. & Messing, J. (1982) *Gene* **19**, 259–268.
21. Bennett, G. N., Brown, K. D. & Yanofsky, C. (1978) *J. Mol. Biol.* **121**, 139–152.
22. Das, A., Crawford, I. P. & Yanofsky, C. (1982) *J. Biol. Chem.* **257**, 8795–8798.
23. Laemmli, U. K. (1970) *Nature (London)* **227**, 680–685.
24. Chamberlin, J. P. (1979) *Anal. Biochem.* **98**, 132–135.
25. Yanofsky, C., Platt, T., Crawford, I. P., Nichols, B. P., Christie, G. E., Horowitz, H., van Cleemput, M. & Wu, A. M. (1981) *Nucleic Acids Res.* **9**, 6647–6668.
26. Shine, J. & Dalgarno, L. (1974) *Proc. Natl. Acad. Sci. USA* **71**, 1342–1346.
27. Crawford, I. P. & Stauffer, G. V. (1980) *Annu. Rev. Biochem.* **49**, 163–195.
28. Yanofsky, C. (1981) *Nature (London)* **289**, 751–758.
29. Kolter, R. & Yanofsky, C. (1982) *Annu. Rev. Genet.* **16**, 113–134.
30. Oxender, D. L., Zurawski, G. & Yanofsky, C. (1979) *Proc. Natl. Acad. Sci. USA* **76**, 5524–5528.
31. Stroynowski, I. & Yanofsky, C. (1982) *Nature (London)* **298**, 34–38.