JOURNAL OF BACTERIOLOGY, Mar. 1978, p. 1457-1466
0021-9193/78/0133-1457$02.00/0

Vol. 133, No. 3
Printed in U.S.A.

Translation of the Leader Region of the *Escherichia coli* Tryptophan Operon

G. F. MIOZZARI AND C. YANOFSKY*
Department of Biological Sciences, Stanford University, Stanford, California 94305

Received for publication 17 October 1977

When the *trp* operon of *Escherichia coli* contains either of two deletions that fuse the initial portion of the leader region to the distal segments of the *trpE* gene, novel fusion polypeptides are produced. The new polypeptides are synthesized efficiently both in vivo and in vitro, and their synthesis is subject to repression by *trp* repressor. Fingerprint analyses of tryptic and chymotryptic digests of the new polypeptides show that both contain *trpE* polypeptide sequences and, despite their different sizes, share the same N-terminal sequence. Our results suggest that synthesis of the new polypeptides is initiated at the AUG-centered ribosome-binding site in the leader region and proceeds in phase to the region coding for the C-terminal end of the *trpE* polypeptide.

Polycistronic mRNA transcribed from the tryptophan (*trp*) operon of *Escherichia coli* contains a leader sequence of 162 nucleotides preceding the translation initiation codon for the *trpE* polypeptide, the first major polypeptide specified by the operon (3). A short transcript also is produced; it contains only the first 142 nucleotides of *trp* mRNA and results from transcription termination at the attenuator, a regulatory site in the leader region of the operon (1). In vitro studies revealed that RNA transcribed from the leader region contains a ribosome-binding site that could serve as a site for initiation of translation (14). The nucleotide sequence of the leader region of the *trp* operon has been determined (21; F. Lee, K. Bertrand, G. Bennett, and C. Yanofsky, J. Mol. Biol., in press); it indicates that there is an AUG codon at positions 27 to 29 in the center of the ribosome-binding site, followed in phase by a UGA translation termination codon at positions 69 to 71. This sequence could code for a leader peptide 14 amino acid residues in length (Fig. 1). Translation of this leader segment has been postulated to play an important role in regulating transcription termination at the attenuator (10).

To determine whether leader RNA is translated, we analyzed the polypeptide products synthesized both in vivo and in vitro from segments of the *trp* operon containing either (i) the entire initial portion of the operon including the leader region, *trpE* and *trpD*, or (ii) the same portion with deletions that fuse the operator-proximal part of the DNA sequence coding for the hypothetical leader peptide to the distal portion of *trpE*. We found that the operons carrying the deletions directed the synthesis of fusion polypeptides, each of which contains the N-terminal end of the hypothetical leader peptide and the C-terminal end of the *trpE* protein. Schmeissner et al. (18) have also obtained evidence indicating that the ribosome-binding site specified by the leader region of the *trp* operon may be used in vivo to initiate polypeptide synthesis.

MATERIALS AND METHODS

Bacterial strains, phage, and plasmids. All bacterial strains used in this study were derivatives of *E. coli* K-12 W3110. The *trp* operon internal deletions *ΔtrpLE1413* and *ΔtrpLE1417* were isolated by Bertrand et al. as described previously (1). The regions of the operon deleted in these strains are indicated in Fig. 1.

The plaque-forming *trp* transducing phage λcI857*trpED10ΔLE1413* and λcI857*trpED10ΔLE1417* were constructed by lysogenizing bacteria containing these *trp* operon deletions with λcI857*trpED10* (5). After heat induction, the lysates were screened for phage containing *trpD* but lacking *trpE*. Up to 50% of the progeny were found to have acquired the deletions.

Plasmid pVH153 was kindly provided by V. Hershfield and D. Helinski. It contains a 4.7-megadalton *Eco*RI fragment carrying *trpPOLED"C"* from the phage λ*trpED10* inserted into the *Eco*RI site in pVH51 (mini-ColE1; 7). Plasmids pGM1 and pGM3 were constructed in the course of this study. They are identical to pVH153 except that they carry the deletions *ΔtrpLE1413* (pGM1) and *ΔtrpLE1417* (pGM3). A diagram of plasmids pVH153, pGM1, and pGM3 is shown in Fig. 2. Plasmid DNA was prepared as described in Selker et al. (19). The preparation of phage and phage DNA has been described (25).

Buffers and solutions. The makeup of the buffers and solutions used was as follows. The polyacrylamide gel electrophoresis sample buffer contained 60 mM tris(hydroxymethyl)aminomethane-hydrochloride buffer (pH 6.8), 3% (wt/vol) sodium dodecyl sul-

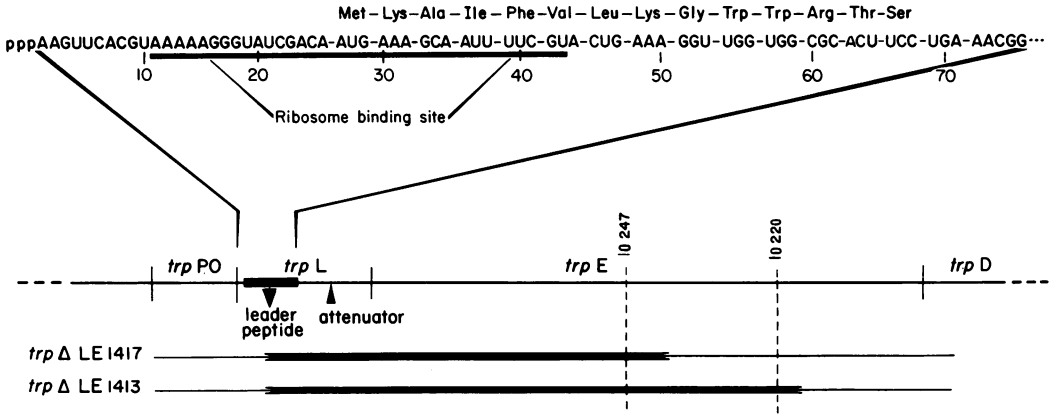

FIG. 1. *Initial portion of the trp operon of E. coli. The bottom part of the figure is a simplified map of the trpPOLE region showing the approximate positions of the attenuator and the sequences coding for the hypothetical leader peptide. trpL and trpE are not drawn to scale. The nucleotide sequence of the RNA transcribed from the initial portion of the leader region is shown at the top of the figure. It shows the location of the ribosome-binding site (14) and the amino acid sequence of the hypothetical leader peptide. The black bars at the bottom of the figure indicate the extent of the internal deletions ΔtrpLE1417 and ΔtrpLE1413. The left-hand end point of each deletion has been determined to fall between base pairs 49 and 52 of the leader region; the right-hand deletion end points were located by genetic recombination analyses with known trpE point mutants (1; Bertrand and Yanofsky, unpublished data).*

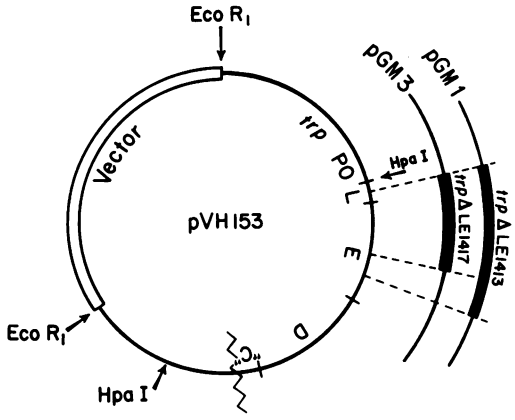

FIG. 2. *MAP of plasmid pVH153 showing the location of trpPOLED"C" and the sites of cleavage by restriction endonucleases EcoRI and Hpa I. The plasmids pGM1 and pGM3 are identical to pVH153, except that they carry the trp operon internal deletions ΔtrpLE1413 (pGM1) and ΔtrpLE1417 (pGM3). The zigzag line next to trp"C" marks the boundary between bacterial trp operon sequences and phage DNA.*

fate (SDS), 5% (vol/vol) 2-mercaptoethanol, 15% (vol/vol) glycerol, and 0.05% (wt/vol) xylene cyanol FF. The staining solution contained 50% (vol/vol) methanol, 10% (vol/vol) acetic acid, and 0.25% (wt/vol) Coomassie brilliant blue. The destaining solution contained 25% (vol/vol) ethanol and 10% (vol/vol) acetic acid. The elution buffer X contained 0.5 M NH_4 acetate, 0.01 M Mg acetate$_2$, 0.1 mM ethylenediaminaetetraacetate, and 0.1% (wt/vol) SDS.

In vitro protein synthesis. The preparation of S30 extracts and the composition of the reaction mixture for in vitro protein synthesis were essentially as described in Zalkin et al. (25), except that the reactions were conducted in the presence of polyethylene glycol (Carbowax-6000, Union Carbide) at a final concentration of 25 mg/ml. For labeling experiments with [^{35}S]methionine and [^{3}H]lysine, the corresponding unlabeled amino acids were either omitted from the reaction mixture (methionine) or reduced to 10% of the normal concentration (lysine). After 60 min of incubation at 34°C, incorporation was terminated by the addition of the corresponding unlabeled amino acid to a final concentration of 5 mM; ribonuclease (RNase) A (50 μg/ml) and deoxyribonuclease (DNase; 50 μg/ml) were then added. Incubations were continued for 5 min at 34°C before the reactions were chilled on ice. Protein was precipitated by the addition of 9 volumes of cold acetone and pelleted by centrifugation. The pellet was dissolved in polyacrylamide gel electrophoresis sample buffer by heating at 95°C for 2 to 3 min.

Phage infection of UV-irradiated cells. UV-irradiated cells (13,000 ergs/mm^2) of the UV-sensitive *E. coli* K-12 strain 159 (12), carrying either the *trpR*$^+$ or the *trpR* allele, were lysogenized with λ*c*I587(*ind*) and infected with purified λ*trp* transducing phages in the presence of tryptophan (50 μg/ml), as described by Ray and Murialdo (16). Proteins synthesized between 15 and 30 min after infection were labeled with [^{35}S]methionine (10 μCi/ml; 315 Ci/mmol). Incorporation of the labeled amino acid was stopped by the addition of unlabeled methionine to a final concentration of 5 mM, followed by sodium azide (50 mM). The cultures were chilled on ice and immediately pelleted

by centrifugation. The pellet was lysed directly in polyacrylamide gel electrophoresis sample buffer by heating for 3 min at 95°C.

Polyacrylamide gel electrophoresis. Proteins labeled in vivo or in vitro were generally separated on 12.5% polyacrylamide gels containing 0.1% SDS by the discontinuous buffer method described by Laemmli (9). In some instances (see figure legends), 10 to 20% linear polyacrylamide gradients containing 0.1% SDS and 8 M urea in both the separating gel and the stacking gel were used. After electrophoresis, the radioactive bands were located directly by autoradiography for preparative runs (see fingerprint analysis below). In analytical runs, the protein bands were fixed by immersion in staining solution for 2 h and the destained for at least 4 h with several changes of destaining solution. The destained gel was briefly rinsed in 25% ethanol to remove excess acetic acid and dried on filter paper. The radioactive bands were located by autoradiography ([^{35}S]methionine) or fluorography ([^{3}H]lysine) (2).

The following polypeptides were used as markers for the determination of the molecular weights of labeled proteins (molecular weights in parentheses): horse heart cytochrome *c* (12,400), lysozyme (14,300), sperm whale myoglobin (17,800), immunoglobulin L-chain (23,500), tryptophan synthetase α (28,700), ovalbumin (45,000), immunoglobulin H-chain (50,000), and bovine serum albumin (67,000). With the exception of lysozyme (Worthington Biochemicals Corp.) and tryptophan synthetase α (prepared in this laboratory), all markers were from a set purchased from Mann Research Laboratories. The molecular weights of immunoglobulin H- and L-chains are taken from Weber and Osborn (23).

Fingerprint analysis of tryptic and chymotryptic digests. Radioactive bands were located on gels by autoradiography. Gel slices containing the labeled proteins were fragmented into small pieces with a glass rod. Portions (1 to 2 ml) of elution buffer X and bovine serum albumin (50 μg) were added, and the mixture was incubated at 37°C overnight. The buffer solution containing most of the labeled protein was separated from the gel pieces by filtration through a pipette tip plugged with glass wool. The proteins were precipitated four times by adding trichloroacetic acid to 20%, suspending in 0.1% NH_4OH, and washing three times with acetone. The dried residue was taken up in 50 μl of performic acid and oxidized for 60 min at 0°C. The acid was removed by several cycles of water addition and drying in a desiccator. Trypsin and chymotrypsin digestions were performed in 0.05 M NH_4HCO_3 (pH 7.8) for 4 to 10 h at 37°C. The volatile solvent was removed by drying in a desiccator, and the dry residue was taken up in a small volume of water. The amount of radioactivity in the final sample was determined by counting a small sample of the solution in a liquid scintillation counter.

Material corresponding to approximately 200,000 cpm was applied to Whatman 3MM paper and fingerprinted as described by Helinski and Yanofsky (6), except that electrophoresis was used in the first dimension. The labeled peptides were located on the fingerprints by autoradiography.

Enzymes and chemicals. Alpha chymotrypsin and trypsin-TPCK were obtained from Worthington Biochemicals Corp. (Freehold, N.J.). Restriction endonuclease *Eco*RI was purchased from Miles Research Laboratories (Elkhart, Ind.). Restriction endonuclease *Hpa* I and T4 DNA ligase were supplied by New England Biolabs (Beverly, Mass.).

L-[^{35}S]methionine (315 Ci/mmol) and L-[4,5-^{3}H-(*N*)]lysine (39 Ci/mmol) were obtained from New England Nuclear Corp. (Boston, Mass.). SDS was purchased from Bio-Rad Laboratories (Richmond, Calif.) and recrystallized twice from ethanol before use in gel electrophoresis.

RESULTS

Construction of phage and plasmids carrying *trpLE* deletions. Bertrand et al. (1) have reported the isolation of mutants with *trp* operon internal deletions that have one end point in the leader region of the operon and the other in one of the five structural genes. Among the mutants isolated were two strains, W3110 Δ*trpLE1417* (1) and W3110 Δ*trpLE1413* (K. Bertrand and C. Yanofsky, unpublished data), that had their left-hand deletion end points (Fig. 1) within the region coding for the hypothetical leader peptide and their right-hand deletion end points within the first major structural gene, *trpE*. The left-hand end points of both deletions were determined to be between base pairs 49 and 52 of the leader region by oligonucleotide analysis of ^{32}P-labeled operator-proximal *trp* mRNA (1; Bertrand and Yanofsky, unpublished data). The deletion strains retain the AUG-centered ribosome-binding sequence early in the leader region as well as the initial seven or eight codons for the hyopthetical leader peptide (Fig. 1). In the absence of functional *trp* repressor, i.e., when a *trpR* allele is introduced, both strains show an 8- to 10-fold increase in the level of *trpB* protein compared with the level in an otherwise isogenic *trpR trp*$^{+}$ strain (1; Bertrand and Yanofsky, unpublished data). This increase in distal gene expression results from the removal of the attenuator, a site of transcription termination in the leader region of the operon (1). The levels of *trp* operon expression observed in W3110 *trpR* strains with Δ*trpLE1413* or Δ*trpLE1417* are among the highest observed in attenuator deletion strains, indicating that the two deletions do not have polar effects on distal gene expression. Because the two deletions terminate in regions of *trpE* within which nonsense mutations cause severe polarity (24), it seems likely that they allow in-phase fusion of translated segments of leader *trp* mRNA and *trpE* mRNA. Alternatively, the deletions could generate efficient translation restart sites within *trpE*. In either case, the carboxy-terminal segment of the *trpE* polypeptide should be synthesized.

To analyze the polypeptide products synthesized from operon segments containing Δ*trpLE1413* and Δ*trpLE1417*, the deletions were transferred onto phage λcI857*trpED10* by genetic recombination. Subsequently, a restriction endonuclease *Eco*RI-generated fragment of this phage DNA containing the operator-proximal part of the *trp* operon was cloned, with pVH51 (mini-ColE1; 7) as a vector. Both resulting plasmids carry *trpD*$^+$ and, therefore, can be selected for on the basis of their ability to allow strain W3110 Δ*trpLD102* to grow on anthranilate (8). Figure 2 shows a diagram of pVH153, a plasmid that carries the initial segment of the wild-type *trp* operon up to the beginning of *trpC* (5), and a diagram of the isogenic plasmids, pGM1 and pGM3, that carry the deletions Δ*trpLE1413* (pGM1) and Δ*trpLE1417* (pGM3). All of the plasmids have two recognition sites for restriction endonuclease *Hpa* I; one is in the *trp* promoter/operator region, and the other is just beyond the *trp* genes in phage DNA (unpublished data from this laboratory). The *Hpa* I digestion pattern of the three plasmids is shown in Fig. 3. Although the larger fragment (molecular weight, 4.3×10^6) is the same in all cases, the smaller fragment, which has a molecular weight of 2.5×10^6 in pVH153, is reduced in size in plasmids pGM1 (molecular weight, 1.75×10^6) and pGM3 (molecular weight, 2.05×10^6), reflecting the sizes of the respective deletions. The differences in molecular weight indicate that the deletions remove about 1,200 base pairs (Δ*trpLE1413*) and 720 base pairs (Δ*trpLE1417*).

Expression of plasmid DNA in vitro. Faithful expression of the *trp* operon DNA of λ or φ80 *trp* transducing phages in a coupled in vitro transcription/translation system (S30) has been reported previously (25). Depending on the nature of the template used, it was found that transcription of the *trp* genes could originate at either a phage promoter or at the *trp* promoter. However, only *trp*-promoted transcription was sensitive to the presence of *trp* repressor. Plasmids pVH153, pGM1, and pGM3 were used as templates in the in vitro transcription/translation system. The proteins made in the presence or in the absence of *trp* aporepressor were labeled with [^{35}S]methionine and separated on a SDS-polyacrylamide gel. The results (Fig. 4) suggest that transcription of the *trp* genes in plasmid pVH153 initiates exclusively at the *trp* promoter, as synthesis of the *trpE* and *trpD* proteins is essentially eliminated if the S30 extract is derived from a strain containing functional *trp* aporepressor (*trpR*$^+$). Figure 5 shows an autoradiogram of a dried gel. Whereas both the *trpE* (molecular weight, 60,000) and *trpD* (molecular weight, 65,000) polypeptides are

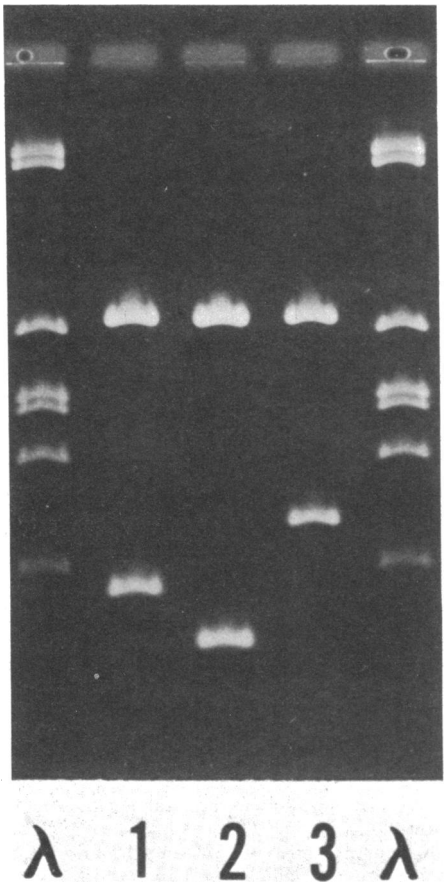

FIG. 3. *Agarose gel electrophoresis of plasmid DNA restricted with restriction endonuclease Hpa I. Restriction digests of plasmids pGM3 (1), pGM1 (2), and pVH153 (3) were run on 0.8% agarose gels in TEA-NaCl buffer as described previously (19). The outside slots (λ) contain λ DNA restricted with restriction endonuclease EcoRI; the fragments serve as molecular weight markers (22).*

present and strongly repressed when pVH153 DNA is provided, there is no band in the position of the *trpE* polypeptide with the deletion plasmids as templates. In each case a new, very intense band appears that is strongly repressed by the presence of *trp* repressor. The new polypeptide bands migrate with molecular weights of 19,000 (pGM1) and 36,000 (pGM3). The difference in molecular weight between the *trpE* protein and the new polypeptides generated by the Δ*trpLE1413* and Δ*trpLE1417* deletions correlates well with the amount of DNA deleted in these strains (see Discussion). The new protein bands could thus represent in-phase fusions between the N-terminal end of the putative leader peptide and the C-terminal end of the *trpE* polypeptide. We will refer to these putative fu-

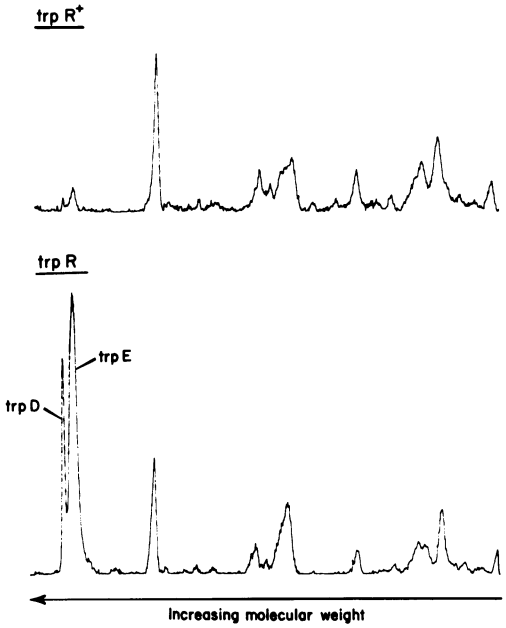

FIG. 4. *Repression of the synthesis of trp enzymes in vitro. Proteins synthesized with pVH153 as template were labeled with [35S]methionine (150 μCi/ml, 315 Ci/mmol) in vitro with S30 extracts derived from strains carrying either the trpR+ or trpR allele. The labeled proteins were separated on a 10 to 20% gradient polyacrylamide gel containing SDS and 8 M urea and were located by autoradiography of the dried gel. The figure is a densitometer scan (Joyce-Loebl) of the autoradiogram; it shows the labeled proteins synthesized in the presence (trpR+) or absence (trpR) of functional trp repressor.*

sion proteins as *LE1413* and *LE1417* polypeptides.

In vivo expression of *trp* operon segments containing *trpLE* deletions. UV-irradiated *E. coli* cells were infected with λcI857 or λ*trp* transducing phages carrying deletions Δ*trpLE1413* or Δ*trpLE1417*. Proteins synthesized after phage infection were labeled with [35S]methionine and analyzed by electrophoresis on SDS-polyacrylamide gels. UV irradiation essentially eliminates host protein synthesis (slot 1) without serious damage to the RNA or protein synthesis machinery (slot 3) (Fig. 6). If the infected host is a λcI857(*ind*) lysogen, the expression of the λ genes of the infecting phage is prevented (slot 4). In addition, transcription from the *trp* promoter can be repressed by infecting a host carrying a functional *trp* repressor in the presence of tryptophan (slots 5 and 7).

Infection of *trpR*+ and *trpR* λ lysogens with λ*trpED10ΔLE1413* and λ*trpED10ΔLE1417* identifies two protein bands that are present in the *trpR* but not in a *trpR*+ host: the *trpD* polypeptide band (molecular weight, 65,000) and one additional very intense polypeptide band. The additional bands have mobilities corresponding to molecular weights of 19,000 for the protein specified by λ*trpED10ΔLE1413* and 36,000 for the protein of λ*trpED10ΔLE1417*; they are identical in size to the presumed *trpLE* fusion proteins observed in vitro (Fig. 5). Although only *trp*-specific products are synthesized after infection of a λ lysogen with λ*trpEDΔLE1417*, λ-specific protein synthesis is only partially repressed when the phage employed is λ*trpED10ΔLE1413*. This incomplete regression may be related to the fact that λ*trpED10ΔLE1413* makes clear plaques on sensitive bacteria at 32°C and, thus, is not normally repressible.

Analysis of trypsin and chymotrypsin digests of *trpLE* fusion proteins. Polypeptides synthesized in vitro with plasmids pVH153,

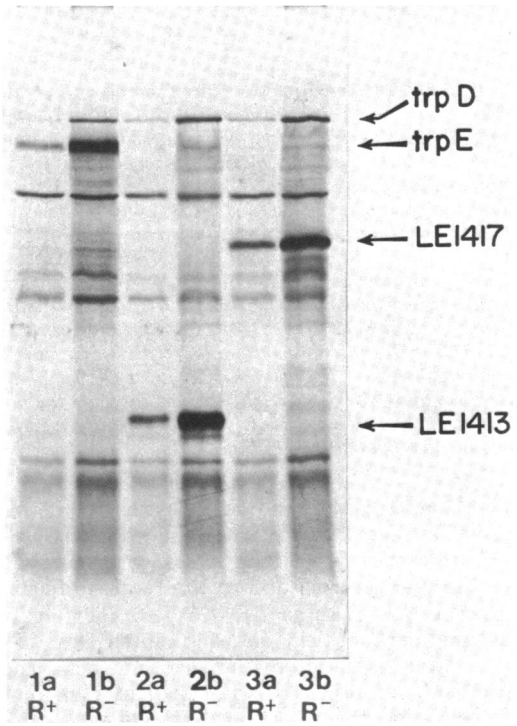

FIG. 5. *Polyacrylamide gel electrophoresis of proteins labeled in vitro. Proteins synthesized in vitro with pVH153 (1a, 1b), pGM1 (2a, 2b), and pGM3 (3a, 3b) as templates were labeled with [35S]methionine (250 μCi/ml, 315 Ci/mmol). The S30 extracts were derived from either trpR+ (1a, 2a, 3a) or trpR strains (1b, 2b, 3b). The labeled proteins were separated on a 12.5% polyacrylamide gel and located by autoradiography. The arrows indicate the positions of the trp-specific polypeptides synthesized by the three plasmids.*

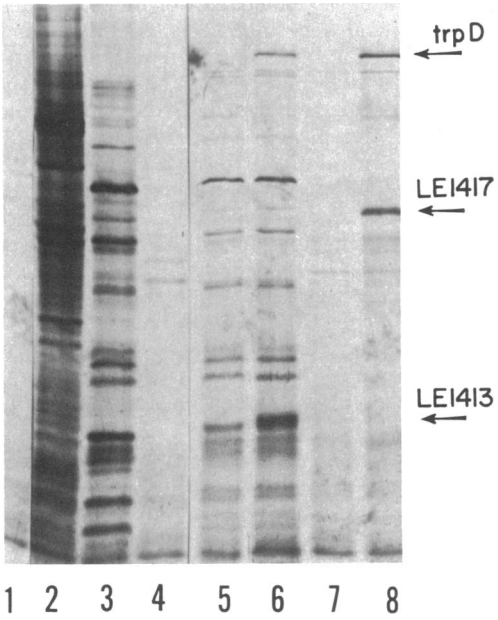

FIG. 6. *Polyacrylamide gel electrophoresis (12.5%) of the proteins labeled 15 to 30 min after infection of UV-irradiated cells with λtrp transducing phage. All samples with the exception of no. 2 (unirradiated control) were irradiated with UV at 13,000 ergs/mm². The slots show: (1) uninfected control, (2) unirradiated control, (3) strain 159 infected with λcI587, (4) strain 159 (λcI857) infected with λcI857, (5) strain 159(λcI857) trpR⁺ infected with λcI857trpED10ΔLE1413, (6) strain 159 (λcI857) trpR infected with λcI857trpED10ΔLE1413, (7) strain 159 (λcI857) trpR⁺ infected with λcI857trpED10ΔLE1417, and (8) strain 159 (λcI857) trpR infected with λcI857trpED10ΔLE1417. The arrows indicate the positions of the trp-specific polypeptides synthesized by the two trp transducing phages.*

pGM1, and pGM3 were labeled with [^{35}S]methionine and separated by electrophoresis on SDS-polyacrylamide gels. The bands corresponding to the *trpE* protein and to the putative fusion proteins *LE1413* and *LE1417* were located by autoradiography, cut out, and eluted from the gel. After oxidation with performic acid, the proteins were digested with trypsin or chymotrypsin, and the digestion products were analyzed by fingerprinting with high-voltage electrophoresis in the first dimension and descending chromatography in the second dimension. The results of this analysis are shown in Fig. 7. It is evident that the majority of the [^{35}S]methionine-labeled peptides generated by trypsin or chymotrypsin digestion of the fusion proteins are also present in the respective digests of the *trpE* protein. Therefore, the *LE1413* and *LE1417* polypeptides must contain part of the *trpE* protein. In addition, digestion with either enzyme yields a prominent spot that is identical for both the *LE1413* and *LE1417* polypeptides but that is not present in the digest of the *trpE* protein (indicated by an arrow in Fig. 7). Because the C-terminal end of both fusion proteins must be identical to the C-terminal end of the *trpE* polypeptide, the additional common spot must arise from the N-terminal portion of the putative fusion proteins. This suggests that the *LE1413* and *LE1417* polypeptides share the same N-terminal sequence.

In a similar experiment in which [^{3}H]lysine was used to label proteins, the same N-terminal tryptic and chymotryptic peptides observed with [^{35}S]methionine as label appeared on fingerprints of the *LE1413* and *LE1417* polypeptides but not on fingerprints of the *trpE* polypeptide (results not shown), indicating that the N-terminal tryptic and chymotryptic peptides of the *LE1413* and *LE1417* fusion proteins contain both methionine and lysine.

DISCUSSION

Two internal deletions in the *trp* operon of *E. coli*, Δ*trpLE1413* and Δ*trpLE1417*, have one terminus in the leader region of the operon and the other within *trpE*. Strains carrying either deletion retain the DNA segment corresponding to the AUG-centered ribosome-binding site at the beginning of the leader mRNA (14), as well as the first seven to eight amino acid codons for the putative leader peptide. The *trp* operon DNA containing the two deletions was shown both in vitro and in vivo to promote the synthesis of new polypeptides, each of which is subject to repression by *trp* repressor and is not coded for by the wild-type *trp* operon. Both polypeptides contain *trpE* sequences and could arise as a consequence of either an in-phase fusion of the residual part of *trpE* to an external translation initiation site or the activation of efficient translation restart sites within *trpE*.

Several observations indicate that the new polypeptides represent the fusion of the N-terminal end of the leader peptide and the C-terminal end of the *trpE* protein.

(i) The molecular weights of the postulated *LE* fusion proteins are consistent with the amount of DNA removed by each of the deletions. A restriction analysis of isogenic plasmids carrying either wild-type *trpE* or the *trpLE* deletions showed that Δ*trpLE1413* and Δ*trpLE1417* remove about 1,200 and 720 base pairs, respectively. Thus, these deletions reduce the coding capacity by 400 and 240 amino acids, and one would expect the molecular weight of the *trpE* polypeptide to be reduced from 60,000 to about 18,000 for the Δ*trpLE1413* deletion and to about

35,000 for the *ΔtrpLE1417* deletion, if the remaining *trpE* sequences are translated. The calculated values of 18,000 and 35,000 are virtually identical to the observed molecular weights (19,000 and 36,000) of the new polypeptides specified by the two deletions, both in vivo and in vitro. This suggests that the residual segment of *trpE* in each of the deletions must be translated starting at the site immediately preceding or following the fusion point between *trpE* and the leader region.

(ii) The presumed *LE* fusion polypeptides are synthesized at a very high rate. Although both *ΔtrpLE1413* and *ΔtrpLE1417* have their right-hand deletion end point in regions of *trpE* within which nonsense mutations have been shown to cause severe polarity (24), neither of the deletions has a polar effect on distal gene expression in vivo. The high levels of *trpB* protein observed in *trpR* strains carrying the two deletions suggest that distal *trpE* mRNA must be translated efficiently in both deletion operons. This conclusion is confirmed by the observation that, under conditions of DNA excess in vitro, between 6 and 10 times more [^{35}S]methionine is incorporated into the fusion polypeptides than into the *trpE* protein (data not shown). This increase closely reflects the observed 8- to 10-fold increase in distal mRNA production associated with removal of the attenuator site (1) and indicates that initiation of translation is just as efficient for the new polypeptides as it is for the wild-type *trpE* protein. Reinitiation of translation at internal sites within structural genes after chain termination has been described for the *rIIB* cistron of phage T4 (17) as well as for the *Z* gene (11, 13) and the *I* gene (4, 15) of the *lac* operon of *E. coli*. However, the efficiency of reinitiation is generally low. The restart sites at the beginning of the *lacI* gene lead to the accumulation of C-terminal repressor fragments at 10% of the wild-type repressor level (4), and the polar effects of chain termination mutants in *lacZ* on transacetylase expression (11, 13, 26) suggest similar or only slightly higher efficiencies for the restart sites within *lacZ*. If the high levels of *trpLE* polypeptides observed with both *LE* deletions were due to reinitiation of translation within *trpE*, internal reinitiations would have to occur essentially at the frequency observed for wild-type *trpE* mRNA. The existence of two such sites within *trpE* is very unlikely, particularly in view of the fact that there appear to be no major discontinuities in the *trpE* polarity gradient (24).

(iii) Additional evidence against internal restarts within *trpE* is provided by the fact that both new polypeptides share an identical sequence at their N-terminal end. Fingerprint analysis of tryptic and chymotryptic digests of both *LE* fusion proteins yields a prominent spot that is identical for both fusions, but is not present in digests of the *trpE* protein. As argued earlier, these spots must arise from the N-terminal portion of the fusion polypeptides. The most likely way both polypeptides can share the same N-terminal tryptic and chymotryptic peptides is for them to be initiated at the same site outside *trpE* and to have the same amino acid sequence up to the first site of trypsin and chymotrypsin cleavage. Among the approximately 50 transcribed nucleotides that precede the fusion point between *trpL* and *trpE* (Fig. 1), the only obvious candidate for this initiation site is the AUG-centered ribosome-binding site early in the leader region. In in vitro experiments, this site has been shown to bind ribosomes with the same efficiency as the *trpE* ribosome-binding site (14). As pointed out previously (14), it resembles a number of other ribosome-binding sites in *E. coli* in that it contains sequences that can potentially base pair with sequences at the 3′ terminus of 16S rRNA (20). Translation initiation at the leader ribosome-binding site would fuse the first seven or eight amino acids of the putative leader peptide in phase to the C-terminal portion of the *trpE* polypeptide. In fact, both the tryptic and chymotryptic N-terminal peptides of the fusion proteins contain both methionine and lysine. Because the N-terminal sequence of the putative leader peptide is Met-Lys-Ala-Ile-Phe-Leu, one would expect both the tryptic peptide (Met-Lys) and the chymotryptic peptide (Met-Lys-Ala-Ile-Phe) to be labeled by either [^{35}S]methionine or [^{3}H]lysine if synthesis of the *LE1413* and *LE1417* polypeptides is initiated at the leader ribosome-binding site.

The high rate of synthesis of both new polypeptides, the fact that they share a common N-terminal tryptic and chymotryptic peptide (which is not present in digests of the intact *trpE* protein), and the existence of an efficient ribosome-binding site immediately preceding the fusion point between *trpL* and *trpE* lead us to conclude that the *LE1413* and *LE1417* polypeptides represent in-phase fusions of the N-terminal end of the leader peptide and the C-terminal end of the *trpE* protein. Schmeissner et al. (18) have reported a similar observation in a fusion of the *lacI* gene to the early part of the *trp* operon. They sequenced the N-terminal segment of the resulting fusion protein and found the first six amino acids to be identical in sequence to the putative leader peptide of the *trp* operon. Our own observations confirm their results and indicate that the ribosome-binding site at the beginning of the *trp* leader sequence is

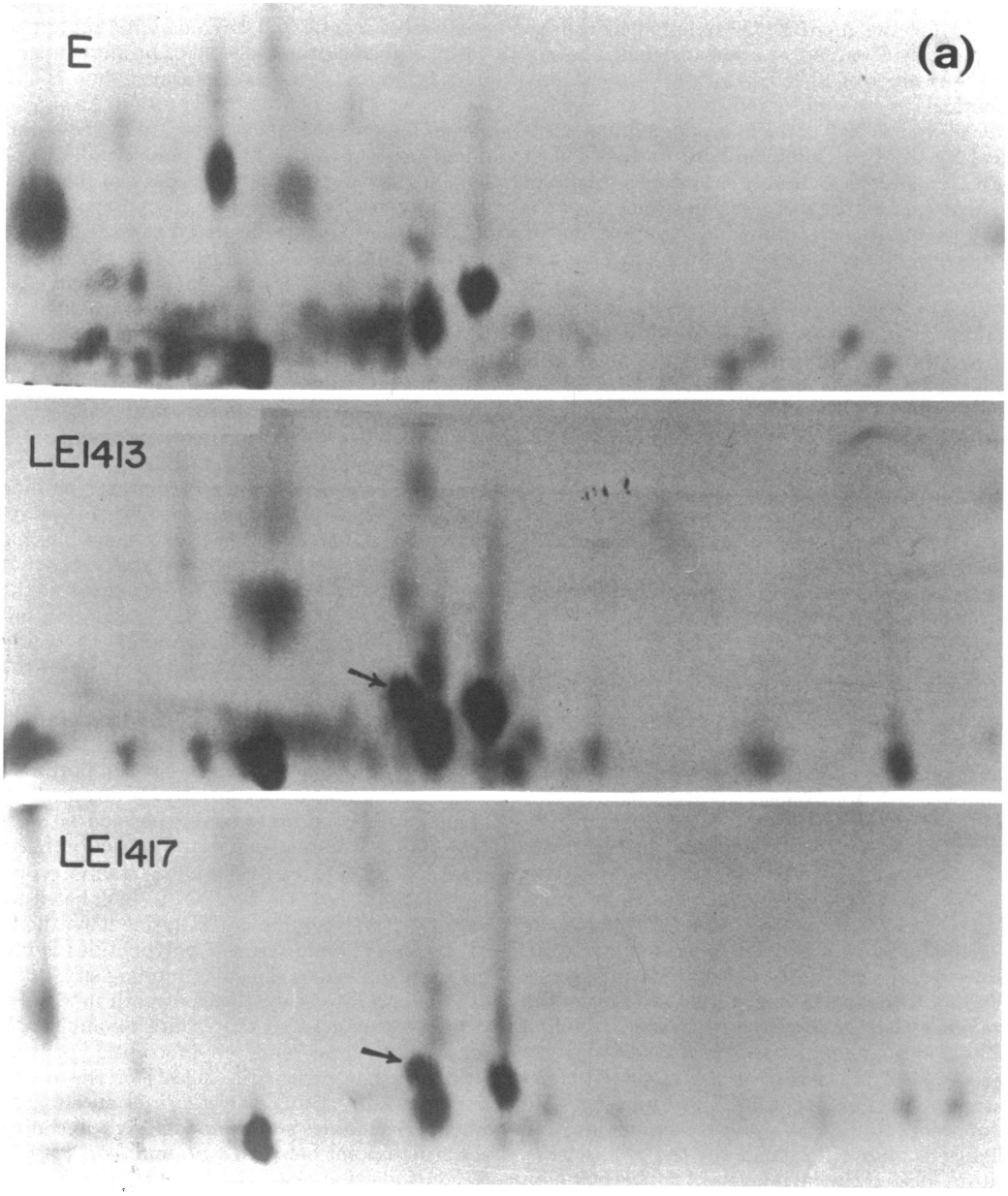

FIG. 7. *Two-dimensional pattern of tryptic (a) and chymotryptic (b) digests of the trpE polypeptide and the LE1413 and LE1417 fusion proteins. Electrophoresis (left to right) was used in the first dimension; descending chromatography (bottom to top) was used in the second dimension. The arrows indicate the positions of the spots that are present in the digests of both fusion proteins but not in the digests of the trpE protein.*

used for the efficient initiation of translation under conditions where the leader peptide is fused to the C-terminal portion of another polypeptide.

All our efforts to demonstrate the synthesis of the intact leader polypeptide itself have been unsuccessful to date (G. Miozzari, unpublished data). In particular, samples labeled both in vivo and in vitro for various amounts of time (2 to 60 min) were analyzed on polyacrylamide gels (10 to 25% with increasing cross-linker concentration, containing 8 M urea to reduce the pore size), on two-dimensional fingerprints (as in Fig. 7), and on peptide columns (Dowex 50). In other

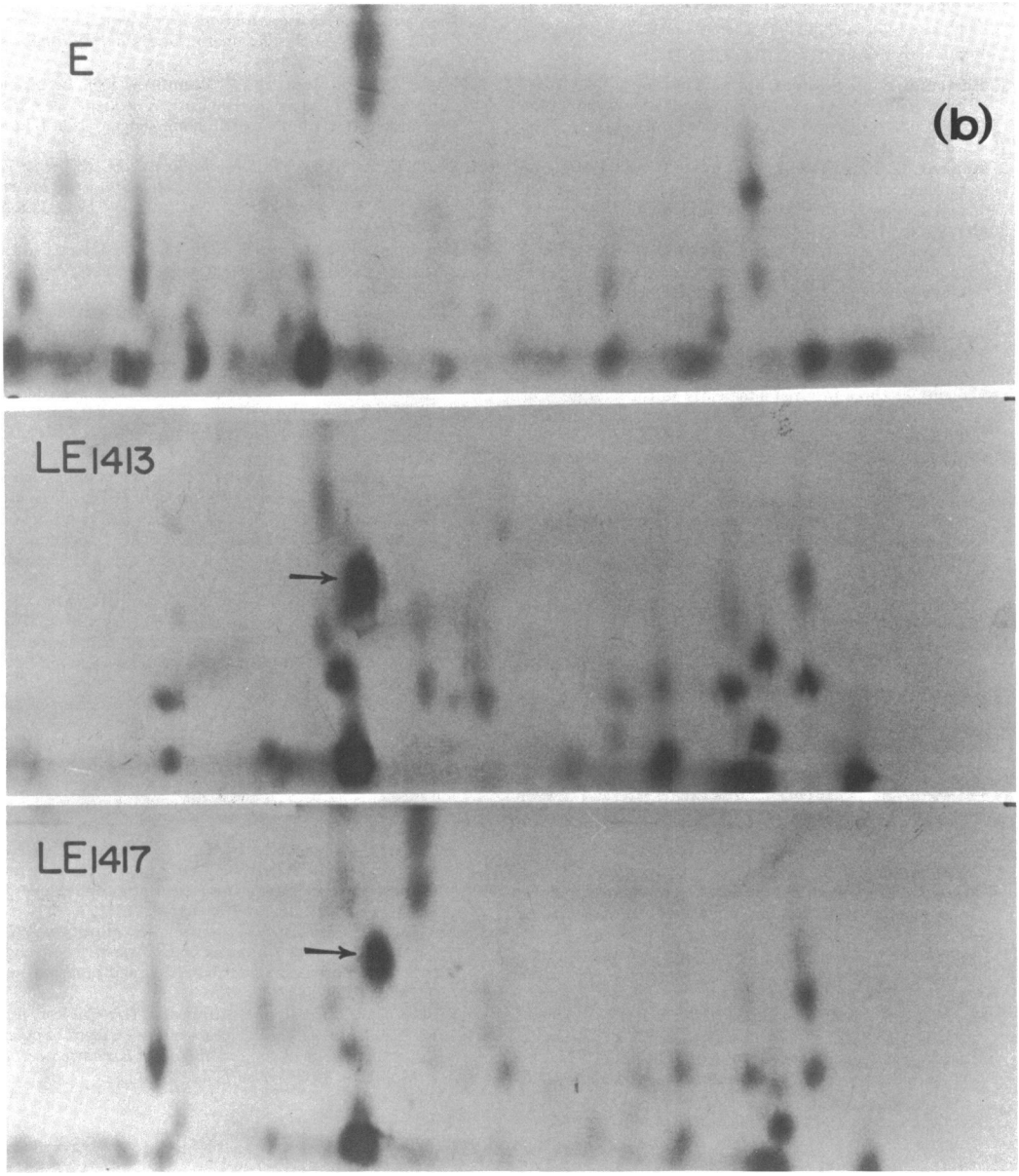

FIG. 7b.

experiments, we tried unsuccessfully to demonstrate the synthesis of the N-terminal portion of the leader peptide by analyzing tryptic and chymotryptic digests of in vitro reactions on fingerprints (as in Fig. 7). We have demonstrated that translation of the leader region can be initiated with high efficiency, and, therefore, one would expect the leader peptide to accumulate in significant amounts. Our inability to detect its synthesis may reflect very rapid degradation of the peptide both in vivo and in vitro. Alternatively, formation of the leader peptide may be prevented in some manner by the secondary structure of the leader RNA (10).

ACKNOWLEDGMENTS

We thank Daniel Oppenheim for help in constructing the *trp* transducing phage.

This work was supported by U.S. Public Health Service grant GM 09738 from the National Institute of General Medical Sciences, grant PCM 73-06774-A03 from the National Science Foundation, and the American Heart Association. G.F.M. is a postdoctoral fellow supported by the Swiss National Foundation. C.Y. is a career investigator of the American Heart Association.

These studies were performed in accordance with the Na-

tional Institutes of Health guidelines on recombinant DNA research.

LITERATURE CITED

1. **Bertrand, K., C. Squires, and C. Yanofsky.** 1976. Transcription termination *in vivo* in the leader region of the tryptophan operon of *Escherichia coli*. J. Mol. Biol. **103**:319–337.
2. **Bonner, W. M., and R. A. Laskey.** 1974. A film detection method for tritium-labeled proteins and nucleic acids in polyacrylamide gels. Eur. J. Biochem. **46**:83–88.
3. **Bronson, M. J., C. Squires, and C. Yanofsky.** 1973. Nucleotide sequences from tryptophan messenger RNA of *Escherichia coli*: the sequence corresponding to the amino-terminal region of the first polypeptide specified by the operon. Proc. Natl. Acad. Sci. U.S.A. **70**:2335–2339.
4. **Files, J. G., K. Weber, and J. H. Miller.** 1974. Translational reinitiation: reinitiation of *lac* repressor fragments at three internal sites early in the *lacI* gene of *Escherichia coli*. Proc. Natl. Acad. Sci. U.S.A. **71**:667–670.
5. **Helinski, D. R., V. Hershfield, D. Figurski, and R. J. Meyer.** 1977. Construction and properties of plasmid cloning vehicles, p. 151–165. *In* R. F. Beers and E. G. Basset (ed.), Recombinant molecules: impact on science and society. Raven Press, New York.
6. **Helinski, D. R., and C. Yanofsky.** 1962. Peptide pattern studies on the wild type A protein of the tryptophan synthetase of *Escherichia coli*. Biochim. Biophys. Acta **63**:10–19.
7. **Hershfield, V., H. W. Boyer, L. Chow, and D. R. Helinski.** 1976. Characterization of the mini-ColE1 plasmid. J. Bacteriol. **126**:447–453.
8. **Jackson, E. N., and C. Yanofsky.** 1973. The region between the operator and first structural gene of the tryptophan operon of *Escherichia coli* may have a regulatory function. J. Mol. Biol. **76**:89–101.
9. **Laemmli, U. K.** 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature (London) **227**:680–685.
10. **Lee, F., and C. Yanofsky.** 1977. Transcription termination at the *trp*-attenuators of *Escherichia coli* and *Salmonella typhimurium*: RNA secondary structure and regulation of termination. Proc. Natl. Acad. Sci. U.S.A **74**:4365–4369.
11. **Michels, C., and D. Zipser.** 1969. Mapping of polypeptide reinitiation sites within the β-galactosidase structural gene. J. Mol. Biol. **41**:341–347.
12. **Murialdo, H., and L. Siminovitch.** 1971. The morphogenesis of bacteriophage lambda. III. Identification of genes specifying morphogenetic proteins, p. 711–723. *In* A. D. Hershey (ed.), The bacteriophage lambda. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
13. **Newton, A.** 1969. Re-initiation of polypeptide synthesis and polarity in the *lac* operon of *Escherichia coli*. J. Mol. Biol. **41**:329–339.
14. **Platt, T., C. Squires, and C. Yanofsky.** 1976. Ribosome protected regions in the leader-*trpE* sequence of *Escherichia coli* tryptophan operon messenger RNA. J. Mol. Biol. **103**:411–420.
15. **Platt, T., K. Weber, D. Ganem, and J. H. Miller.** 1972. Translational restarts: AUG reinitiation of the *lac* repressor fragment. Proc. Natl. Acad. Sci. U.S.A. **69**:897–901.
16. **Ray, P., and H. Murialdo.** 1975. The role of gene *Nu3* in bacteriophage lambda head morphogenesis. Virology **64**:247–263.
17. **Sarabhai, A., and S. Brenner.** 1967. A mutant which reinitiates the polypeptide chain after chain termination. J. Mol. Biol. **27**:145–162.
18. **Schmeissner, U., D. Ganem, and J. H. Miller.** 1977. Genetic studies of the *lac* repressor. II. Fine structure deletion map of the *lacI* gene, and its correlation with the physical map. J. Mol. Biol. **109**:303–326.
19. **Selker, E., K. Brown, and C. Yanofsky.** 1977. Mitomycin C-induced expression of *trpA* of *Salmonella typhimurium* inserted into the plasmid ColE1. J. Bacteriol. **129**:388–394.
20. **Shine, J., and L. Dalgarno.** 1974. The 3′-terminal sequence of *Escherichia coli* 16S ribosomal RNA: complementarity to nonsense triplets and ribosome binding sites. Proc. Natl. Acad. Sci. U.S.A. **71**:1342–1346.
21. **Squires, C., F. Lee, K. Bertrand, C. L. Squires, M. J. Bronson, and C. Yanofsky.** 1976. Nucleotide sequence of the 5′ end of tryptophan messenger RNA of *Escherichia coli*. J. Mol. Biol. **103**:351–381.
22. **Thomas, M., and R. W. Davis.** 1975. Studies on the cleavage of bacteriophage lambda DNA with *EcoRI* restriction endonuclease. J. Mol. Biol. **91**:315–328.
23. **Weber, K., and M. Osborn.** 1969. The reliability of molecular weight determinations by dodecyl sulfate-polyacrylamide gel electrophoresis. J. Biol. Chem. **244**:4406–4412.
24. **Yanofsky, C., V. Horn, M. Bonner, and S. Stasiewsky.** 1971. Polarity and enzyme functions in mutants of the first three genes of the tryptophan operon of *Escherichia coli*. Genetics **69**:409–433.
25. **Zalkin, H., C. Yanofsky, and C. L. Squires.** 1974. Regulated *in vitro* synthesis of *Escherichia coli* tryptophan operon messenger ribonucleic acid and enzymes. J. Biol. Chem. **249**:465–475.
26. **Zipser, D., S. Zabell, J. Rothman, T. Grodzicker, and M. Wenk.** 1970. Fine structure of the gradient of polarity in the *Z* gene of the *lac* operon of *Escherichia coli*. J. Mol. Biol. **49**:251–254.