CHAPTER 18 LECTURE NOTES: CONTROL OF GENE EXPRESSION
PART A: CONTROL IN PROKARYOTES

I.   Introduction
A. Up to now we have investigated HOW genetic information is inherited, WHAT genetic information is composed of, and HOW genetic information is expressed. This chapter examines HOW the expression of genetic information is REGULATED.
B. In higher organisms, gene regulation is frequently tissue specific and/or developmentally specific. In lower organisms, gene regulation is frequently in response to environmental stimuli.
C. We will examine several regulatory systems in *E. coli* as paradigms for regulatory systems in general.
        1. Cell must be able to turn a gene on/off
        2. Cell must "recognize" when to do (1)
        3. Other interesting aspects
                a) Use of mutants to elucidate biological processes
                b) Cleverness of bacteria
                c) Context of discovery

II.   The lactose operon (an inducible operon under both negative and positive control; regulation at the level of transcriptional initiation).
Francois Jacob and Jacques Monod 1950s

        A. Introduction
                1. *E. coli* can use lactose (glc + gal) as a carbon source (although glucose is preferred).
                2. The enzyme β-galactosidase (β-gal for short) cleaves lactose into glc and gal and is encoded by the *lacZ* gene.
                3. The enzyme lactose permease transports lactose into the cell and is encoded by the *lacY* gene.
                4. The genes are transcribed from a polycistronic mRNA along with the *lacA* gene.
                5. There is a promoter region that controls *lac* expression.
                6. There is an <u>operator</u> region to which a repressor binds (see below).
                7. The entire unit (POZYA) is called an <u>operon</u> (a set of adjacent genes transcribed as one polycistronic mRNA plus the adjacent regulatory signals)
        B. Negative regulation
                1. Original observation was that β-gal was not synthesized unless lactose (or a non-metabolizable analogue of lactose called <u>IPTG</u>) was in the growth media. This also applied to permease synthesis. In otherwords, β-gal synthesis was INDUCED in response to the presence of lactose or IPTG. What was the mechanism behind this regulation?

2. The discovery of regulatory mutants in *E. coli* that made β-gal <u>constitutively</u> (all the time - independent of the presence or absence of IPTG).

    a) One set of mutations mapped close to *lacZY* but not in them. Gene was designated *lacI* (for controlling inducibility). Complementation studies using <u>partial diploids</u> (chromosomal copy of the genes plus a copy on the F plasmid) were done:

Genotype	β-gal synthesis Non-induced	Induced	Conclusion
$I^+ Z^+ Y^+$	-	+	β-gal expression is inducible
$I^- Z^+ Y^+$	+	+	**LacI is required for repression of b-gal synthesis**
$I^+ Z^- Y^+ \ / \ F\,I^- Z^+ Y^+$	-	+	**LacI can complement in trans**
$I^- Z^- Y^+ \ / \ F\,I^- Z^+ Y^+$	+	+	Control
$\Delta IZY / I^- Z^+ Y^+$	+	+	Control

These data suggested that LacI was a trans acting repressor of *lacZ* transcription. This was later confirmed by examining the *lacZ* mRNA levels and by in vitro transcription reactions with purified components.

    b) The second set of mutations mapped directly before *lacZ*. The region was designated *lacO* for operator region (not a gene). These mutants had the same phenotype as the *lacI* mutants: expression of β-gal was constitutive. Complementation studies:

| Genotype	β-gal synthesis Non-induced	Induced	Permease synthesis Non-induced	Induced	Conclusion
$O^+ Z^+ Y^+$	-	+	-	+	control
$O^c Z^+ Y^+$	+	+	+	+	**Operator is required for repression**
$O^+ Z^+ Y^- \ / \ F\,O^c Z^+ Y^+$	+	+	+	+	**Operator region functions in cis**
$O^+ Z^+ Y^+ \ / \ F\,O^c Z^- Y^+$	-	+	+	+	**Operator region functions in cis**
$O^+ Z^- Y^+ \ / \ F\,O^c Z^+ Y^-$	+	+	-	+	**Operator region functions in cis**

These data suggested that *lacO* was a cis acting site at which repression occurred. This was later confirmed with in vitro binding assays.

3. A different mutation in *lacI* was also identified that <u>prevented</u> the induction of β-gal expression by IPTG. The mutation was designated *lacI*^{S}. Complementation studies:

Genotype	β-gal synthesis		Conclusion
	Non-induced	Induced	
I$^+$ Z$^+$ Y$^+$	-	+	β-gal expression is inducible
I^S Z$^+$ Y$^+$	-	-	**Mutation eliminates ability of IPTG to induce**
I^S Z$^+$ Y$^+$ / F I$^+$	-	-	***lacI*^{S} mutation is dominant to *lacI***
I^S Z$^+$ Y$^+$ / F I$^-$	-	-	
I^S O$^+$ Z$^+$ Y$^+$ / F I$^+$ O^c Z$^+$ Y$^+$	+	+	Operator mutations are epistatic to *lacI*^{S}

These data suggested that the mutation in *lacI*^{S} destroyed the binding site for an inducer. Thus, even in the presence of IPTG or lactose, repression of *lacZ* transcription still occurs. Subsequent in vitro binding studies showed that IPTG bound LacI.

4. Model for negative regulation based on above data (and other data too)
   a) In the absence of lactose, LacI repressor binds to the *lacO* site and represses *lacZYA* transcription.
   b) In the presence of lactose, allolactose (an isomer of lactose) binds to LacI repressor causing a conformational change in the protein that prevents binding to *lacO*. Transcription of *lacZYA* genes can now occur. This change in the conformation in response to the binding of a small molecule is called <u>allostery</u>.
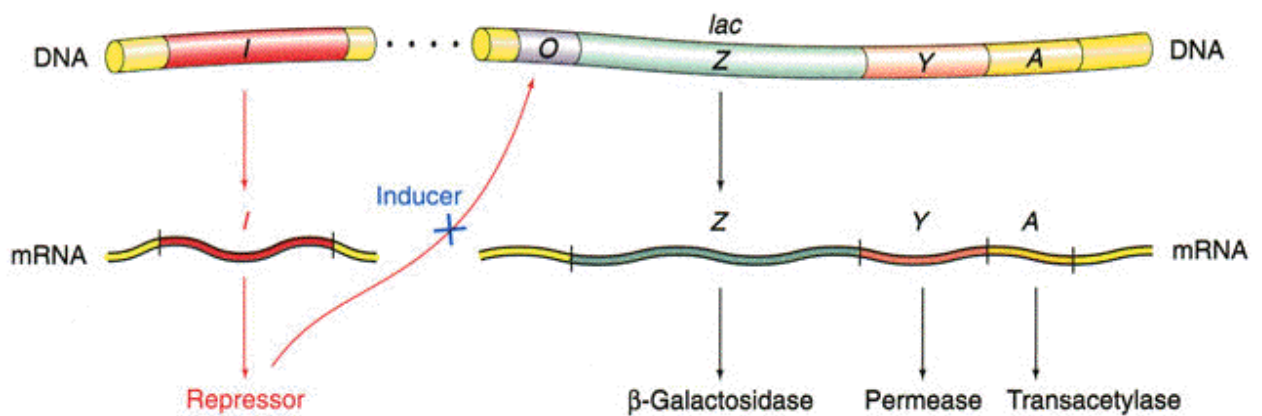
(From: AN INTRODUCTION TO GENETIC ANALYSIS 6/E BY Griffiths, Miller, Suzuki, Leontin, Gelbart © 1996 by W. H. Freeman and Company. Used with permission.)

C. Positive regulation
  1. Initial observation was that when glucose and lactose were both present in the media, there was no induction of β-gal expression until all the glucose was metabolized. This phenomenon was dubbed <u>catabolite repression</u> because it was presumed that a catabolic breakdown product of glucose repressed lactose utilization.
  2. Catabolite repression is modulated by cAMP levels in the following way:
  ↑ glucose = ↓ cAMP = ↓ β-gal expression = catabolite repression

  ↓ glucose = ↑ cAMP = ↑ β-gal expression

  3. Mutants were isolated that resulted in decreased β-gal expression. They mapped to:
      a) Adenylate cyclase (*cya*) gene – makes cAMP
   b) *crp* gene (encoding CAP or CRP)
  4. Model is that the cAMP-CRP complex binds to the lac promoter and aids RNA polymerase in activation of transcription.
          (1) When ↑ glucose: ↓ cAMP = ↓ cAMP-CAP = no binding to the promoter = no activation of β-gal expression. This makes sense because glucose (not lactose) is the preferred carbon source.
   (2) When ↓ glucose: ↑ cAMP = ↑ cAMP-CAP = binding to the promoter = activation of β-gal expression.
D. Overall model for regulation of the *lac* operon
  1. based on genetic analysis described above
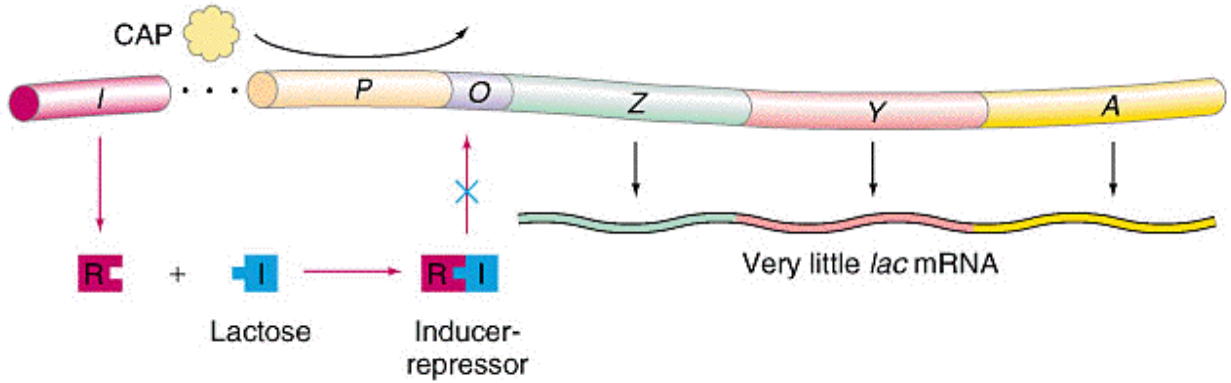 2. based on in vitro binding and transcription studies
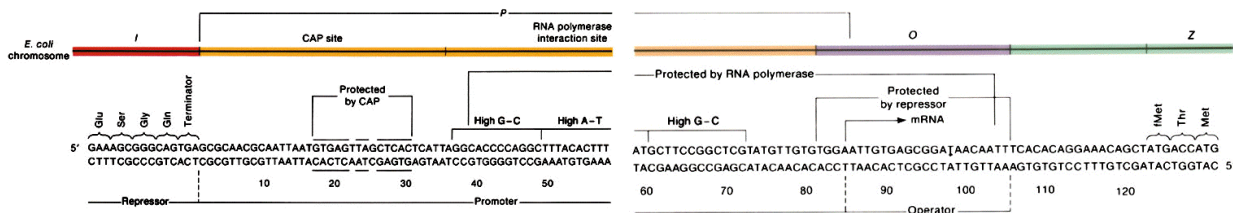
(From: AN INTRODUCTION TO GENETIC ANALYSIS 6/E BY Griffiths, Miller, Suzuki, Leontin, Gelbart © 1996 by W. H. Freeman and Company. Used with permission.)
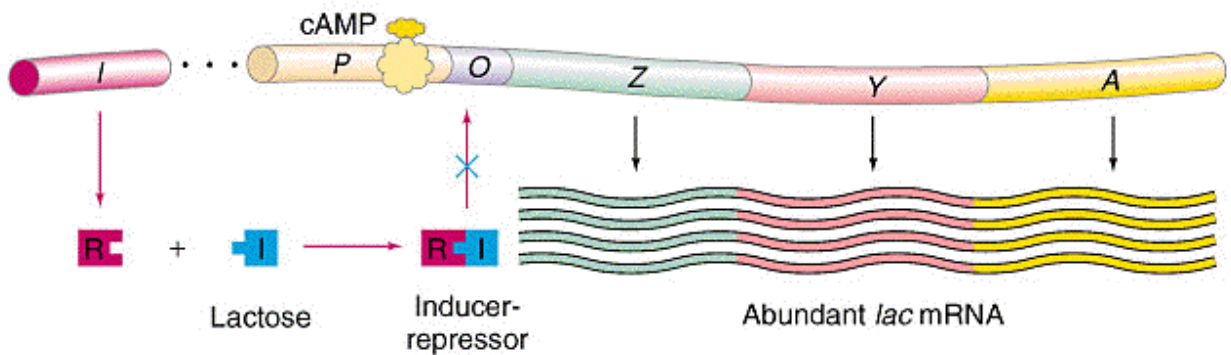
(a) Glucose present (cAMP); no lactose; no *lac* mRNA

CAP

I ... P O Z Y A

Repressor
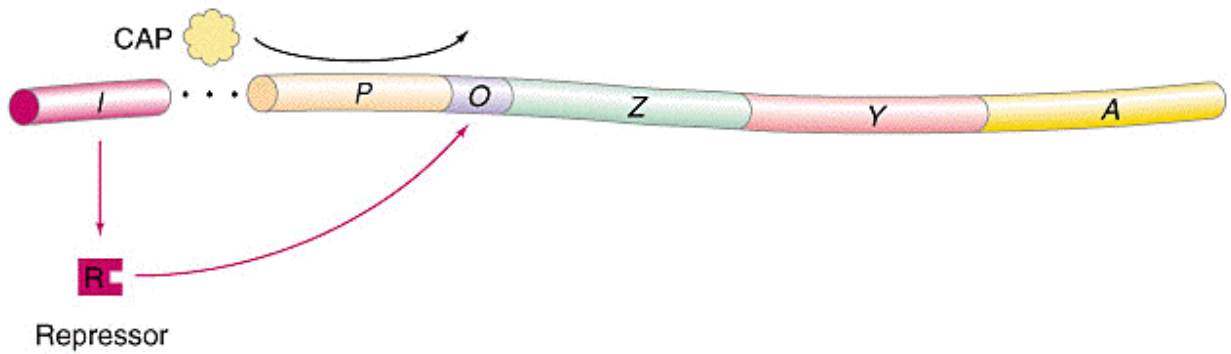
(b) Glucose present (cAMP low); lactose present

CAP

I ... P O Z Y A

Very little *lac* mRNA

R + I → R I

Lactose   Inducer-
           repressor

(c) No glucose (cAMP high); lactose present

cAMP

I ... P O Z Y A

Abundant *lac* mRNA

R + I → R I

Lactose   Inducer-
           repressor

E. Reminder of the differences between positive and negative control
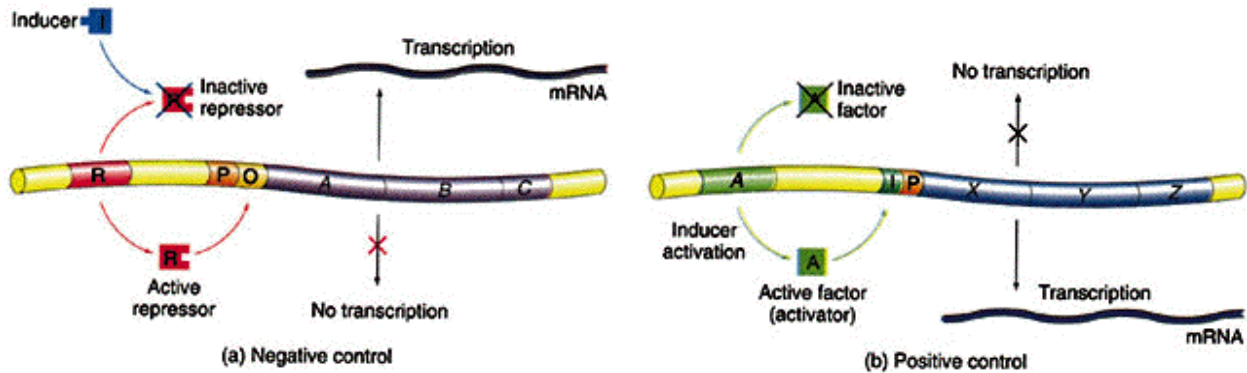
(From: AN INTRODUCTION TO GENETIC ANALYSIS 6/E BY Griffiths, Miller, Suzuki, Leontin, Gelbart © 1996 by W. H. Freeman and Company. Used with permission.)

III.   The arabinose operon (an inducible operon under positive and negative control via ONE protein; regulation at the level of transcription initiation)

    A. The arabinose operon encodes the *araBAD* gene which are required for utilization of the sugar arabinose.

    B. AraC can act as a repressor or activator of *araBAD* transcription.

        1. In the presence of arabinose, AraC forms a complex with arabinose which alters the conformation of the protein so that it binds the *araI* site which stimulates *araBAD* transcription.

        2. In the absence of arabinose, AraC binds to both the *araO* and *araI* regions and represses transcription via the formation of a repression loop.

        3. Allostery again

    C. Activation of the arabinose operon also requires cAMP-CAP as a positive activator.
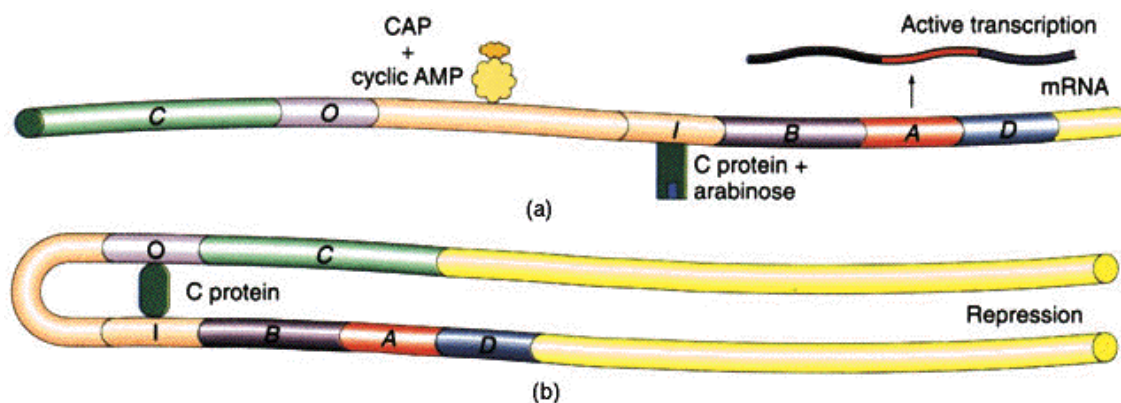
(From: AN INTRODUCTION TO GENETIC ANALYSIS 6/E BY Griffiths, Miller, Suzuki, Leontin, Gelbart © 1996 by W. H. Freeman and Company. Used with permission.)

IV.   The tryptophan operon (a repressible operon; regulation at the level of transcriptional initiation, feedback inhibition, and premature transcriptional termination)

     A. The *trp* operon encodes genes that are required for the synthesis of tryptophan (Trp) when it is not available in the growth medium.

     B. Physiological observation was that excess Trp in the media represses *trp* expression. This makes sense because there is no need to make Trp if it is in the media.

     C. Levels of regulation

          1. Transcriptional initiation: Trp repressor protein (TrpR) binds to an operator in the *trp* operon and represses transcription. This is a similar situation to the *lac* operon except that the <u>presence</u> of Trp in the media results in the binding of Trp to the repressor which <u>activates</u> the repressor so that it can bind the operator and repress transcription. In the case of the Lac repressor, the binding of lactose to the repressor <u>inactivates</u> the repressor so that it can NOT bind the operator.   The evidence that TrpR was a repressor was similar to that for the Lac repressor, namely that deletion mutations *trpR* and in the operator region preceding the *trp* genes resulted in constitutive expression of the operon.

          2. <u>Feedback inhibition</u>: When too much Trp is made, Trp binds to the first enzyme in the enzymatic pathway, thereby inactivating it. This results in a temporary shutoff of trp biosynthesis until the levels of Trp have decreased sufficiently so that Trp no longer binds to the first enzyme. This is a way to temporarily "tweak" the system to be as efficient as possible.

          3. Transcriptional attenuation (see below)
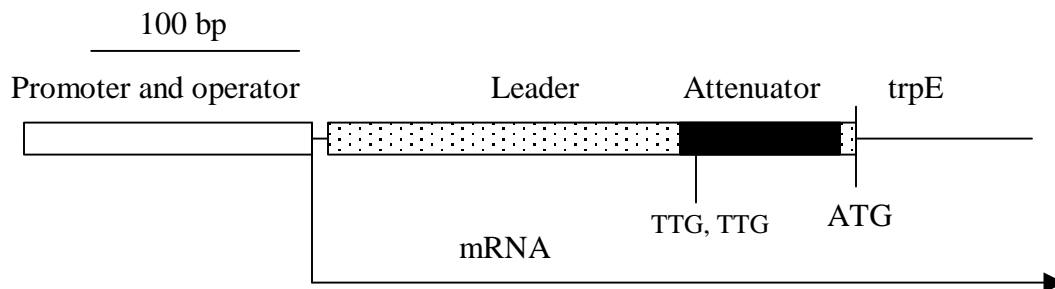
     D. <u>Transcriptional attenuation</u> (note most of these experiments were done in a *trpR* deletion mutant to eliminate any repression caused by TrpR)
Charles Yanofsky et al (1970s)

          1. Observed that in a *trpR* mutant, there was still a 10X increase in *trpE* mRNA when Trp was removed from the media. Why?

          2. Found a cis acting mutation that had a 10X increase in *trpE* mRNA <u>even</u> when Trp is present in the media. Mutation was a deletion in the leader sequence (see

picture above). The deleted element was designated an attenuator because its function appeared to be to reduce the amount of *trpE* mRNA when Trp was present.

3. Further experiments showed that while *trpE* mRNA was reduced 10X, initiation of transcription and transcription of the leader was not reduced. Thus, it appeared that the 10X reduction in transcription was due to termination of transcription somewhere in the leader.

4. Now the question became, what prevents termination of transcription in the absence of Trp (or what causes termination in the presence of Trp)? Yanofsky et al proposed a model based on sequence analysis and site directed mutagenesis in the attenuator region. The model was based on the secondary structure of the leader RNA. There are two MUTALLY EXCLUSIVE possibilities for the structure of the leader – only one can exist at a time: (A) Hairpins can form between regions 1 and 2   AND between regions 3 and 4 (termination hairpin) **OR** (B) A hairpin can form between regions 2 and 3 (antitermination hairpin)
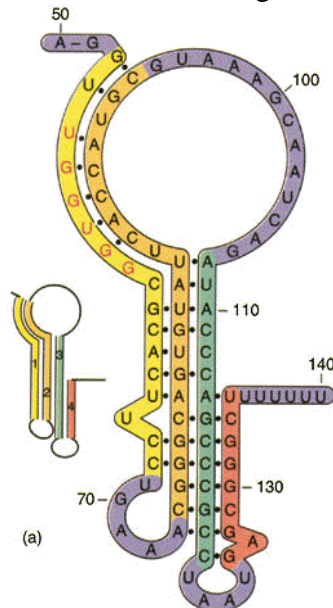
(From: AN INTRODUCTION TO GENETIC ANALYSIS 6/E BY Griffiths, Miller, Suzuki, Leontin, Gelbart © 1996 by W. H. Freeman and Company. Used with permission.)


Model for transcriptional attenuation:

Transcription of the leader region begins. As soon as regions 1 and 2 are transcribed they form a hairpin that temporarily pauses RNAP. The ribosome hops on the mRNA and begins translating the leader. The moving ribosome disrupts the hairpin and "unpauses" the paused RNAP so that transcription resumes. Now there is a precise coupling of transcription and translation. The next events depend on whether excess Trp is present:

If there is excess Trp in the cell (high Trp levels), the ribosome translates the leader peptide until it hits a stop codon located at the end of region 2 of the RNA. At this stop codon, the ribosome pauses thereby covering up region 2 RNA. Since region 2 RNA is covered up, it can not form the antitermination hairpin with region 3 RNA. This leaves region 3 RNA free to pair with region 4 RNA which has just been transcribed. The RNA hairpin formed by region 3 and 4 is a termination hairpin that destabilizes the DNA:RNA hybrid causing the mRNA to dissociate from the transcription complex. TERMINIATION OF TRP TRANSCRIPTION OCCURS.

If the cell is starved for Trp(low Trp levels), the ribosome pauses at the 2 Trp codons located in region 1 of the RNA because there is not enough Trp available. Meanwhile, transcription of the leader is continuing and now region 3 has been transcribed. Region 3 RNA can form the antitermination hairpin with region 2 RNA. While this RNA hairpin is formed the termination hairpin (3&4) can NOT form. Thus, transcription continues through region 4 and into the *trp* genes. NO TERMINATION OF TRP TRANSCRIPTION OCCURS.
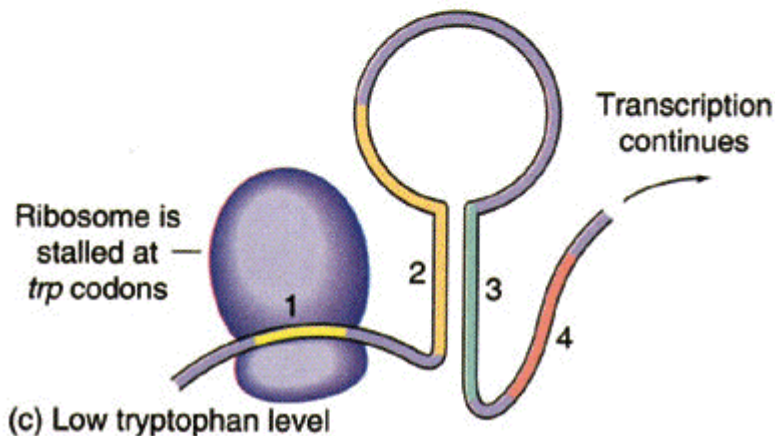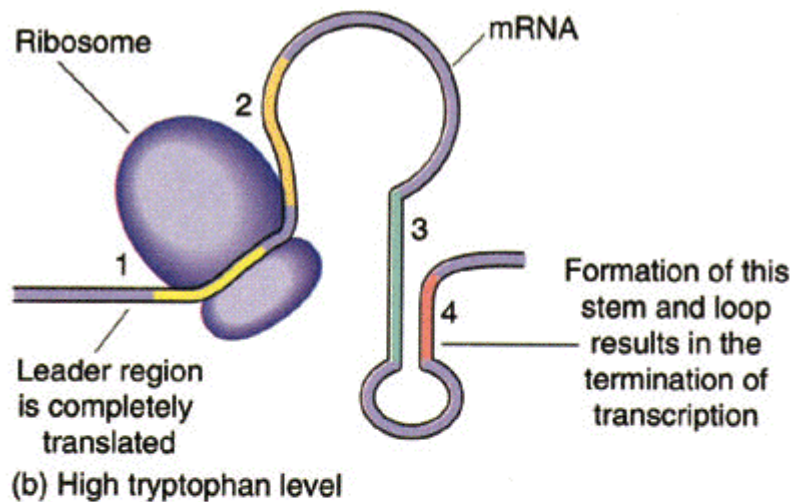
(From: AN INTRODUCTION TO GENETIC ANALYSIS 6/E BY Griffiths, Miller, Suzuki, Leontin, Gelbart © 1996 by W. H. Freeman and Company. Used with permission.)

     5.   Several biosynthetic operons have regulatory mechanisms similar to this one.

V.   Multioperon control: One way that genes that are all needed at the same time are coordinately expressed is to put them all in the same operon. Another way is by having 1 regulatory protein regulate multiple operons, all encoding genes needed for a certain function.

VI.   Two component regulatory systems
        A. Environmental stimulus affects one domain of the sensor protein which causes an alteration in the structure of another domain in the sensor protein. The altered domain in the sensor protein phosphorylates a domain in the response regulator protein. The phosphorylation of a domain in the response regulator alters the response regulatory protein conformation so that it can now activate transcription of a gene.
        B. Processes regulated by 2-component regulatory systems include nitrogen assimilation, chemotaxis, sporulation, flagella, phosphate utilization, pathogenesis, cellular communication

PO_4

Output =
activation of transcription

Sensor protein

Response regulator
protein

STIMULUS