

B M B 400
Part Four: Gene Regulation
Section IV = Chapter 18
REGULATION AFTER INITIATION OF TRANSCRIPTION

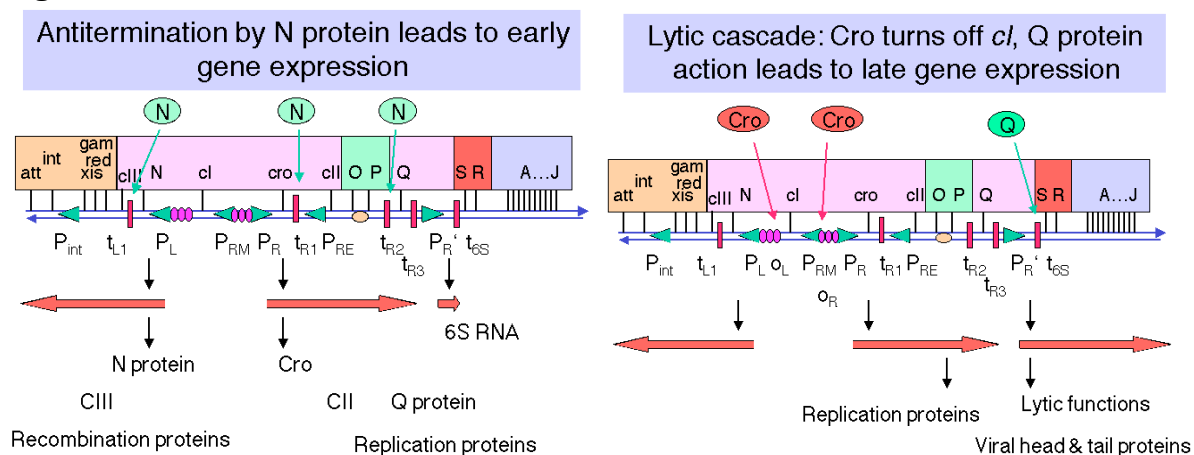
Although regulation of the initiation of transcription appears to be a dominant factor in control of expression of many genes, the importance of regulation after initiation is becoming better appreciated in an increasing number and variety of systems. The classic systems in which these issues have been explored are antitermination in bacteriophage λ and in attenuation of transcription in bacterial biosynthetic operons, in particular the *trp* operon in *E. coli*. Although some of the mechanistic details may be peculiar to bacteria, especially the need for coupled transcription and translation in the *trp* attenuation system, the phenomenon of regulation after initiation is seen in a wide variety of organisms, ranging from bacteria to humans. Some of this work was discussed in the sections on elongation of transcription in Chapter II of Part Three.

Both systems discussed in this chapter control the frequency of **termination** of transcription. Antitermination in bacteriophage λ can prevent RNA polymerase from stopping at **ρ -dependent** terminators, thus leading to transcription of downstream genes. Attenuation in the *trp* operon also controls the frequency at which RNA polymerase stops at an early terminator in the operon, hence regulating the transcription of downstream genes. In contrast to the system in λ , attenuation in *trp* regulates termination at a **ρ -independent** terminator.

A. Antitermination in bacteriophage λ

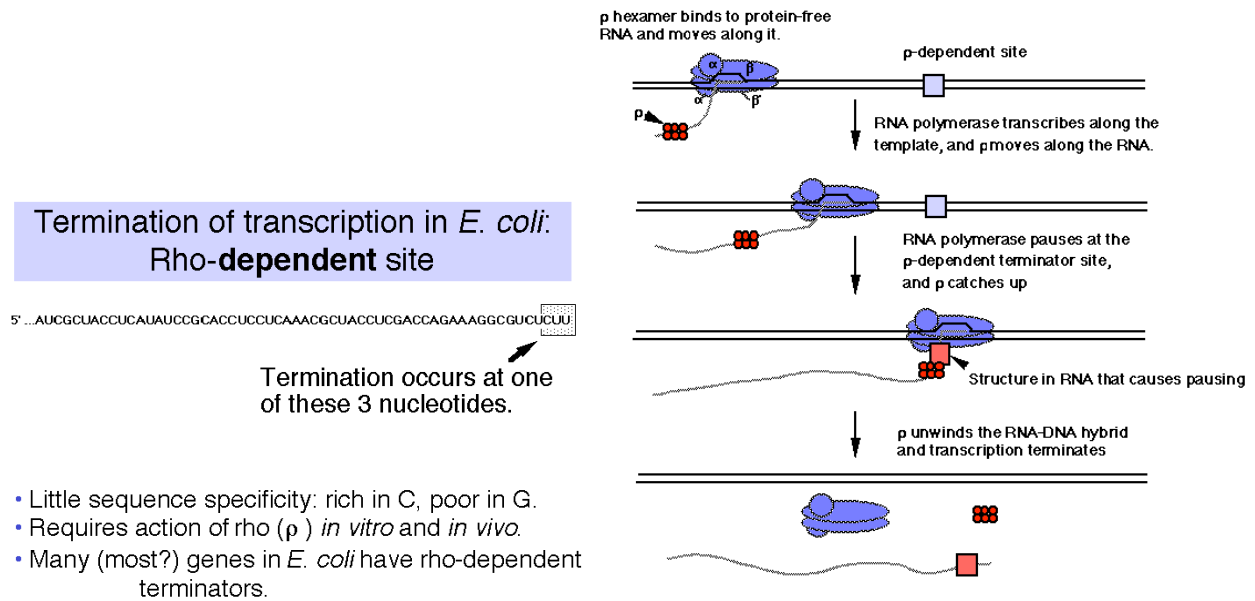
1. Just to quickly review one of the points in Chapter III, **antitermination occurs at two different times in the λ life cycle**. The N protein allows read-through transcription in the shift from immediate-early to early transcription, and the Q protein allows read-through transcription of the late genes.

Fig. 4.4.1



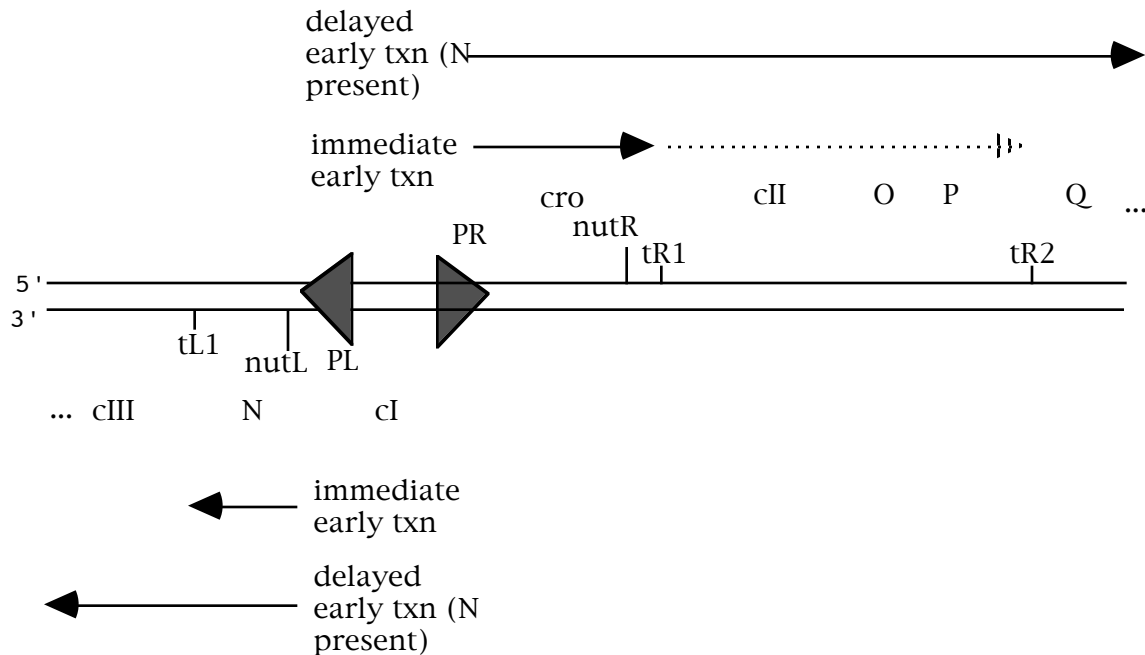
2. Recall from Part Three of the text that ρ -dependent terminators do not have a well-conserved sequence or secondary structure. Also, the protein ρ tracks along protein-free regions of the RNA until it hits a paused transcriptional complex at a ρ -dependent site, at which point its RNA helicase activity can cause termination and dissociation of the polymerase and transcript from the template DNA.

Fig. 4.4.2. Action of ρ protein at terminators of transcription

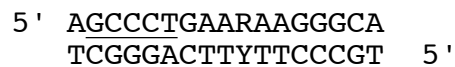


3. **Sites on the DNA needed for antitermination in bacteriophage λ**

- nut* sites (**N** utilization sites) for pN, *qut* site for pQ.
- The *nut* sites are within the transcription unit, not at the promoter and not at the terminator.
 - (1) *nutL* is in the 5' untranslated region of the *N* gene, and *nutR* is in the 3' untranslated region of the *cro* gene.
 - (2) In both cases, the *nut* site precedes the terminator at which pN will act.

Figure 4.4.3. nut sites are located within transcription units

(3) Both *nutL* and *nutR* are 17 bp sequences with a dyad symmetry.



c. Model

The protein pN recognizes the *nut* site and binds to RNA polymerase as it transcribes through the site. The complex of pN with the RNA polymerases is highly processive and overrides the efforts of ρ at the terminator.

4. *E. coli* (host) proteins needed for action of pN.

- a. These were isolated as host functions that when mutated prevented action of pN.
- b. NusA (encoded by *nusA*, for N utilization substance, complementation group A) is the best characterized.

(1) Can form part of the transcription complex

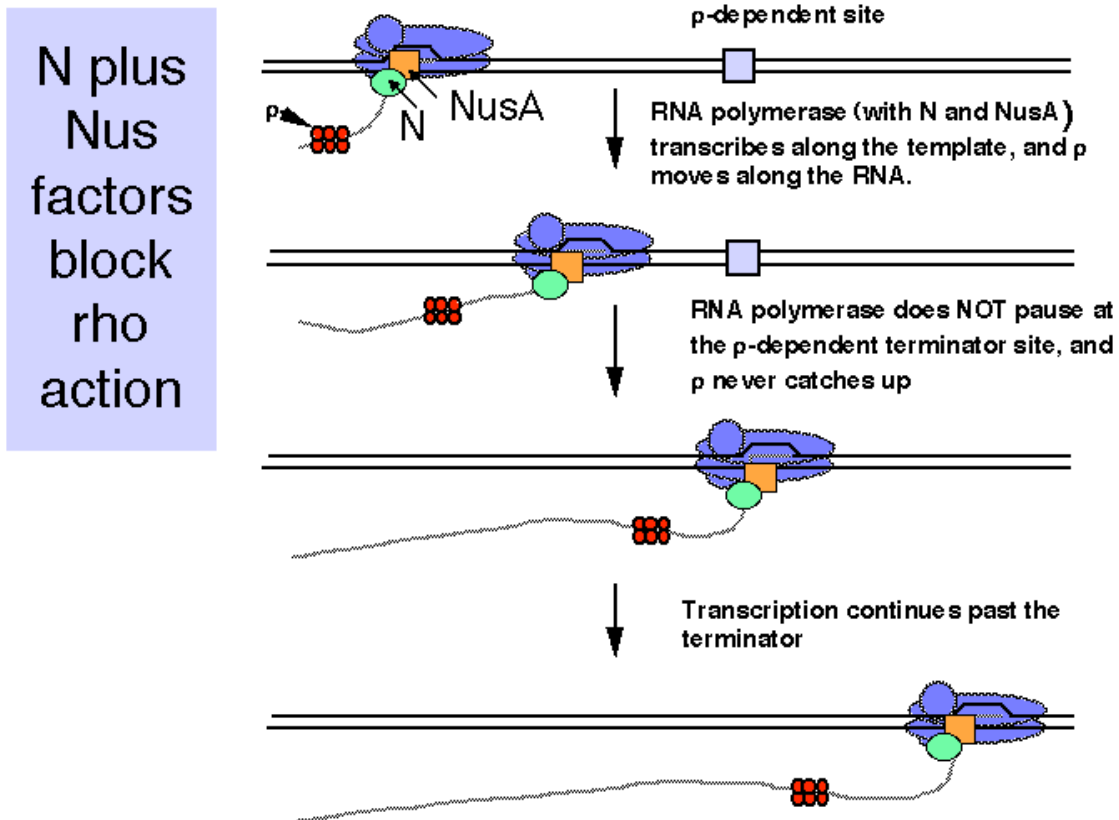
(2) Has been proposed to bind to the core RNA polymerase after σ dissociates.

(3) Can also bind pN.

(4) Model:

NusA binds the core polymerase after σ dissociates. As this complex transcribes through a *nus* site, pN binds also. The complex $\alpha_2\beta\beta'$ -NusA-pN prevents ρ -dependent termination at *tR1*, *tR2*, and *tL1*.

Figure 4.4.4.



c. Several other *nus* genes have been identified,

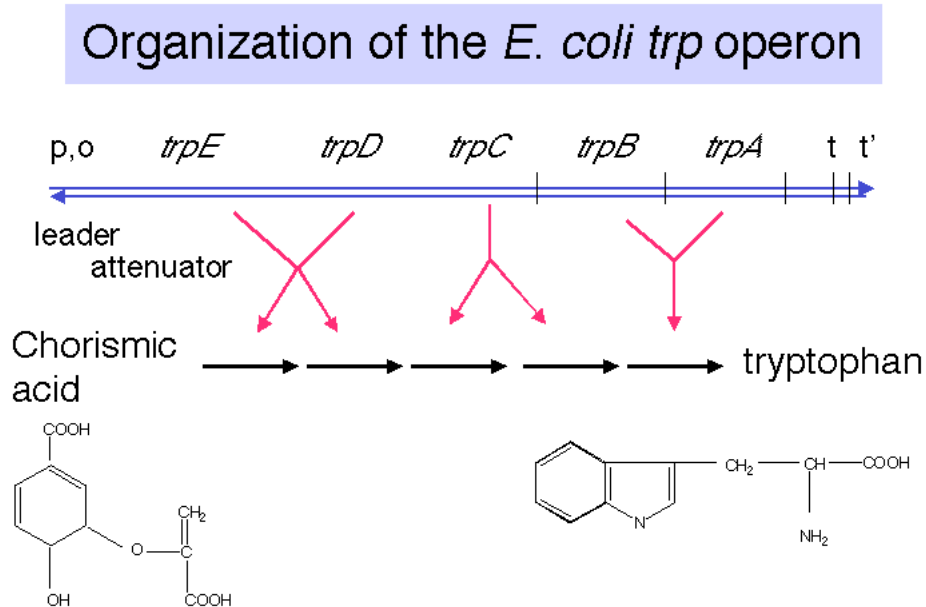
NusG is the bacterial homolog of a family of conserved proteins involved in elongation. It is homologous to the large subunit of DSIF, which is an elongation factor in mammals. DSIF is the DRB-sensitivity inducing factor. Current studies implicate it in both negative and positive effects on elongation. It has two subunits, one of 160 kDa that is homologous to the yeast transcriptional regulatory protein Spt5, and one of 14 kDa that is homologous to the yeast Spt4 protein.

Another *nus* gene encodes a ribosomal protein. Much more needs to be learned about both termination and antitermination. The *nus* phenotype of mutations in a gene encoding a ribosomal protein suggests that translation is also coupled to this process.

B. Components of the *E. coli trp* operon

1. The *trp* operon encodes the enzymes required for biosynthesis of tryptophan. More specifically, its five genes (*trpEDCBA*) encode five subunits of proteins that in total catalyze five enzymatic steps, converting chorismic acid to tryptophan. However, there is not a 1:1 correspondence between a cistron and an enzyme. For example, *trpB* and *trpA* encode, respectively, the β and α subunits of tryptophan synthase, which catalyzes the replacement of glycerol-3-phosphate from indole-3-glycerol-phosphate with serine to form tryptophan, with glyceraldehyde-3-phosphate as the other product

Figure 4.4.5.



2. A leader sequence separates the promoter and operator from the first structural gene of the operon, *trpE*.
3. An attenuator of transcription follows the leader. As we will see in more detail below, the efficiency of "premature" termination at this attenuator is determined by the extent of translation of the leader, which in turn is determined by the availability of Trp-tRNA^{trp}. This is an important part of the regulation of the operon.
4. Two terminators of transcription follow the structural genes, one dependent on ρ and one independent of ρ .

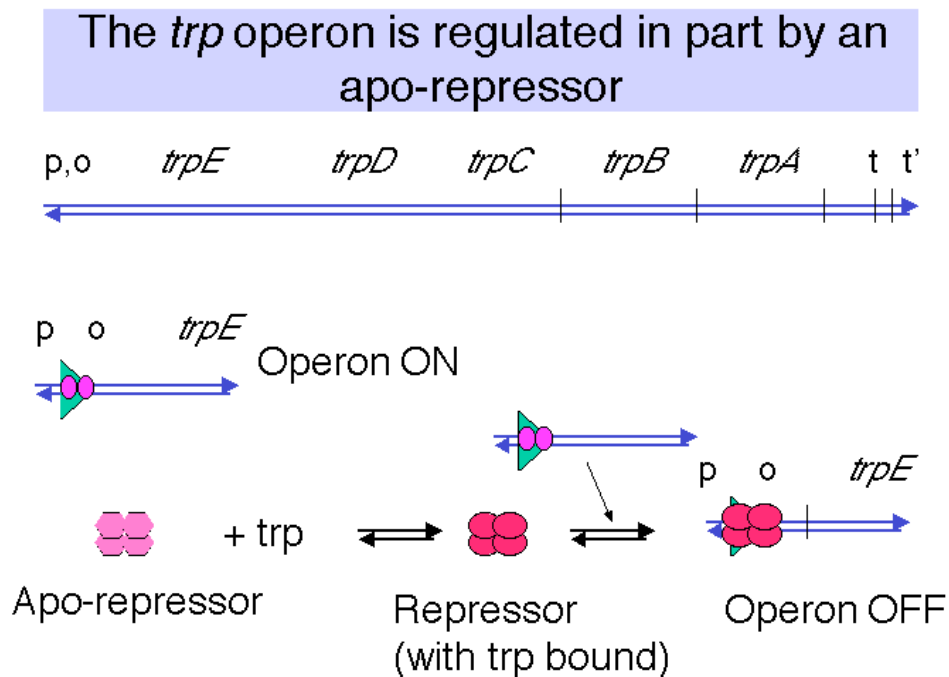
C. Modes of regulation: turn operon off in presence of Trp

1. Repressor-operator: requires a protein binding to a specific site in the presence of Trp to decrease the efficiency of initiation of transcription.
2. Attenuation: the elongation (and termination) of transcription by RNA polymerase is linked to the progress of translation by a ribosome. In the presence of Trp, the translation by the ribosome causes transcription of the subsequent genes in the operon to terminate.

D. Repressor: apo-repressor and co-repressor (Trp)

1. The apo-repressor is encoded by *trpR* at a distant locus. The apo-repressor is a homo-tetramer. It has a high affinity for the operator only when it is bound by the amino acid Trp, which serves as a co-repressor. Thus the active repressor is a tetramer of (formerly apo-) repressor in complex with Trp. The active repressor binds to the operator to prevent initiation of transcription.
2. The operator overlaps the promoter, including the -10 region of the promoter. It has a dyad axis of symmetry.

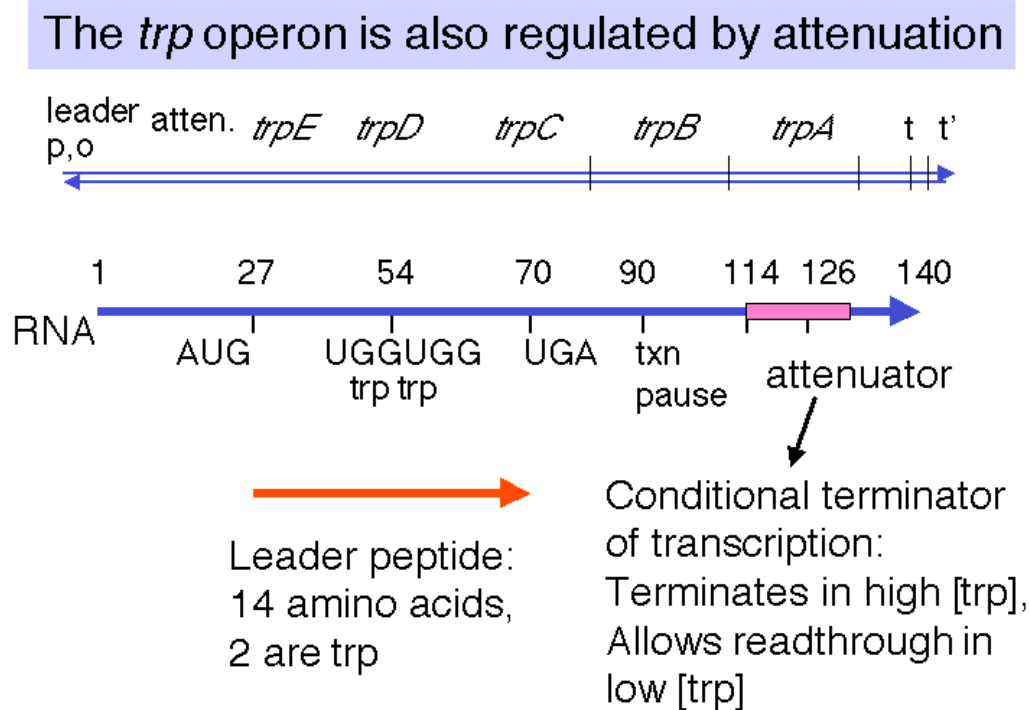
Figure 4.4.6.



E. Attenuation

1. The **attenuator** is a **conditional transcriptional terminator used to regulate expression** of biosynthetic operons in bacteria.
 - a. The attenuator is upstream of the structural genes *trpEDCBA*.
 - b. It is a ρ -independent termination site. Its ability to terminate transcription is dependent on its ability to form the stem of duplex RNA that is characteristic of ρ -independent termination sites.

Fig. 4.4.7



2. **The fraction of transcripts that read through the attenuator is determined by the [Trp-tRNA^{trp}].**
 - a. The concentration of charged tRNAs is a measure of the amount of Trp available for protein synthesis. If most tRNA^{trp} is charged, there is an abundance of Trp, and the cell does not need to make more.
 - b. Low [Trp-tRNA^{trp}] allows read-through transcription through the attenuator, so that *trpEDCBA* is expressed.
 - c. High [Trp-tRNA^{trp}] causes termination of transcription at the attenuator.

3. **The [Trp-tRNA^{trp}] determines the progress of ribosomes as they translate a short leader peptide.**
 - a. The leader peptide is a short 14 amino acid polypeptide encoded by *trpL*.
 - b. Two codons for Trp are in the leader, and the progress of ribosomes past these Trp codons will be determined by the availability of Trp-tRNA^{trp}. When the concentration of tryptophanyl-tRNA is high, translation of the *trp* leader will be completed, but when it is low, translation will stall at the tryptophan codons.

4. **The extent of progress of the ribosomes determines the secondary structures formed in the leader RNA.**
 - a. When the [Trp-tRNA^{trp}] is high, the ribosomes translate past the Trp codons to complete the synthesis leader of the peptide. This allows the nascent RNA to form the structure for ρ -independent terminator. Thus transcription terminates before the RNA polymerase reaches *trpEDCBA*.
 - b. When the [Trp-tRNA^{trp}] is low, the ribosomes stall at the Trp codons, which prevents formation of the secondary structures in the RNA necessary for termination at the attenuator. Thus read-through transcription continues through *trpEDCBA* and the operon is expressed, so that more Trp is made.

Table 4.4.1. The basic components of regulation at the attenuator of the *E. coli trp* operon are tabulated below.

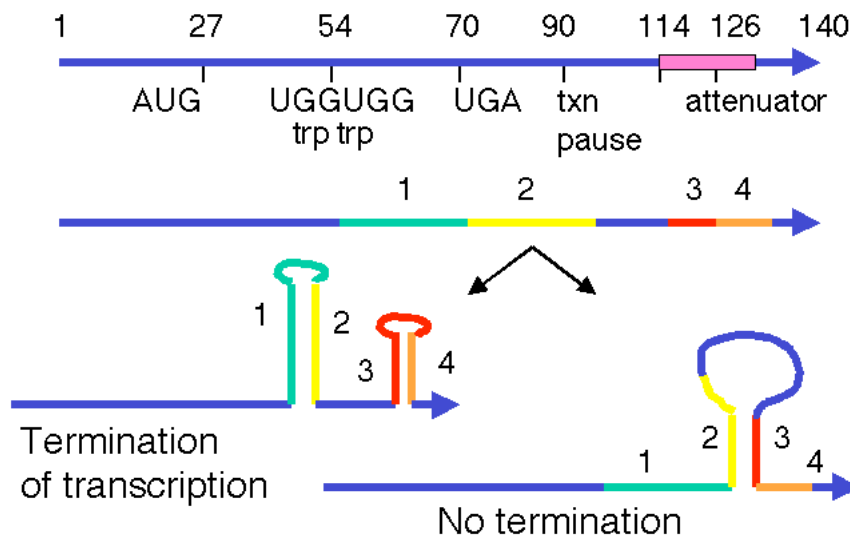
[trp-tRNA]	translation of <i>trpL</i>	secondary structures formed in RNA	Attenuator	Operon
High	complete	3-4 stem	terminate transcription	OFF
Low	stalls at Trp codons	2-3 stem	allow read-through transcription	ON

5. Alternative base-paired structures in leader RNA

- Four regions of the leader RNA can be involved in secondary structure formation, in particular base-paired stems. These are referred to simply as regions 1, 2, 3, and 4.
- Potentially, 1 can pair with 2, 2 can pair with 3, and 3 can pair with 4.

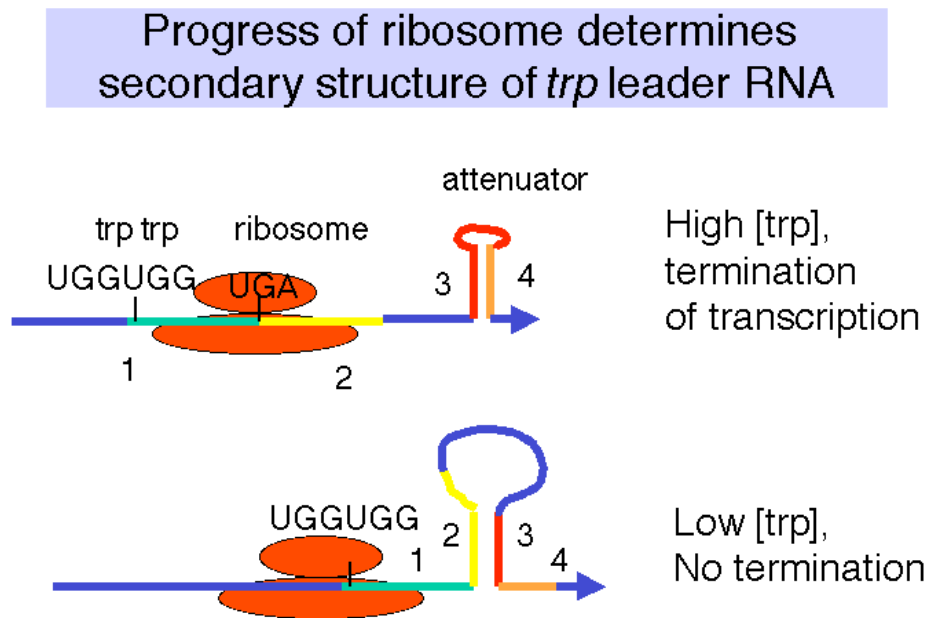
Fig. 4.4.8

Alternative base-paired structures in leader RNA



- A stem formed by pairing between 3 and 4 makes a G+C rich stem followed by U's, which is sufficient for ρ -independent termination of transcription. When the $[\text{Trp-tRNA}^{\text{trp}}]$ is high, the 3-4 base-paired structure forms, and transcription terminates at the attenuator. This turns the operon OFF.
- The formation of a base-paired stem between regions 2 and 3 precludes formation of the 3-4 terminator, and transcription will continue into the structural genes *trpEDCBA*. This turns the operon ON.

Fig. 4.4.9



- e. The choice between a 2-3 stem or a 3-4 stem is dictated by the progress of the ribosome.
- (1) If the ribosome can translate past the Trp codons (when the [Trp-tRNA^{Trp}] is high), then it will reach a natural translation termination codon. When the ribosome is in that position, region 2 of the leader RNA is covered by the ribosome, so the 2-3 stem cannot form but the 3-4 stem can. This generates the secondary structure needed for termination of transcription at the attenuator.
 - (2) In contrast, if the ribosome stalls at the Trp codons in the leader, because the [Trp-tRNA^{Trp}] is low, then region 2 of the leader RNA is not covered by the ribosome. It can then base pair with region 3. This prevents formation of the 3-4 terminator, and RNA polymerase can continue elongation through *trpEDCBA*.

E. Mutational analysis (selected examples)

1. Translation of *trpL* is needed for regulation by attenuation

Mutation of the AUG for initiation of translation of the leader RNA prevents transcription past the attenuator.

In the absence of translation, both the 1-2 and 3-4 stems can form. The latter 3-4 stem is the terminator.

2. Charged tRNA^{trp} is required for regulation

Mutation of the genes for tRNA^{trp} or Trp-tRNA^{trp} synthetase leads to constitutive expression of *trpEDCBA*.

In these mutants, translation will stall at Trp codons regardless of the intracellular [Trp], and no terminator will form at the attenuator.

3. Specific secondary structures in the *trp* leader RNA are needed for regulation

E.g. mutations that decrease the number of base pairs between the 3 and 4 regions will decrease the amount of transcriptional termination (i.e. increase expression of the operon). Compensatory mutations that increase the number of base pairs between 3 and 4 will suppress the original mutations.

F. Attenuation requires coupled transcription and translation

1. Requires no regulatory proteins: charging of cognate tRNA is the regulatory signal.
2. Need a transcriptional pause site at +90 to allow the ribosomes to catch up with the RNA polymerase and thereby affect the secondary structures in the nascent RNA.

G. Attenuation is a common mechanism for regulating biosynthetic operons

Many operons that encode the enzymes catalyzing biosynthesis of amino acids are regulated by attenuation. In each case, the leader polypeptide is rich in the amino acid that is the product of the pathway.

E.g. *his*, *phe*, *leu*, *thr*, *ilv*.

Additional readings

Friedman, D.I. and Count, D.L. (1995) Transcriptional antitermination: The lambda paradigm updated. *Molecular Microbiology* 18: 191-200.

Henkin, T. (2000) Transcriptional termination in bacteria. *Current Opinions in Microbiology* 3: 149-153.

Gusarov, I. and Nudler, E. (2001) Control of intrinsic transcriptional termination by N and NusA: The basic mechanism. *Cell* 107: 437-449.

Questions on Chapter 18. Regulation after initiation of transcription

- 18.1** Which of the following statements concerning the action of N protein are true?
- 1) N action requires sequences on the λ DNA called nut_L^+ and nut_R^+ .
 - 2) N activity requires a host function encoded by $nusA^+$.
 - 3) N protein acts to promote *rho*-dependent termination.
 - 4) N protein can relieve the polarity of certain amber mutations.
- 18.2** Antitermination at t_{L1} of λ by N protein allows read-through transcription through *int*, which encodes the integrase enzyme. However, large amounts of the Int protein are not produced lytic infection, because these transcripts continue past the ρ -dependent terminator t_{int} . This allows the formation of a secondary structure in the RNA that serves as a signal for RNases to degrade the transcripts from the 3' end. Why are large amounts of Int made during lysogeny?
- 18.3** Sketch the RNA secondary structures in the *trp* leader/attenuator region being translated by a ribosome under conditions of low and high concentrations of tryptophan. What determines the progress of the ribosome, and how does this affect *trp* expression?
- 18.4** Which of the following events occur when *E. coli* is starved for the amino acid tryptophan?
- 1) No tryptophanyl-tRNA is made.
 - 2) The ribosome translates the leader peptide completely (to the UGA stop codon).
 - 3) A G+C rich stem-loop structure forms in the nascent RNA (regions 3 and 4) at the attenuator site.
 - 4) A stem-loop structure forms in the nascent RNA (regions 2 and 3) that precludes formation of the G+C rich stem-loop at the attenuator site.
 - 5) Transcription reads through the attenuator into *trp* *EDCBA*.
- 18.5** (POB) Transcription attenuation.
In the leader region of the *trp* mRNA, what would be the effect of:
- a) Increasing the distance (number of bases) between the leader peptide gene and sequence 2?
 - b) Increasing the distance between sequences 2 and 3?
 - c) Removing sequence 4?