Proc. Natl. Acad. Sci. USA
Vol. 75, No. 10, pp. 4833–4837, October 1978
Biochemistry

Single base-pair alterations in the *Escherichia coli trp* operon leader region that relieve transcription termination at the *trp* attenuator

(termination signal/RNA secondary structure/attenuation)

GEORGE V. STAUFFER*, GERARD ZURAWSKI, AND CHARLES YANOFSKY

Department of Biological Sciences, Stanford University, Stanford, California 94305

Contributed by Charles Yanofsky, July 28, 1978

ABSTRACT **We have isolated a set of regulatory mutants defective in transcription termination at the attenuator in the leader region of the *Escherichia coli* tryptophan (*trp*) operon. *In vivo* the mutants have 2- to 4-fold increased levels of expression of the *trp* operon above the level of the *trpR* parental strain. These levels are increased an additional 1.5- to 2-fold when the mutants are starved of tryptophan. Transcription termination at the *trp* attenuator was analyzed *in vitro* with DNA restriction fragments containing the termination-relief mutations. Whereas the frequency of readthrough transcription beyond the termination site is 5% with the wild-type DNA template, it is 46–76% when mutant DNAs are used as templates. The base change in the leader region of each mutant was determined by RNA and/or DNA sequencing. All the changes were between base pairs +116 and +132, in the G·C-rich segment of the leader region. The RNA residues between +114 and +134 of the leader transcript can form a stable stem and loop structure [$\Delta G \cong -20$ kcal (-84 kJ)]. All of the termination-relief mutations destablilize this structure ($\Delta G \cong -9.0$ to -10.5 kcal). These results suggest that the efficiency of transcription termination may be dependent on the integrity of the secondary structure of the above segment of the transcript of the leader region.**

Transcription of the *trp* operon in *Escherichia coli* is regulated at two sites, the *trp* promoter-operator and the *trp* attenuator. The frequency of transcription initiation at the promoter is regulated by the interaction of the tryptophan-activatable Trp aporepressor (the product of *trpR*) and the operator (1–4). RNA polymerase molecules that have initiated transcription at the *trp* promoter either proceed into the structural genes of the operon or terminate transcription at the attenuator, a segment of the leader region that precedes the first structural gene (5–7). In cells growing in the presence of excess tryptophan, 80–90% of the RNA polymerase molecules that initiate transcription at the *trp* promoter terminate transcription at the attenuator, whereas in cells starved of tryptophan most RNA polymerase molecules continue beyond the attenuator and transcribe the structural genes of the operon (8). Transcription termination at the attenuator is regulated in response to changes in the levels of charged and uncharged $tRNA^{Trp}$ (7, 9).

In all known transcription termination regions that are recognized by *E. coli* RNA polymerase *in vitro*, a G·C-rich region containing dyad symmetry precedes an A·T-rich region within which termination occurs (10). In this paper we show that single base-pair changes in the *trp* operon leader region which reduce the efficiency of transcription termination at the *trp* attenuator *in vivo* and *in vitro* are located in the G·C-rich segment of the leader region.

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.

MATERIALS AND METHODS

Isolation of *trpL* Mutants. *E. coli* K-12 strain W3110 *trpR trpD159am trpA9761am lacZ U118* was mutagenized with ethyl methanesulfonate or ultraviolet light (11), washed, and plated on glucose minimal medium (12) supplemented with indole (5 μg/ml) and 7-aza-DL-tryptophan (1 μg/ml). Survivors that arose as a result of nonsense suppressors, polarity suppressors, deletion of the *trpD159* mutation, or reversion of *trpD159* or *trpA9761* were discarded (13). P1 lysates prepared on the remaining isolates were plated on strain W3110 *trpR cysB* Δ*trpEA97*, and *cysB*$^+$ transductants were selected. Approximately 50% of the *cysB*$^+$ transductants from the nine independent mutants examined were able to grow on indole plus 7-aza-DL-tryptophan. This is the expected cotransduction frequency for mutations that occur in the *trp* operon (14).

Assay of Anthranilate Synthetase. Cells were grown and extracts were prepared as described (9). Anthranilate synthetase activity was determined by the procedure of Creighton and Yanofsky (15).

Preparation of Tryptophan-Starved Cells. A 40-ml culture of cells grown to 5×10^8/ml in glucose minimal medium supplemented with acid casein hydrolysate (0.05%) and L-tryptophan (50 μg/ml) was collected on a Millipore filter, washed with 2.5 vol of warmed growth medium minus tryptophan, resuspended in 1 vol of warmed growth medium minus tryptophan plus indolyl-3-acrylic acid (10 μg/ml) to induce tryptophan starvation (16), and incubated at 37° for 1.5 hr prior to harvesting.

***In Vitro* Transcription.** The *trpPOL* region from each mutant was recombined onto the *trpP*$^+$*O*$^+$Δ(*LE*)*D*$^+$ plasmid pGM3 (unpublished data). Plasmid DNA was isolated by a sodium dodecyl sulfate lysis procedure followed by ethidium bromide/cesium chloride equilibrium density gradient centrifugation (17, 18). The *Hpa* II_{570} and *Hin*f_{470} DNA restriction fragments (19), which carry the *trpPOL* region, were isolated from the above plasmids by the procedures described by Bennett *et al.* (19). *In vitro* transcription reactions were performed as described by Lee and Yanofsky (20).

Sequence Analysis of *trpL* Region. RNA sequence analysis of the leader transcripts was performed as described (21). The DNA sequence of the *trpL* region of each mutant was determined by the method of Maxam and Gilbert (22). The details of this procedure and related techniques used in this analysis were as described (19).

All experiments using recombinant DNA were performed in accordance with the Guidelines of the National Institutes of Health.

* Present address: Department of Microbiology, The University of Iowa, Iowa City, IA 52242.

RESULTS

To identify the segment of the leader region of the *E. coli trp* operon essential for transcription termination, we isolated and analyzed leader point mutants defective in the termination process. The mutations were isolated in a *trpR trpD* polar *trpA* nonsense mutant in which high levels of tryptophan synthetase β_2 are required for utilization of indole (23). This strain cannot grow on indole because it lacks tryptophan synthetase α and has low tryptophan synthetase β_2 levels due to the polar effect of the *trpD* mutation. The tryptophan analog 7-aza-DL-tryptophan increases the stringency of the selection by increasing transcription termination at the attenuator (7) and by competing with tryptophan in the charging of $tRNA^{Trp}$ and for incorporation into protein (24).

The *trpL* mutations were recognized by screening mutants for those with changes linked to the *trp* operon which were responsible for growth on indole plus 7-aza-DL-tryptophan and by eliminating those which did not retain the parental *trpD* and *trpA* alterations (*Materials and Methods*). Nine independent strains with the above characteristics, which also had 2- to 4-fold increased levels of *trpE* expression when grown with excess tryptophan (Table 1), were chosen for study. Each of these strains responded to tryptophan starvation by producing higher levels of anthranilate synthetase (Table 1). However, in some cases the increase was small, probably because the strains already had increased levels of this enzyme.

To distinguish between mutations that increase the frequency of transcription initiation and those that relieve transcription termination at the *trp* attenuator, we measured readthrough transcription at the *trp* attenuator *in vitro* with DNA restriction fragments carrying the *trp* regulatory region of each of the nine mutants. When the *Hpa* II_{570} DNA fragment carrying the wild-type *trp* regulatory region is transcribed by RNA polymerase in a purified *in vitro* system, 95% of the RNA transcripts are the terminated, 140-nucleotide-long species (20). The remaining transcripts are 260 nucleotides long and are formed by RNA polymerase molecules which readthrough to the end of the DNA fragment (Fig. 1). When *Hpa* II_{570} DNA fragments prepared from the mutant plasmids were transcribed *in vitro*, 46–76% of the RNA molecules synthesized were readthrough transcripts (Fig. 2 and Table 1). This demonstrates that each of the nine mutations decreases the extent of transcription termination at the *trp* attenuator.

We determined the alteration in each mutant by using the DNA sequencing method of Maxam and Gilbert (22). The mutant DNAs were cut at the *Hin*fI site (base pair +180, Fig. 1) and the sequence for approximately 100 residues towards the transcription start site was determined. A single base-pair change was found in the leader region of each mutant (Table 1). In *trpL80*, *trpL81*, and *trpL82* the base-pair change was determined by RNA sequencing (21) and confirmed by DNA sequencing (22). All of the changes were located between base pairs +116 and +132 of the leader region (Table 1), in the G·C-rich segment of the attenuator.

DISCUSSION

Transcription of the *trp* operon of *E. coli*, both *in vivo* and *in vitro*, is generally terminated in a series of A·T base pairs located about 140 base pairs from the site of transcription initiation (refs. 8 and 20, and Fig. 1). *In vivo*, the extent of transcription termination is regulated in response to changes in the levels of charged and uncharged $tRNA^{Trp}$ (7, 9) and appears to be influenced by mutations affecting the termination release factor ρ (13). The event signalling transcription termination presumably involves the recognition by RNA polymerase of some feature of the template, the transcript, or the template–transcript complex. A G·C-rich region in the leader region immediately precedes the A·T-rich region within which transcription termination occurs (Fig. 1). Previous studies have shown that deletions entering this G·C-rich region from the structural gene side give full relief of transcription termination *in vivo* and *in vitro* (8, 27). One similar deletion with a terminus in the A·T-rich region partially relieves termination *in vivo*, and totally relieves it *in vitro* (27). These results suggest that both regions play roles in the transcription termination process. We have isolated and characterized single base-pair mutations in the *trp* leader region that relieve transcription termination at the *trp* attenuator. These mutations are all located in the G·C-rich segment of the *trp* leader region, allowing us to define the features of this region that may serve to signal transcription termination. None of these single base-pair changes relieves termination completely, suggesting that multiple base changes in the G·C-rich region may be necessary for full termination-relief.

The precise nature of the transcription termination signal is unknown. At least three plausible alternatives are apparent: (*i*) polymerase may pause or stop in the G·C-rich region, and this may cause termination in the A·T-rich region immediately beyond; (*ii*) a sequence of bases, in the DNA template, the RNA transcript, or the transcript–template complex may be recog-

Table 1. Characteristics of termination-relief mutants

trpL mutation*	Mutagen	Anthranilate synthetase† Trp excess	Anthranilate synthetase† Trp starved	% readthrough transcription	Base pair change	Position	ΔG values‡ First stem and loop (residues +74 to +119)	ΔG values‡ Second stem and loop (residues +119 to +134)
trpL⁺	—	0.42	1.41	5	—	—	−12.2	−20.1
trpL80	EMS	1.23	2.12	76	ΔCG	+116§	−4.4	−9.3
trpL150	EMS	1.27	2.59	69	ΔCG	+116§		
trpL78	EMS	1.11	2.40	43	CG → TA	+117	−2.6	−10.5
trpL81	EMS	0.99	2.03	58	CG → TA	+117		
trpL82	EMS	1.24	1.82	62	CG → TA	+117		
trpL151	EMS	1.21	2.99	63	CG → TA	+117		
trpL77	EMS	1.15	2.46	76	CG → AT	+118	−4.4	−10.3
trpL153	EMS	1.61	2.40	74	GC → AT	+132	−12.2	−9.0
trpL154	UV	1.61	2.52	ND	GC → TA	+132	−12.2	−9.0

EMS, ethyl methanesulfonate; ND, not determined.

* In addition to the *trpL* mutation, each strain is also *trpR trpD159 trpA9761*.

† One unit of specific activity corresponds to the conversion of 0.1 μmol of chorismate to anthranilate in 20 min at 37°C per mg of protein.

‡ ΔG values were calculated according to Tinoco and coworkers (25, 26).

§ Deletion of a C·G base pair between +116 and +118.

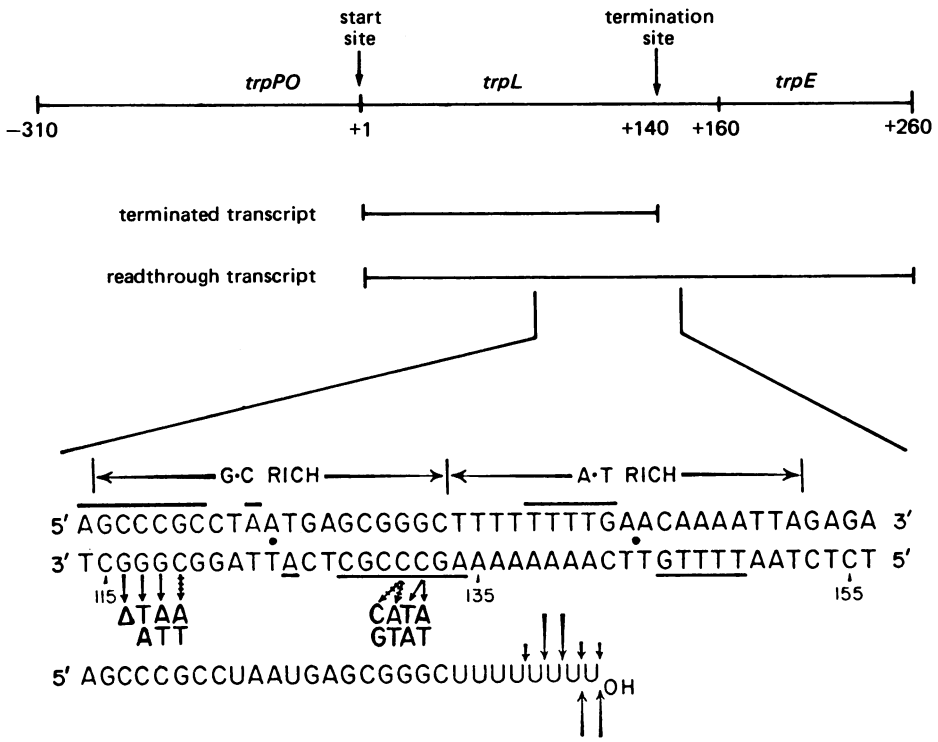

FIG. 1. Regulatory region of the tryptophan operon of *E. coli* and the locations of termination-relief mutations. Shown is the *Hpa* II_{570} DNA fragment which contains the *trp* promoter-operator, the site of transcription initiation (base pair +1), the 160 base-pair leader region (*trpL*) containing the transcription termination site (base pair +140), and the initial 100 base pairs of *trpE*, the first structural gene of the operon. Transcription *in vitro* of this fragment yields two major products; the 140-residue terminated leader transcript and the 260-residue readthrough transcript. The DNA nucleotide sequence in the region of the transcription termination site is also shown. It includes both a G·C-rich region and an A·T-rich region containing dyad symmetry. The centers of symmetry are marked by dots; the extent of the symmetry is indicated by the bars above and below the sequence. The RNA nucleotide sequence shown below the DNA sequence is the 3′-OH end of the terminated leader transcript. Arrows above the U residues indicate the termini *in vivo*; arrows below the U residues mark the termini *in vitro* (27). Base pair changes in the termination-relief mutants described in this paper are indicated by arrows. Δ identifies a single G·C base pair deletion at position 116, 117, or 118. Three termination-relief mutations isolated in parental strains with increased transcription termination (unpublished data) are indicated by crosshatched arrows.

nized by RNA polymerase as the signal to terminate transcription; or (*iii*) a structural feature of the transcript other than its nucleotide sequence may be recognized by RNA polymerase as the termination signal. We shall discuss these alternatives in the following paragraphs.

(*i*) RNA polymerase does pause and terminate transcription in the G·C-rich segments of natural and synthetic DNA templates (28). In fact, a feature common to all ρ-independent terminators is the existence of a G·C-rich region immediately preceding an A·T-rich region within which termination occurs (10). Gilbert (28) and Neff and Chamberlin (29) have suggested that since RNA·DNA hybrids with high G·C content are more stable than the corresponding DNA·DNA hybrids, RNA polymerase will pause in regions of high G·C-content, perhaps thus becoming amenable to termination. The termination-relief mutants described in Table 1 all reduce the G·C content of the region between base pairs +114 and +134. This finding is therefore consistent with the view that the G·C content of the template is crucial to transcription termination. However, all but one of our mutations were isolated after mutagenesis with ethyl methanesulfonate, an agent favoring GC → AT transitions (30). In other studies we have isolated similar termination-relief mutants without mutagenesis, and one of these has a GC → CG transversion at base pair +131 in the *trp* leader region (see Figs. 1 and 3; unpublished data). The properties of this mutant are indistinguishable from those of the mutants described in Table 1. This result suggests that G·C content alone is not sufficient to signal transcription termination (see also *iii* below).

(*ii*) A comparison of the nucleotide sequences of the termination regions of the *trp* operons of *E. coli* (20), *Shigella dysenteriae* (31), *Salmonella typhimurium* (20), and *Serratia marcescens* (unpublished data) and the *phe* operon of *E. coli* (32) reveals that the sequence $GCCN_{12}GG/UGCN_{0-1}U_6$ is conserved. Fig. 1 shows this conserved region from the *E. coli trp* operon leader region (residues +115 to +141). This figure also shows that all of the termination-relief mutations fall in the conserved region. Thus, if RNA polymerase does recognize a nucleotide sequence as a termination signal, it is evident that two regions, base pairs +116 to +119 and base pairs +131 and +132, contribute greatly to this signal. Alternatively, conservation of this sequence of the G·C-rich region may be indicative of a function in termination other than as the termination signal.

(*iii*) The G·C-rich regions of ρ-independent transcription termination sites all have dyad symmetry (10). Because of this symmetry, hairpin loops can form in the RNA transcript. In the *E. coli* and *S. typhimurium trp* leader transcripts, two alternative stem-loop structures can be formed (ref. 20; Fig. 3). Base pairing consistent with the existence of such hairpin loops has been observed *in vitro* (20). The regions involved are residues +74 to +119 or, alternatively, residues +114 to +134 (Fig. 3). Table 1 shows that seven of the termination-relief mutations analyzed destabilize both stem and loop structures, whereas the changes in mutants *trpL153* and *trpL154* destabilize only the second stem and loop (residues +114 to +134). These two mutants exhibit 4-fold relief of transcription termination *in vivo* compared to the 2- to 3-fold termination relief observed with mutants with changes that destabilize both stems. Although this

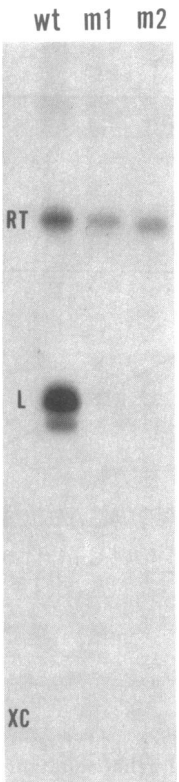

FIG. 2. RNA transcribed from *Hpa* II_{570} DNA fragments prepared from wild-type or from termination-relief mutant strains. The RNA was resolved on a 10% polyacrylamide/7 M urea/TBE gel as described (20). L, 140-residue leader transcript; RT, 260-residue readthrough transcript; wt, the template was the *Hpa* II_{570} fragment prepared from wild-type DNA; m, the template was a *Hpa* II_{570} fragment prepared from DNA from a mutant strain; ml is *trpL151*; m2 is *trpL153*. XC, position of xylene cyanol dye marker.

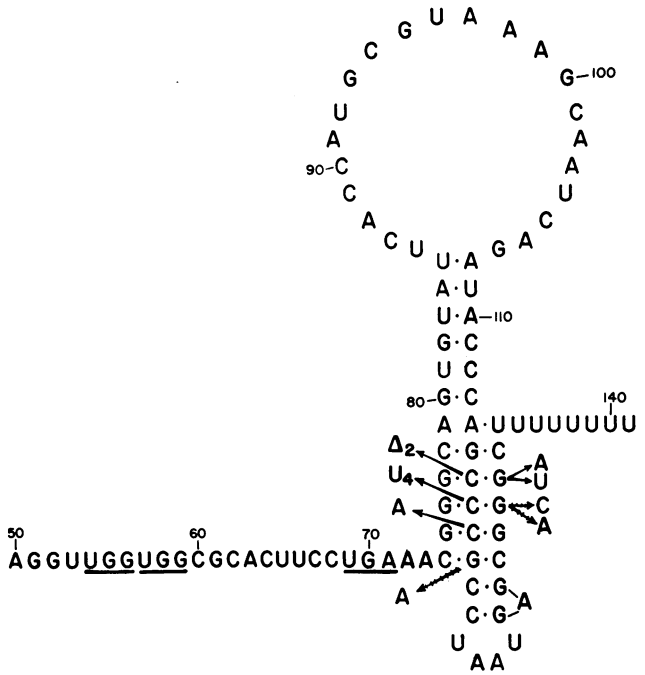

FIG. 3. Transcript locations of the nucleotide changes in termination relief mutants. The 3'-OH half of the *trp* leader transcript is shown as the double stem-loop structure presented by Lee and Yanofsky (20). Arrows are explained in the legend for Fig. 1. Subscript numerals 2 and 4 indicate that the changes indicated occurred twice and four times, respectively.

difference is not seen *in vitro* (Table 1), these results suggest that the inability to form a stable second stem and loop is conducive to termination relief and, hence, that this structure may function as the termination signal. This interpretation is supported by recent studies on transcription termination at the ρ-dependent termination site of phage λ, t_{R1} (10). Two mutations that destabilize a potential stem and loop structure in the termination region result in decreased termination efficiencies. Furthermore, these mutations (AT → GC and TA → CG) increase the G-C richness of the termination region, supporting the view that G-C content alone is not sufficient for transcription termination.

In other studies with the *trp* operon we have determined whether the attenuator is a transcription pause site. The RNA polymerase–DNA template–leader RNA complex does not dissociate *in vitro* when transcription terminates at the *trp* attenuator (F. Lee and C. Yanofsky, unpublished data). To determine whether such polymerase molecules have paused rather than stopped, we added rifampicin to block reinitiation and measured the relative amounts of the 140-nucleotide terminated leader transcript and the 260-nucleotide readthrough transcript during a 0- to 60-min incubation period. The 140-nucleotide RNA species could not be chased into the 260-nucleotide species when either wild-type or mutant *Hpa* II_{570} DNA fragments were used as templates (unpublished data). These findings indicate that polymerase stops rather than pauses at the *trp* attenuator.

The extent of transcription termination in the leader region of the *trp* operon of *E. coli* is regulated by a mechanism termed attenuation (9, 33). A logical means of accomplishing this regulation is either to alter or expose the structure that serves as the signal for transcription termination. The studies described in this report suggest that a hairpin loop formed between residues +114 to +134 may serve as the termination signal. Transcription termination at the attenuator is relieved by tryptophan starvation (resulting in a deficiency of charged $tRNA^{Trp}$) and it is thought that the stalling of ribosomes at the tandem Trp codons of the coding region of the leader transcript is responsible for this relief (20). How this is accomplished is not known, but recent studies with the *trp* and *phe* operons (32) have led to the proposal that interaction between the Trp or Phe coding region and a distal segment of the transcript interferes with RNA polymerase's recognition of the termination signal.

These studies were supported by grants from the U.S. Public Health Service (GM 09738), the National Science Foundation (PCM 77-24333), and the American Heart Association (69C-15). G.S. is a postdoctoral trainee of the U.S. Public Health Service. G.Z. is a Career Investigator Fellow of the American Heart Association. C.Y. is a Career Investigator of the American Heart Association.

1. Bennett, G. N., Schweingruber, M. E., Brown, K. D., Squires, C. & Yanofsky, C. (1976) *Proc. Natl. Acad. Sci. USA* **73**, 2351–2355.
2. McGeoch, D., McGeoch, J. & Morse, D. (1973) *Nature (London) New Biol.* **245**, 137–140.
3. Rose, J. K., Squires, C. L., Yanofsky, C., Yang, H.-L. & Zubay, G. (1973) *Nature (London) New Biol.* **245**, 133–137.
4. Squires, C. L., Lee, F. & Yanofsky, C. (1975) *J. Mol. Biol.* **92**, 93–111.
5. Bertrand, K., Korn, L., Lee, F., Platt, T., Squires, C. L., Squires, C. & Yanofsky, C. (1975) *Science* **189**, 22-26.
6. Bertrand, K. & Yanofsky, C. (1976) *J. Mol. Biol.* **103**, 339–349.
7. Morse, D. E. & Morse, A. N. C. (1976) *J. Mol. Biol.* **103**, 209–226.
8. Bertrand, K., Squires, C. & Yanofsky, C. (1976) *J. Mol. Biol.* **103**, 319–337.
9. Yanofsky, C. & Soll, L. (1977) *J. Mol. Biol.* **113**, 663–677.

10. Rosenberg, M., Court, D., Shimatake, H., Brady, C. & Wulff, D. L. (1978) *Nature (London)* **272,** 414–423.
11. Miller, J. (1972) *Experiments in Molecular Genetics* (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY).
12. Vogel, H. J. & Bonner, D. M. (1956) *J. Biol. Chem.* **218,** 97–106.
13. Korn, L. J. & Yanofsky, C. (1976) *J. Mol. Biol.* **103,** 395–409.
14. Yanofsky, C. & Lennox, E. S. (1959) *Virology* **8,** 425–447.
15. Creighton, T. E. & Yanofsky, C. (1970) *Methods Enzymol.* **17A,** 365–380.
16. Morse, D. E., Mosteller, R., Baker, R. F. & Yanofsky, C. (1969) *Nature (London)* **223,** 40–43.
17. Guerry, P., LeBlanc, D. J. & Falkow, S. (1973) *J. Bacteriol.* **116,** 1064–1066.
18. Selker, E., Brown, K. & Yanofsky, C. (1977) *J. Bacteriol.* **129,** 388–394.
19. Bennett, G. N., Schweingruber, M. E., Brown, K. D., Squires, C. & Yanofsky, C. (1978) *J. Mol. Biol.* **121,** 113–137.
20. Lee, F. & Yanofsky, C. (1977) *Proc. Natl. Acad. Sci. USA* **74,** 4365–4369.
21. Squires, C., Lee, F., Bertrand, K., Squires, C. L., Bronson, M. J. & Yanofsky, C. (1976) *J. Mol. Biol.* **103,** 351–381.
22. Maxam, A. & Gilbert, W. (1977) *Proc. Natl. Acad. Sci. USA* **74,** 560–564.
23. Jackson, E. N. & Yanofsky, C. (1972) *J. Mol. Biol.* **71,** 149–161.
24. Mosteller, R. D. & Yanofsky, C. (1971) *J. Bacteriol.* **105,** 268–275.
25. Tinoco, I., Jr., Borer, P. N., Dengler, B., Levine, M. D., Uhlenbeck, O. C., Crothers, D. M. & Bralla, J. (1973) *Nature (London) New Biol.* **246,** 40–41.
26. Borer, P. N., Dengler, B., Tinoco, I., Jr. & Uhlenbeck, O. C. (1974) *J. Mol. Biol.* **85,** 843–853.
27. Bertrand, K., Korn, L. J., Lee, F. & Yanofsky, C. (1977) *J. Mol. Biol.* **117,** 227–247.
28. Gilbert, W. (1976) in *RNA Polymerase,* eds. Losick, R. & Chamberlin, M. (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY), pp. 193–205.
29. Neff, N. F. & Chamberlin, M. J. (1978) *J. Biol. Chem.* **253,** 2455–2460.
30. Coulondre, C. & Miller, J. H. (1977) *J. Mol. Biol.* **117,** 525–567.
31. Miozzari, G. F. & Yanofsky, C. (1978) *Proc. Natl. Acad. Sci. USA* **75,** in press.
32. Zurawski, G., Brown, K., Killingly, D. & Yanofsky, C. (1978) *Proc. Natl. Acad. Sci. USA* **75,** 4271–4275.
33. Kasai, T. (1974) *Nature (London)* **249,** 523–527.