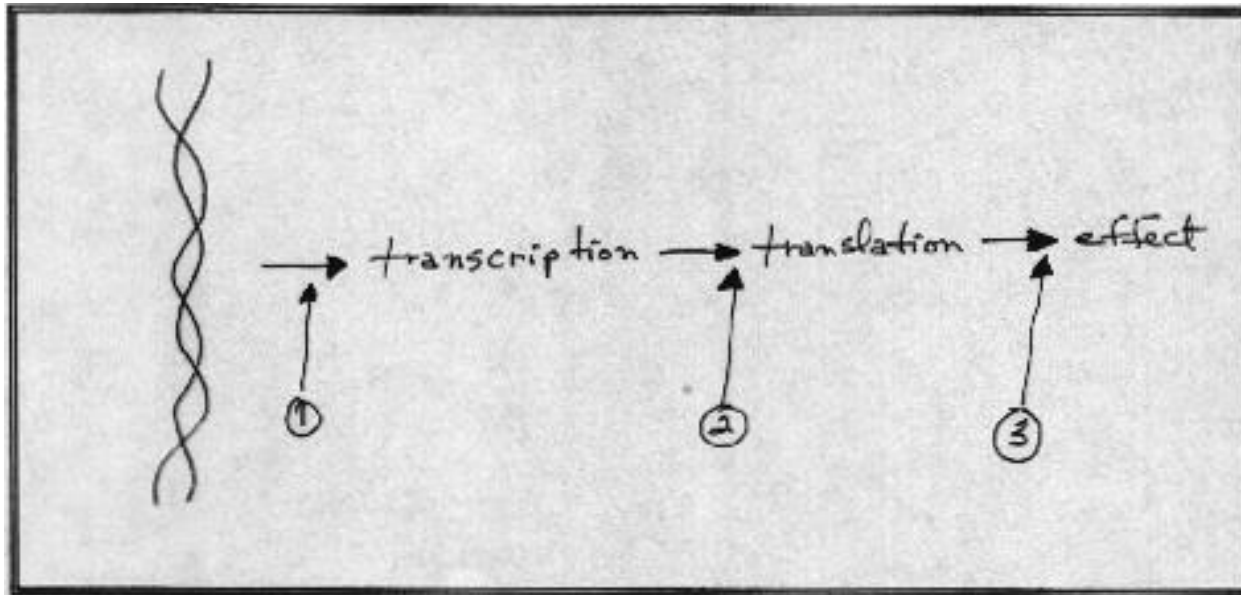


# Gene Switches



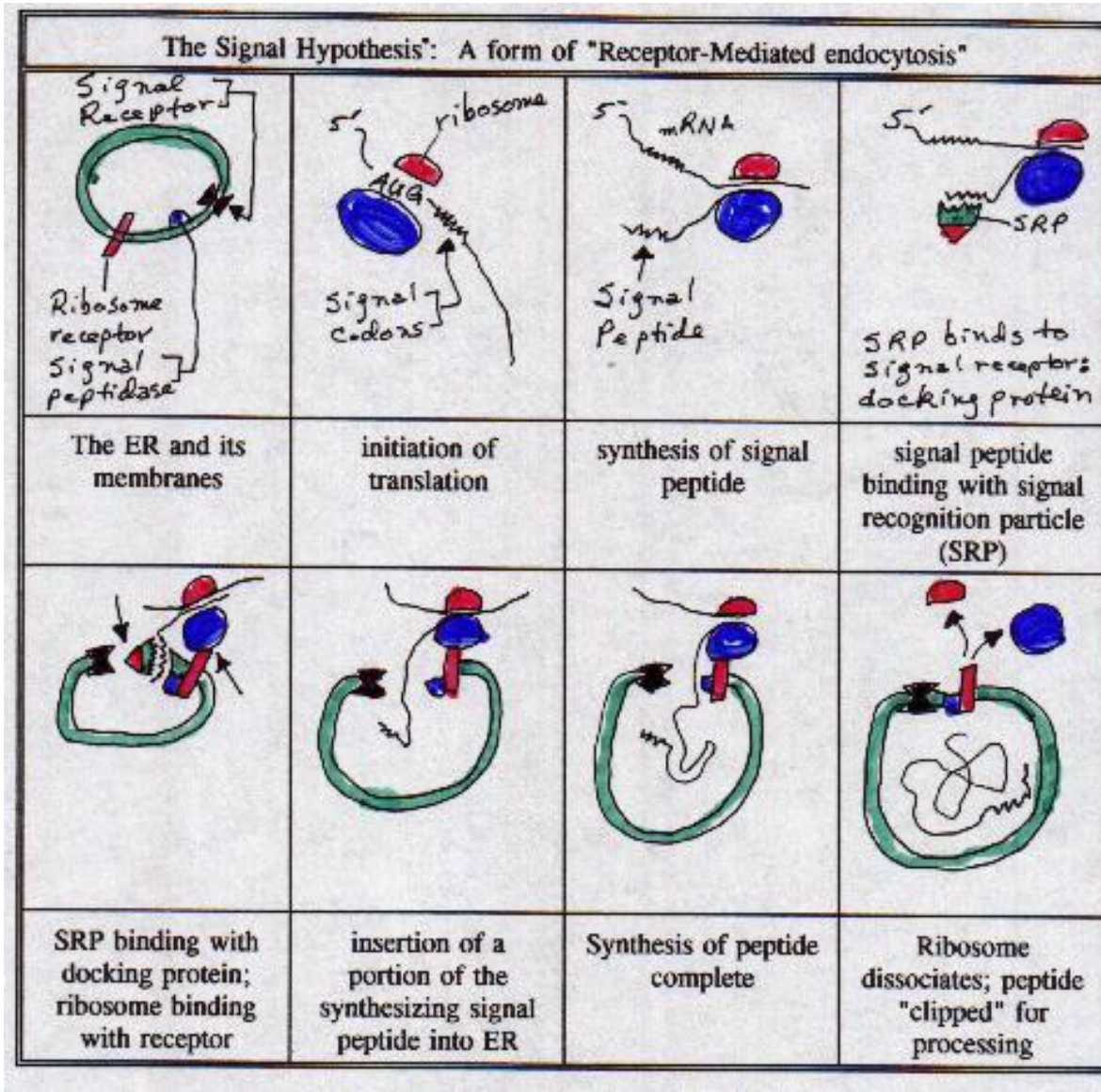
- When regulating genetic activities, there are 3 big regulatory mechanisms (Figure, above):
  - 1) inhibit transcription,
  - 2) inhibit translation and/or
  - 3) inhibit protein activity.

- The most **energy saving** mechanism is to inhibit transcription. The end result is that the information of the gene remains unexpressed.
- 
- Translation inhibition may be due to destroyed mRNA, may be due to inactivated mRNA molecules or may be due to blocking ribosomes from binding with/to mRNA. The end result is that the protein won't be synthesized.

# The Signal Hypothesis

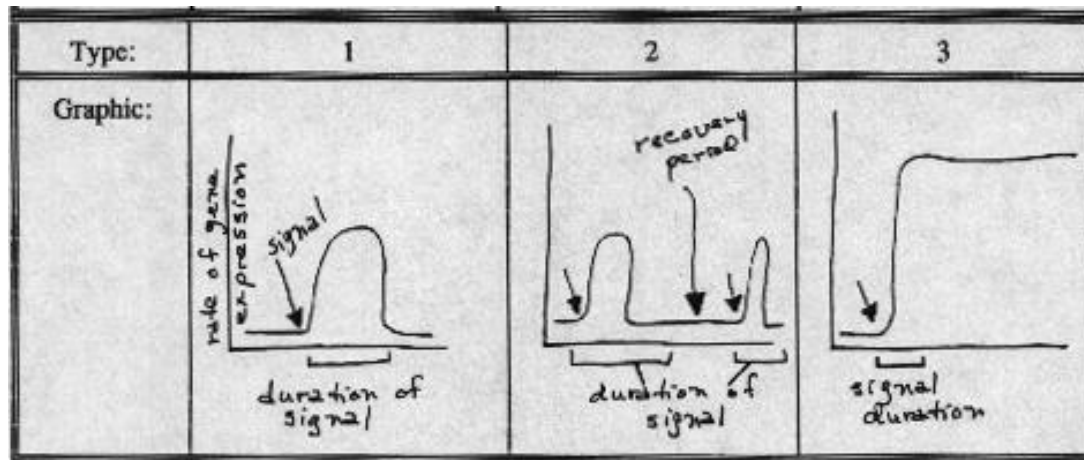
- To inhibit protein activity requires intracellular inhibitors and may require post-translational modifications that inactivates or reduces the activity of the protein. To accomplish this task, one mechanism that has received a great deal of attention is to stop transport of the protein to the cell site.

# The Signal Hypothesis



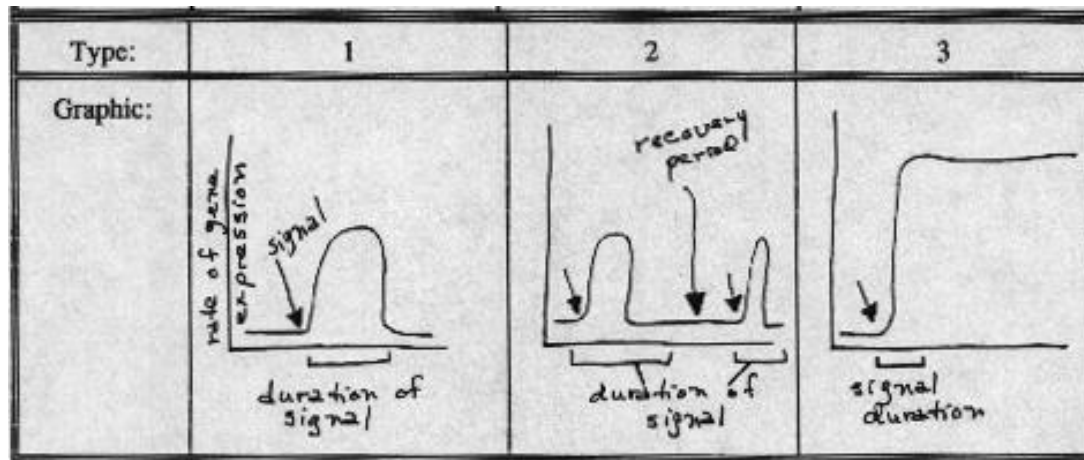
- Another mechanism whereby the protein's activity is inhibited is through negative feedback, i.e., as the concentration of substrate ( $[S]$ ) increases it inhibits enzymes further up the biochemical pathway.
- In negative feedback, a substrate builds up in sufficient quantities to cause one or more enzymes to be allosterically inhibited (review enzyme inhibition), temporarily halting enzymatic activity.

# Type 1



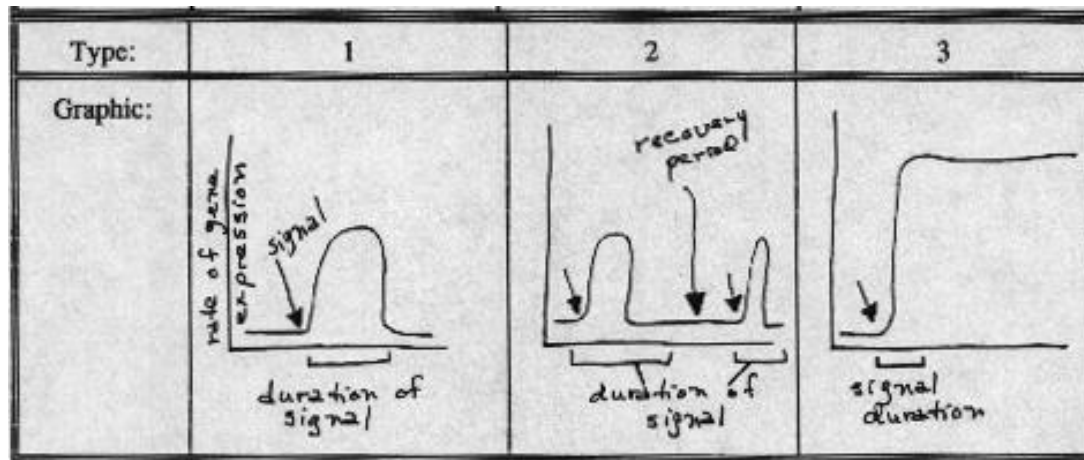
- Type 1 expression shows that the rate of expression is proportional to the presence of the inducer and lasts only as long as there is inducer present causing a signal to drive expression. Aka Type A, e.g., Steroid hormones

# Type 2



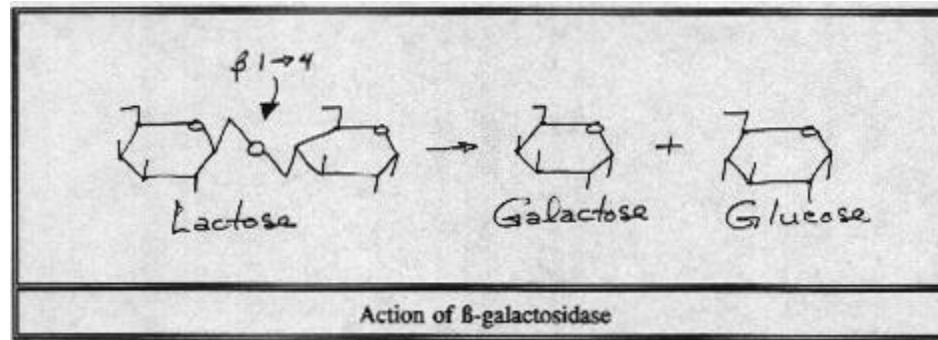
- Type 2 expression involves multiphasic expressions separated by recovery periods even if the inducer is present. This is also called "response-desensitization-recovery" as it involves a large response to a signal, followed by a recovery period, then a smaller response occurs. Aka Type B, e.g., Pharmacologic agents

# Type 3



- Type 3 expression is an expressive activity that gets a short signal, but the gene[s] continue[s] expressing even after the inducer is no longer providing the signal to drive expression. Aka Type C, e.g., Cancer

# Definitions

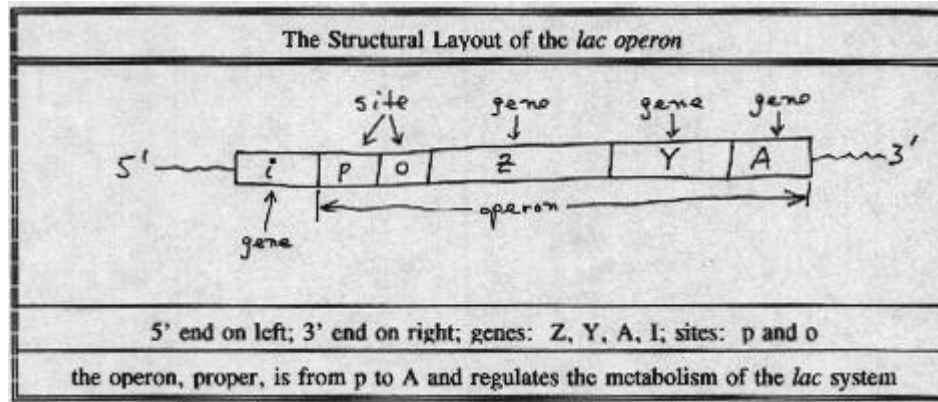


- An **operon** is a piece of DNA containing 2 or more genes next to each other and which respond to joint regulation at specific regulatory sites near to genes that direct polypeptide synthesis (structural genes).
- The **lac operon** is a region of DNA that codes for the hydrolysis of lactose to glucose and galactose ( $\beta$ -galactosidase is the enzyme that does this in *E. coli*, Figure, above).
  - The lac operon is the best studied "switch" for genetics and has been studied extensively in *E. coli*.
  - It was first elucidated by J. Monod and F. Jacob in Paris in 1961.
- The phenomenon of gene switches has been best studied in prokaryotes, but the ideas seem to, more or less, carry over into eukaryotes, as well.

- **Positive regulation** causes an increase of genetic expression due to a specific regulator.
- **Negative regulation** causes a reduction of genetic expression due to a specific regulator.
- **Double negative regulation** occurs when an inhibitor of a negative regulator binds to the negative regulator. This is roughly positive regulation. It is **also known as *derepression***.

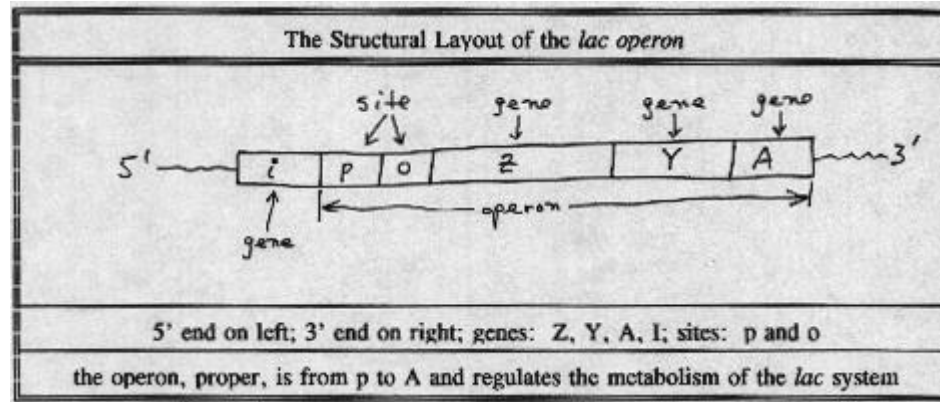
- A **cistron** is the smallest unit of genetic expression; it codes for the sub-unit structure of a protein. It seems, now, that "one gene, one enzyme" is NOT so good any more as a way to explain protein synthesis. "One cistron, one sub-unit" is probably more appropriate.
- An **inducible gene** is a gene that is induced, i.e., whose expression is increased.
- **Constitutive expression** occurs when genes are expressed at a reasonably constant rate; this is not known to be subject to regulation.
- **Constitutive mutation** occurs when a regulated gene that has mutated resulting in constitutive mutation causing constitutive expression of that gene.
- A **site** (proper) is a section of DNA that is NOT transcribed into mRNA.
- **Polycistronic mRNA** is a large piece of mRNA that codes for 2 or more genes with independent AUG and UAA sites for EACH cistron. These are primarily found in prokaryotes.

# The lac operon



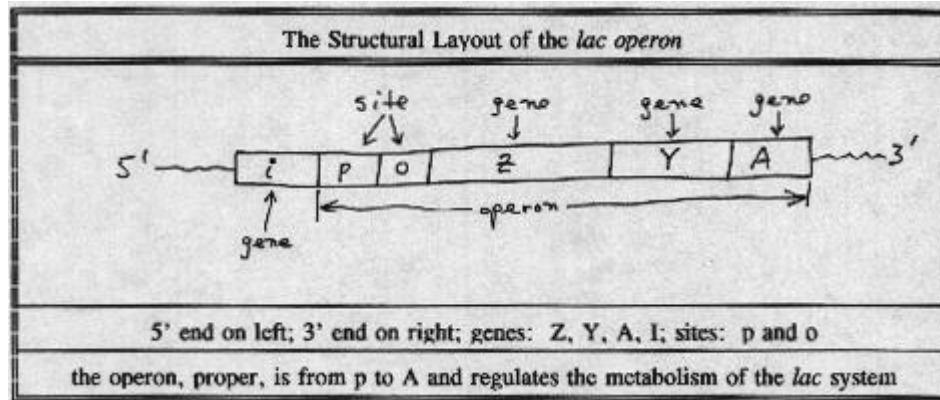
- The layout of the lac operon. The 5' end is on the left; the 3' end on the right.
- There are four genes: Z, Y, A, I (i).
- There are two sites: p and o.
- The operon, proper, is from p to A and regulates the metabolism of the lac system.
- Specifically, the gene, lac Z, codes for  $\beta$ -galactosidase. This enzyme hydrolyzes lactose to glucose and galactose, as already discussed.
- lac Y codes for lactose permease. This enzyme transports and concentrates lactose in the cell.
- lac A codes for galactoside acetylase. It is not present in all lactose metabolizing prokaryotes. Its function is not presently known. It may, perhaps, be used to synthesize an intermediate in lactose catabolism (??????).
- All the above genes are mRNA coded and are polycistronic.

# The lac operon



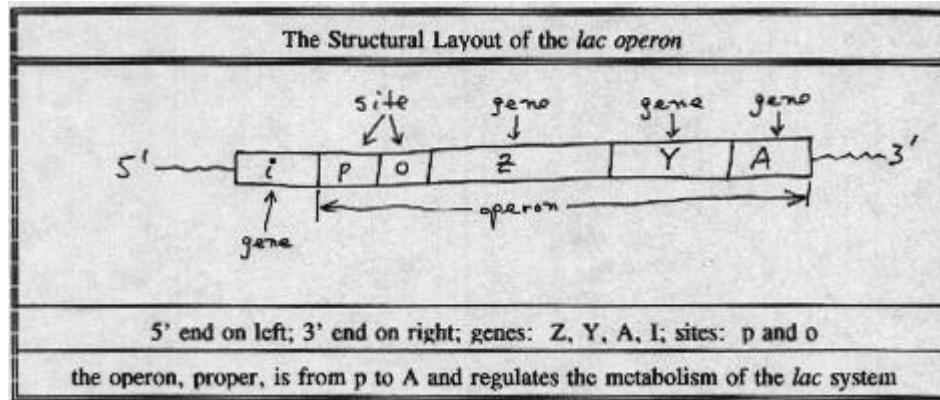
- The regulatory gene, *I* or *i*, is a regulatory element.
- It codes for the "off switch" under normal conditions.
- It codes, specifically, for repressor proteins.
- It is also known as (aka) *lac i*.
- It is monocistronic and a constitutive gene.
- It codes for the synthesis of the sub-units of the ***lac repressor*** (**NEGATIVE repression**) which are tetramers that seem to function as dimers by binding to the operator (the *o*) site.

# The lac operon



- There are 2 control sites for the lac operon: the "p" site and the "o" site.
- The p site is the promoter for RNA Pol binding (DNA-dependent).
- About 40-45 base pairs (bp's) are recognized by RNA Pol for transcription.
- When the o site is capped by "i" gene products, RNA Pol is not able to bind to the "p" site and transcription is blocked.

# The lac operon



- The "o" site is the operator site. It is adjacent to the structural genes that it regulates. It is 27 base pairs long, more or less. It has a two-fold axis of symmetry about 21 bp's. The two-fold axis of symmetry is shown below (DS DNA):

5'-AATTGTGAGC G GATACAATT-3'  
 3'-TTAACACTCG C CTATTGTTAA-5'

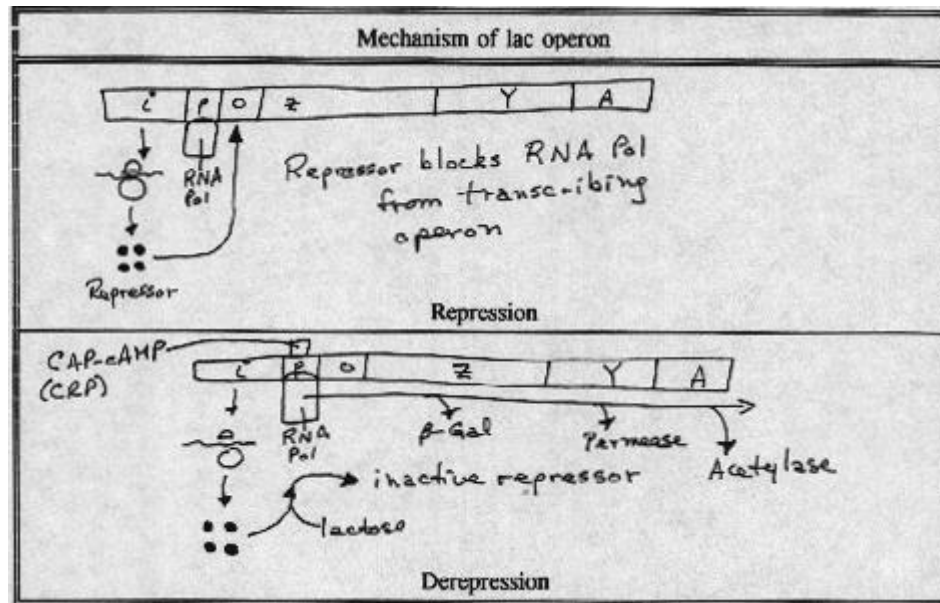
- One axis of symmetry is underlined; the second axis of symmetry is *italicized*.

- The lac repressor is recognized by 1 base pair in the 17 base pairs (illustrated, below; SS DNA shown) of the o site.

5'-TGTGTGGAATTGTGAGCGGATAACAATTCACACA-3'

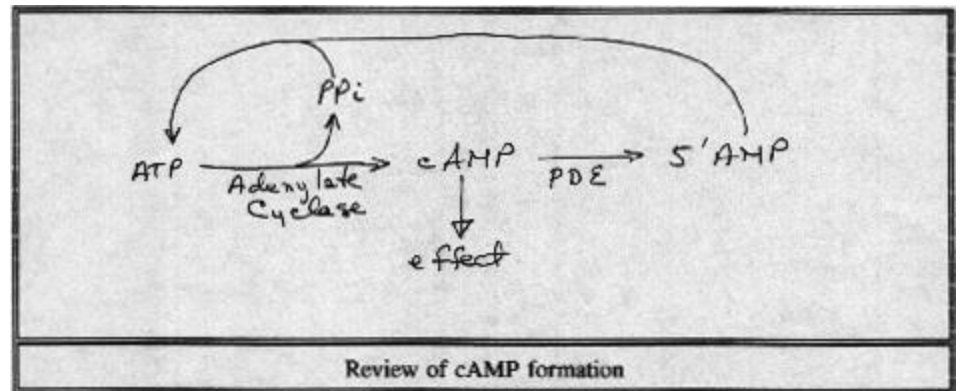
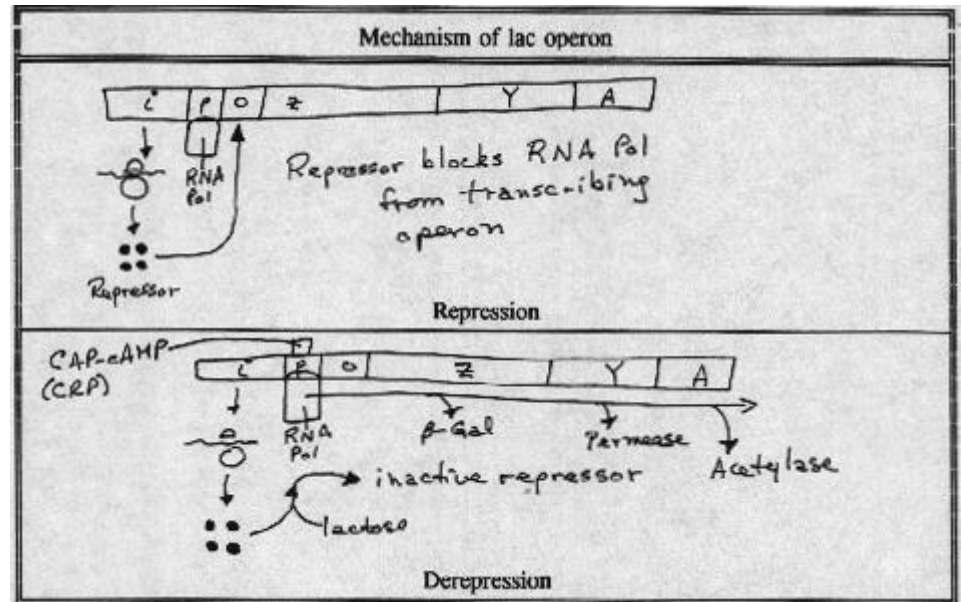
- When the repressor is bound to the operator, this causes steric hindrance for the RNA Pol and blocks initiation of transcription.
- The symmetry of the repressor is probably complimentary to the o site.
- The repressor seems to also protect the above-marked region of the o site from pancreatic enzymatic degradation (DNA'ase).
- There are 20-40 repressor tetramers and 1 operator site per *E. coli* cell.
- The repressor binds in the major groove and does not disturb the B-DNA structure.

- When lac genes are expressed, they give a Type 1 response on the order of 10-fold to 100-fold increase in activity.
- When the inducer is removed, lac Z, Y, A syntheses decrease markedly, i.e., the gene products drop off drastically.

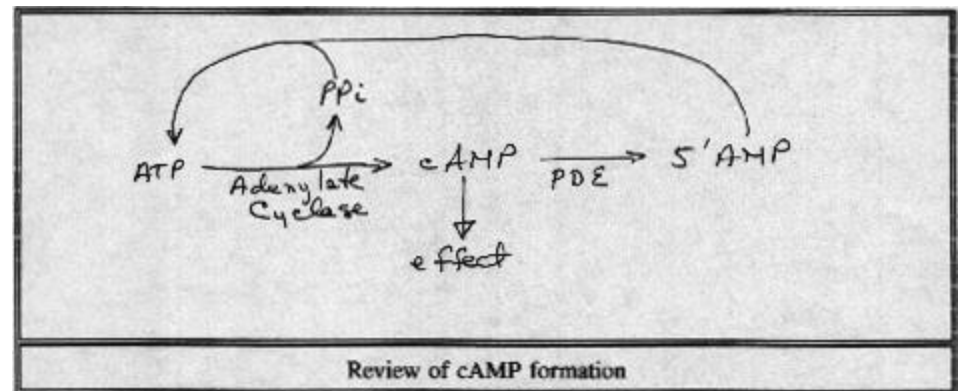
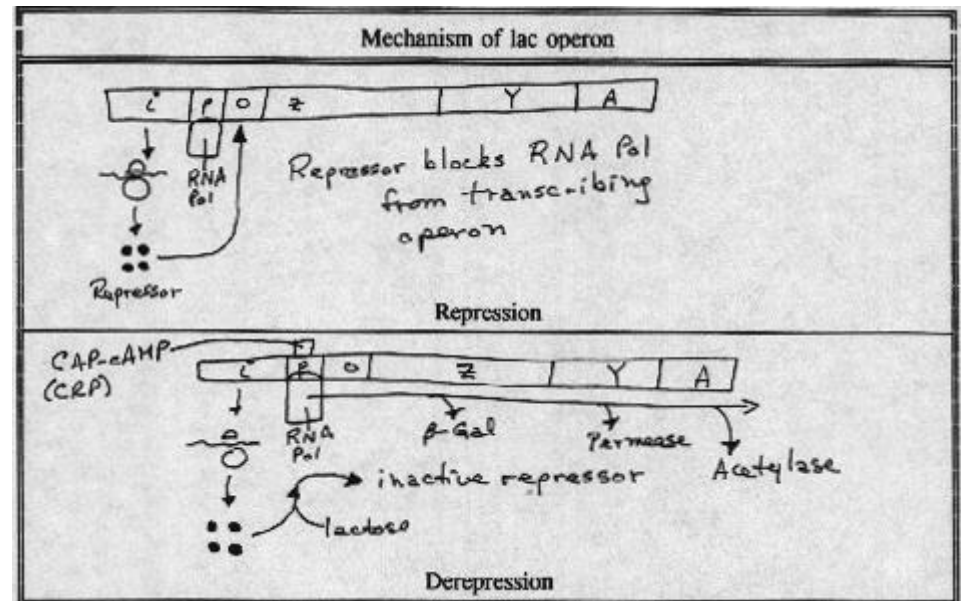


- The mechanisms of repression and derepression of the lac operon.
- The "i" gene is transcribed and the resulting mRNA is translated.
- The **repressor** results.
- The repressor binds to the "o" site and blocks the further movement of RNA Pol, halting operon transcription.

- Under conditions causing derepression, a catabolite gene activator protein (**CAP; aka cAMP receptor protein (CRP)**) is bound with cAMP (Synthesis shown/reviewed in Figure, bottom) to mediate "catabolite repression".
- "Catabolite repression" was thought to involve some glycolytic catabolite which did not permit *E. coli* to use lactose UNTIL ALL the glucose was utilized.
- We now know it is the **CAP-cAMP/CRP-cAMP complex** that regulates this "catabolite repression".



- During derepression, the **CAP-cAMP (CRP-cAMP) complex** binds to the promoter site, which allows RNA Pol to bind and continue with transcription, i.e., transcribing the operon so that active lactose metabolizing enzymes are synthesized - this protein-cAMP complex is a positive regulator.



- *E.coli*, like many other bacteria, metabolize carbohydrates like glucose.
- Unlike many other bacteria, *E. coli* metabolizes lactose.
- In general (as alluded to, above) this bacterium, when placed in a medium containing glucose and lactose, catabolizes glucose, then temporarily stops growing.
- After the lac operon is induced, *E. coli* begins to catabolize lactose and continues to grow until toxic catabolites build up in the medium, killing it off.

- In order to understand how the repressor worked, some simple experiments were set-up and observed.
- Geneticists first mutated the *i* gene to code for an inappropriate/ineffective repressor.
- Under the reaction conditions, there was constitutive expression of the operon because the repressor was unable to bind to the "o" site.
- Then the *i* gene was mutated, again, this time to make an inappropriate repressor that would bind to the o site, but that would not bind inducer.
- This resulted in continual repression.
- By placing the genetically mutated bacteria in media with/without glucose and/or lactose, researchers were able to show that the lac operon was not transcribed/translated until AFTER the glucose had been all used up.

- Researchers then coupled these sorts of experiments with experiments to increase or decrease cAMP levels in the presence of *E. coli*. These are summarized in the table, below:

Experiment 1	Experiment 2	Experiment 3	Experiment 4	Experiment 5	Experiment 6
No glucose	+ glucose	+ glycerol	No glucose	+ glucose	No glucose
No lactose	No lactose	No lactose; No glucose	+ lactose	+ lactose	+ lactose + glycerol
cAMP levels resulting in increased cAMP-CAP	↓ cAMP levels	cAMP levels resulting in increased cAMP-CAP	cAMP levels resulting in increased cAMP-CAP	↓ cAMP levels	cAMP levels resulting in increased cAMP-CAP
lac operon turned ON	lac operon turned OFF	lac operon turned ON	lac operon turned ON	lac operon turned OFF*	lac operon turned ON

- \*The results from experiment #5 (glucose and lactose shutting off the operon) suggest that a glucose-based compound is an adenylate cyclase inhibitor.
- Remember that adenylate cyclase is the enzyme that cyclizes ATP to cAMP and PPi

# Academic Aside:

- The repressor, itself, has an association constant ( $K_{\text{association}}$ ) of  $7 \times 10^9$  L/mol-sec (2d order rate constant)

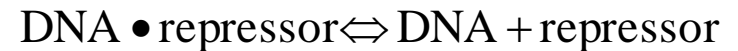
Dissociation (separation) :

**This is a very high value that indicates that association of the DNA with the repressor is highly favored.**

**It also indicates that the repressor slides along the DNA to the "o" site of the operon (a one-dimensional search) rather than diffusing through the aqueous medium for a chance encounter (a three-dimensional search)**

**HENCE, this process is HIGHLY organized!**

**End of Aside**

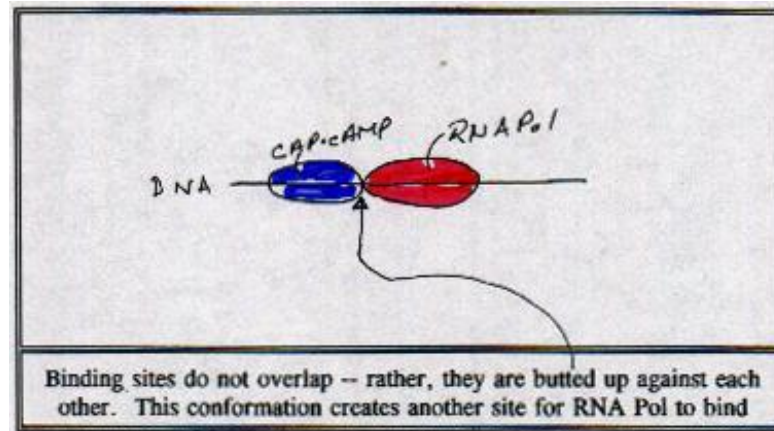


$$K_d = \frac{[\text{DNA}] [\text{repressor}]}{[\text{DNA} \bullet \text{repressor}]}$$

$$\text{remember: } K_a = \frac{1}{K_d}$$

$\therefore$

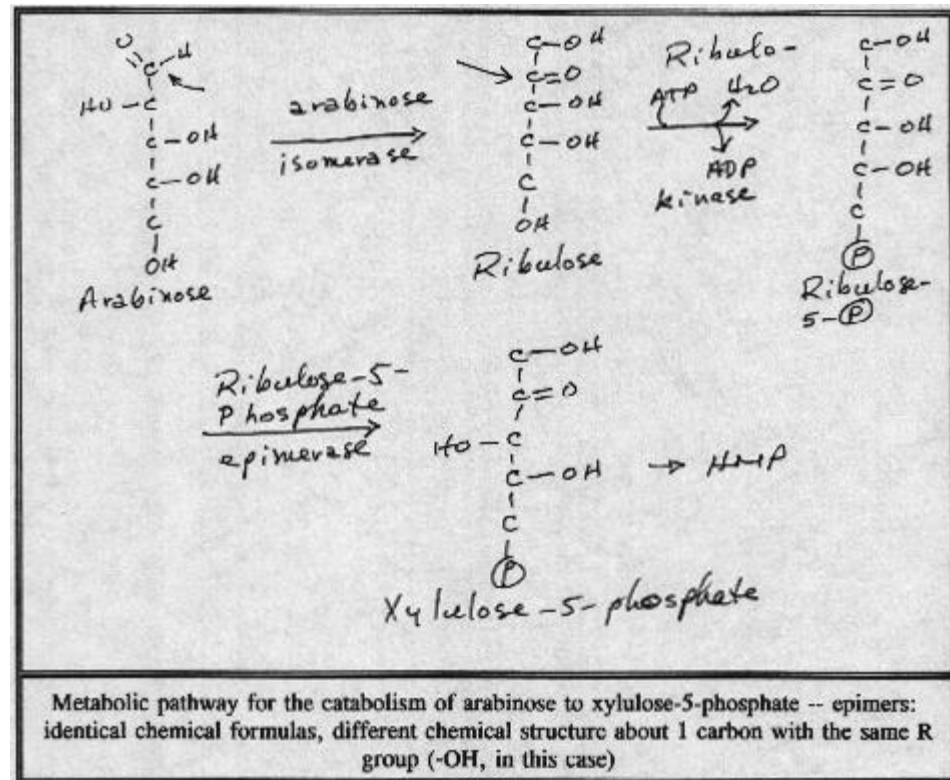
$$\frac{1}{K_d} = \frac{[\text{DNA} \bullet \text{repressor}]}{[\text{DNA}] [\text{repressor}]} = 7 \times 10^9 \text{ L/mol-sec}$$

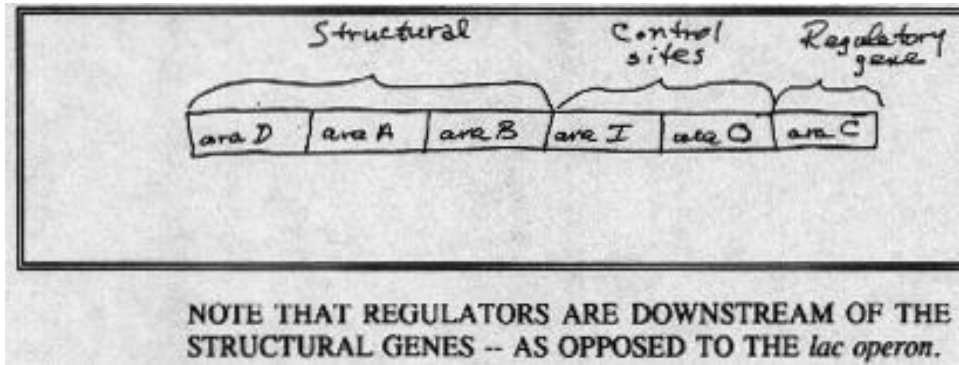


- Additional information is known about CRP: it functions as a dimer with RNA Pol.
- It binds upstream of RNA Pol on -87 to -49 (nucleotides -87 to -49, i.e., on a sequence of nucleotides that are 49-87 nucleotides behind the RNA Pol binding site).
- RNA Pol binds/protects -48 to +5 (it covers the first triplet to be transcribed and two extra for good measure), Figure, above.
- CAP increases mRNA initiation by more than 50-fold.
- How does it do this? The binding sites do not overlap; rather they are butted up against each other.
- This conformation creates another site for RNA Pol to bind.
- **Alternatively, the repressor binds from -3 to +21, which overlaps the RNA Pol site (-48 to +5), so the repressor functions by literally blocking RNA Pol from binding.**

# The Ara Operon: A Slightly Different Approach

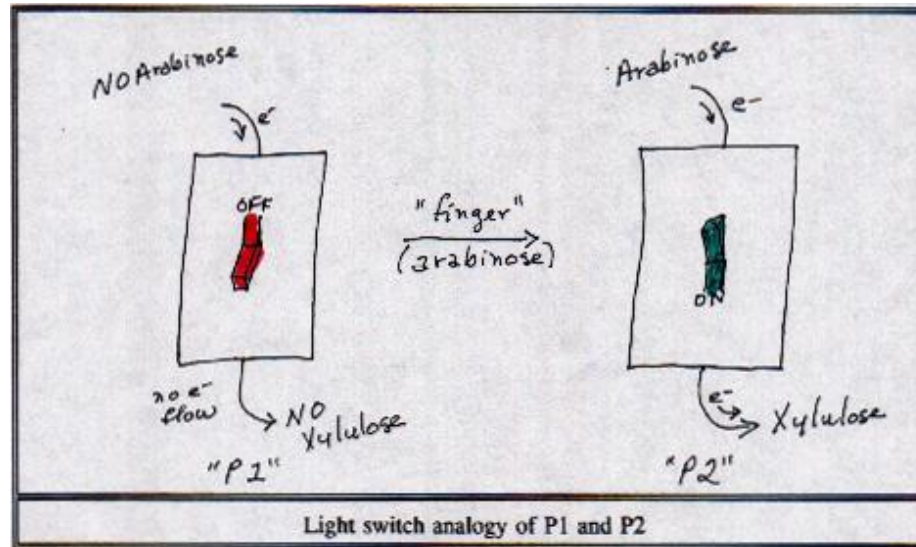
- The **ara operon** codes for the enzymes necessary to catabolize arabinose to xylulose-5-phosphate in microorganisms, Figure, right.
- In brief, arabinose is isomerized to ribulose which is phosphorylated to ribulose-5-phosphate.
- Ribulose-5-phosphate is epimerized to xylulose-5-phosphate (an epimer is a compound that has an identical chemical formula that differs by the orientation of identical R-groups on the identical carbon, e.g., -OH on carbon 3 in ribulose-5-phosphate and xylulose-5-phosphate).





- The layout of the ara operon.
- Note that regulators are downstream of the structural genes as opposed to the location of the regulators in the lac operon.
- Ara A, B and D code for arabinose isomerase, ribulokinase and ribulose-5-phosphate epimerase, respectively.
- The product of ara c is a protein that has two different conformations: P1 and P2.
- - P1 is present in the absence of arabinose. It is a repressor. It binds to the operator and blocks the ara operon from "working".
  - P2 is present in the presence of arabinose (it is P1 that has undergone a conformational change, i.e., a shape change). It is an activator that binds with CAP and cAMP. It bonds to the promoter and activates the ara operon.

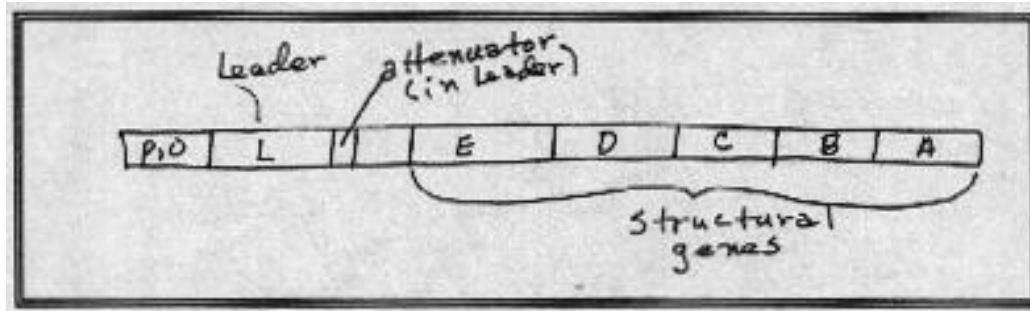
# The Light Switch Analogy for P1 and P2



- In short, arabinose acts like electrons flowing to a light switch.
- If there is **no arabinose**, there will be no xylulose and the light switch (P1) stays closed (or "off").
- If there **is arabinose**, then the light switch is turned on (P2) and xylulose is formed.
- Hence, the ara operon is positively and negatively regulated by the SAME protein with a DIFFERENT conformation for each set of conditions. The ara operon was the first operon to be shown to be both positively and negatively regulated.

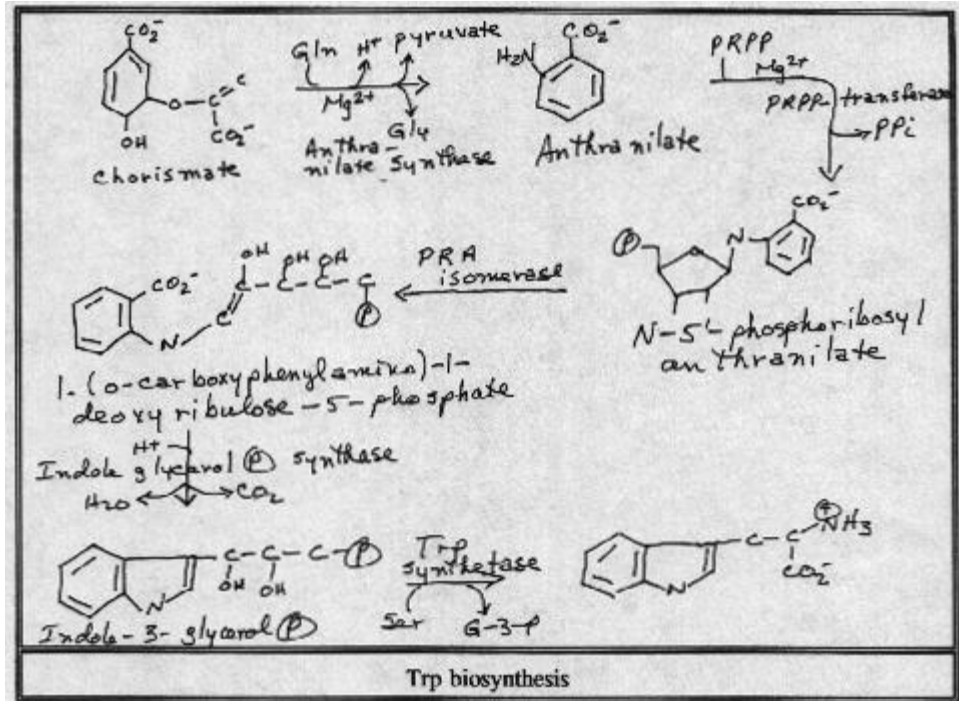
# The trp Operon: An Even Different Approach

- The key to the trp operon is to remember that when the concentration of trp is high ( $\uparrow[\text{trp}]$ ), the synthesis in trp mRNA is stopped, i.e., this is an operon necessary for the synthesis of trp.
- Secondly when  $\downarrow[\text{trp}]$  are present, trp mRNA synthesis is turned on.



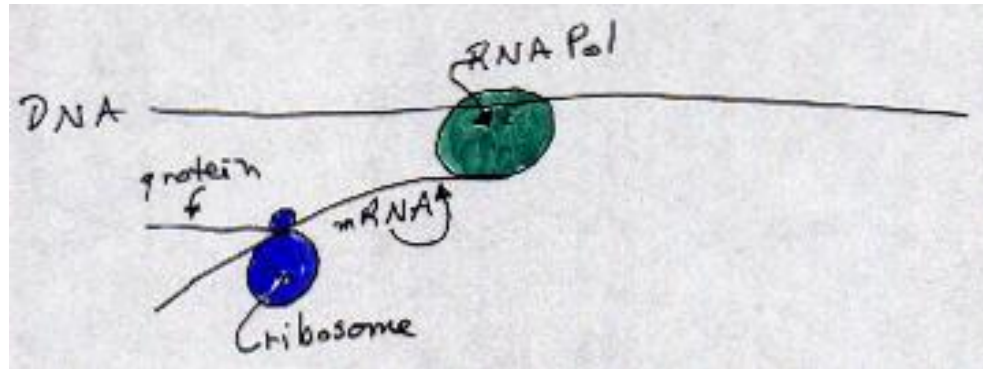
- The layout of the trp operon.
- Note that there is a leader sequence upstream of the structural genes and that there is an attenuator sequence in the leader sequence.
- An attenuator reduces, weakens, diminishes the output or effect or product synthesis.
- The trp operon codes for the 5 enzymes necessary for trp synthesis.

- The biosynthetic pathway for trp.
- In brief, the synthesis begins with chorismate (the ionized form of chorismic acid) as it is reacted with gln in the presence of  $Mg^{2+}$  with anthranilate synthase.
- Anthranilate is then reacted with phosphoribosylpyrophosphate to form N-5'-phosphoribosylanthranilate (NPRA).
- NPRA is isomerized to its open chain derivative (1-(o-carboxyphenylamino)-1-deoxyribulose-5-phosphate).
- Following the removal of water and carbon dioxide and the addition of ser, trp is synthesized.

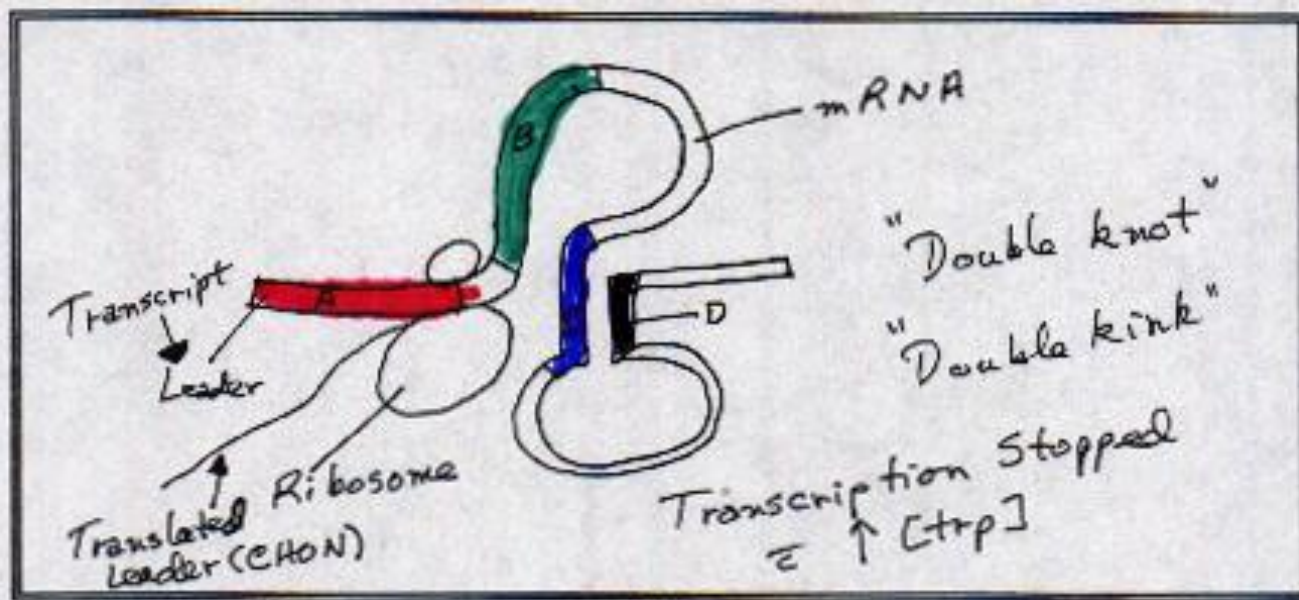


- Translation of the trp operon literally follows behind transcription of the trp mRNA.
- This takes about 4 minutes to synthesize trp mRNA and it takes about 3 minutes to degrade trp mRNA.
- The trp R gene codes for a repressor protein.
- This gene is a great distance removed from the trp operon.
- The repressor AND trp (in high concentrations) combine for a tightly bound complex that binds to the operator.
- Trp is required for this activity and is a corepressor.
- 
- Since the "p" and "o" sites overlap, when the trp-repressor complex binds, RNA Pol can not bind.

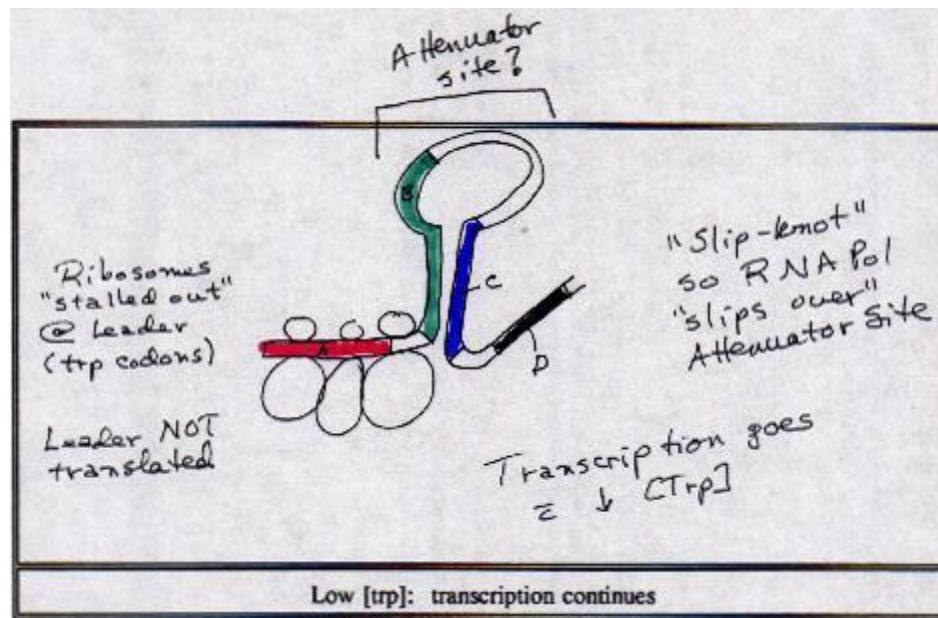
- The trp operon also has a "new twist": a leader sequence.
- This leader sequence is intimately involved in regulating the trp operon as we shall see later.
- The leader sequence butts up against an attenuator sequence.
- The attenuator sequence/site works with the "o" site to regulate trp genes.
- An interesting side bar is that as trp becomes scarcer, there is a greater number of RNA Pol molecules that "make it past" the attenuator site.



- The attenuator is a unique mode of regulation in that it requires completely translated leader sequence for activation when  $\uparrow[\text{trp}]$ .
- This is where gene regulation gets messy.
- Figure, top, illustrates the cotranslation of the trp mRNA as it is being transcribed.



- When  $\uparrow$  [trp] occur, transcription ceases.
- It ceases because 1) the complete leader sequence is synthesized in response to the elevated [trp] and 2) because the combination of translated leader sequence and ribosome sense an increase in the tryptophan concentration.
- Between these two forms, a double knot or a double kink is introduced into the partially transcribed mRNA that does not permit RNA Pol to continue transcribing the trp operon, Figure, above.
- Note that regions "C" and "D" are permitted to bind, inducing the double-knot conformation in the incompletely transcribed mRNA, inhibiting further transcription.



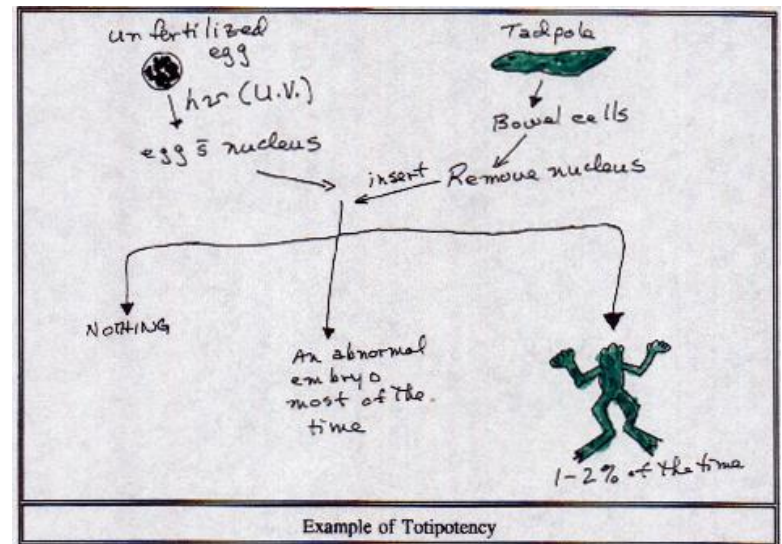
- When tryptophan levels are low, however, the mechanism is altered to permit transcription of the trp operon. When levels of trp are low and when there is a reduction in the numbers of tryptophanyl tRNA, the ribosome following the RNA Pol "stalls out" on the leader sequence. With the reduction in trp-tRNA, the leader sequence can not be translated. Without transcription of the leader sequence, two regions in the attenuator sequence are not permitted to interact, thereby making a "slip-knot" in the mRNA so that translation may continue. This "slip-knot" permits the ribosome[s] to slip right across the attenuator site and continue providing the cell with the machinery to make trp, Figure, above.
- Note that regions "C" and "D" are NOT permitted to interact, although regions "B" and "C" now interact.
- It is this latter interaction that makes the "slip-knot" over which the ribosome[s] slides to continue translating the trp operon.
- In other words, it seems that regions "C" and "D" when combined are the attenuator and regions "B" and "C" "lock-out" the attenuator so transcription may continue.

# An Introduction to Regulation of Gene Expression in Animal Cells

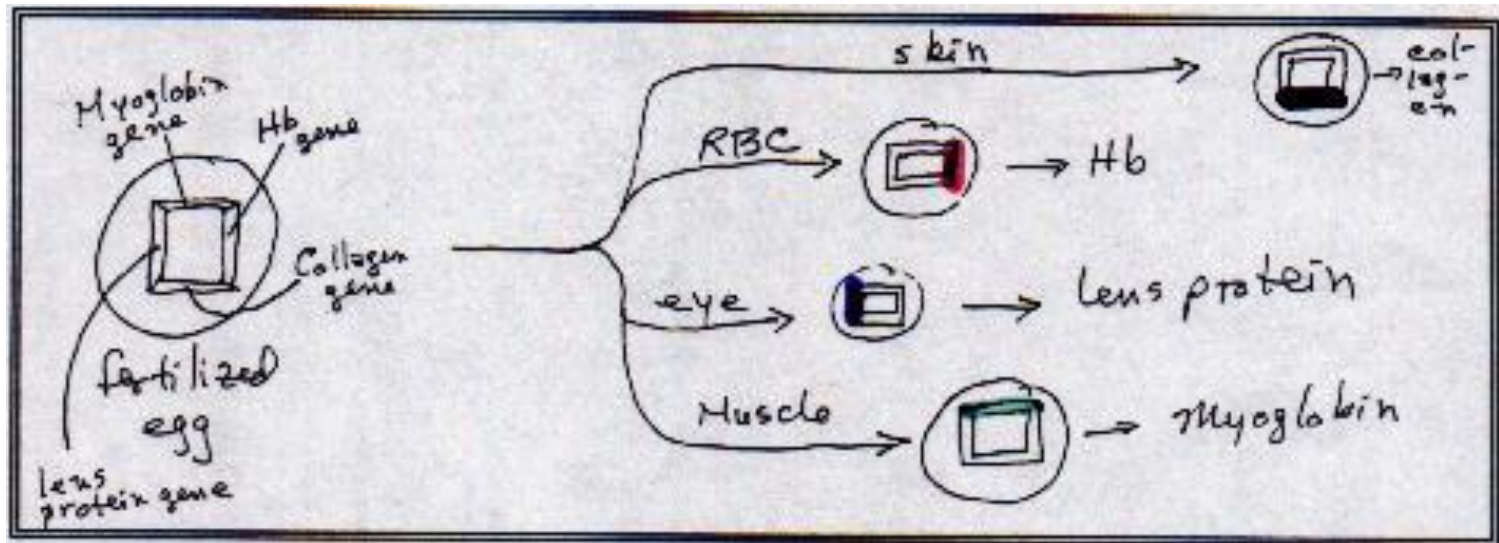
- Question: Do all the different animal cells contain different genes and different amounts of DNA, or are the cell differences due to the expression of genetic material from cells that have the same genomes?
- 
- John Gurdon answered this question in 1964. Gurdon performed a number of experiments with the African toad that allowed, for the first time, people to observe totipotency. Totipotency is the ability of a cell to proceed through all the variable stages of cellular, histological, organ and organismal development to produce a "normal" adult animal.



- In his studies with the African toad, he demonstrated that each and every cell in an animal's body has all the genetic information necessary for making the entire animal, all over, again, if that animal's cell's nucleus is introduced into an unfertilized egg.
- His BIG experiment used the nucleus from a bowel cell and inserted it into an unfertilized egg.
- The nucleus of this egg had previously been destroyed by UV radiation.
- Approximately 1-2% of these inserted eggs actually developed into mature toads, Figure, right.
- The biggest point of this experiment was NOT that only 1-2% of the eggs produced mature animals, but that these animals were produced from a SOMATIC cell, NOT a reproductive cell.



- Therefore, the answer to the question is that each cell in our being contains all the necessary information (DNA) to allow for the production of a mature animal.
- Since each toad produced in this manner was genetically identical, they were called CLONES.
- This process has met with great failure in humans.
- One question to be answered, then, is that if the somatic cells contain all this information, why do only some cells produce specific proteins for that tissue, but not for others?



- This question is easily answered in the following manner:
- The gene[s] for hemoglobin are actively producing hemoglobin in red blood cells during their development.
- The same gene is present in bone cells' nuclei, however, this segment in bone cells is "inactive" and not expressed.
- The same may be said for the collagen gene in bone cells: this gene is actively being expressed in bone tissue for healthy bone formation -- hence, it is "active".
- This gene is inactive in red blood cells since they have no function that requires them to utilize connective tissue.

# Regulatory Proteins: DNA Binding for Transcriptional Control

- There are three motifs of DNA binding proteins: Helix-turn-helix, zinc (metallo) fingers and leucine zippers. All three of these bind to DNA.
- Each have been exhaustively discussed.
- 
- There are some general rules, though, that are applicable to these DNA binding proteins:

- They have a high affinity binding (binds tightly and favorably) to a specific site on the DNA.
- They have a low affinity for the rest of the DNA.
- Only a "small bit" of these proteins may directly contact the DNA.
- The majority of the protein is:
- - scaffolding to hold the protein in the correct conformation for DNA-protein orientation and/or,
  - strapping for interprotein interaction, i.e., dimerization, when required

- Hydrogen bonds and van der Waals forces maintain the protein-DNA interactions.
- The 3 "styles" (motifs) of DNA binding proteins are unique. If these motifs are identified in proteins with function unknown, they are *prima facie* evidence that these proteins in all likelihood bind with DNA in a regulatory manner.
- Helix-turn-helix proteins dimerize in a symmetric manner; the DNA binding site is a symmetrical palindrome between monomers.
- Leucine (leu) zipper proteins do as helix-turn-helix (i.e., as above).
- Zinc finger proteins repeat the binding sites 2-9 times.
- The styles of binding sites for all 3 motifs increase cooperative binding and increase the affinity of the protein of/for/with DNA -- analogy: cooperative binding of oxygen onto hemoglobin as you learned previously.

# Aside

- Metalloproteins (in addition to Zn fingers) regulate numerous genes, particularly in bacteria. Some of the functions of these metalloproteins follow:
  - - metal ion detoxification
    - metal ion uptake
    - metal ion storage
    - metal ion homeostasis
    - metal ions regulate metabolic pathways, as well

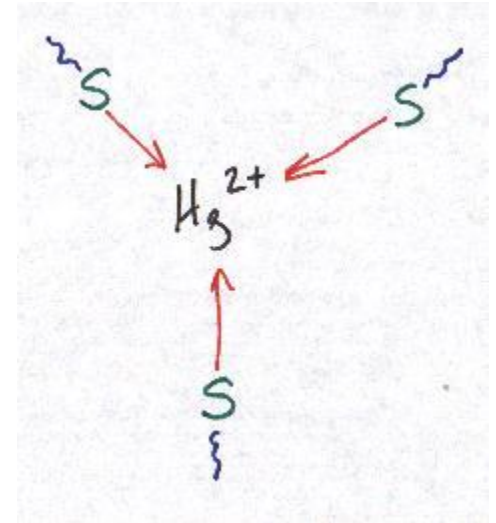
## Examples of metal ions and the role it/they plays are as follow:

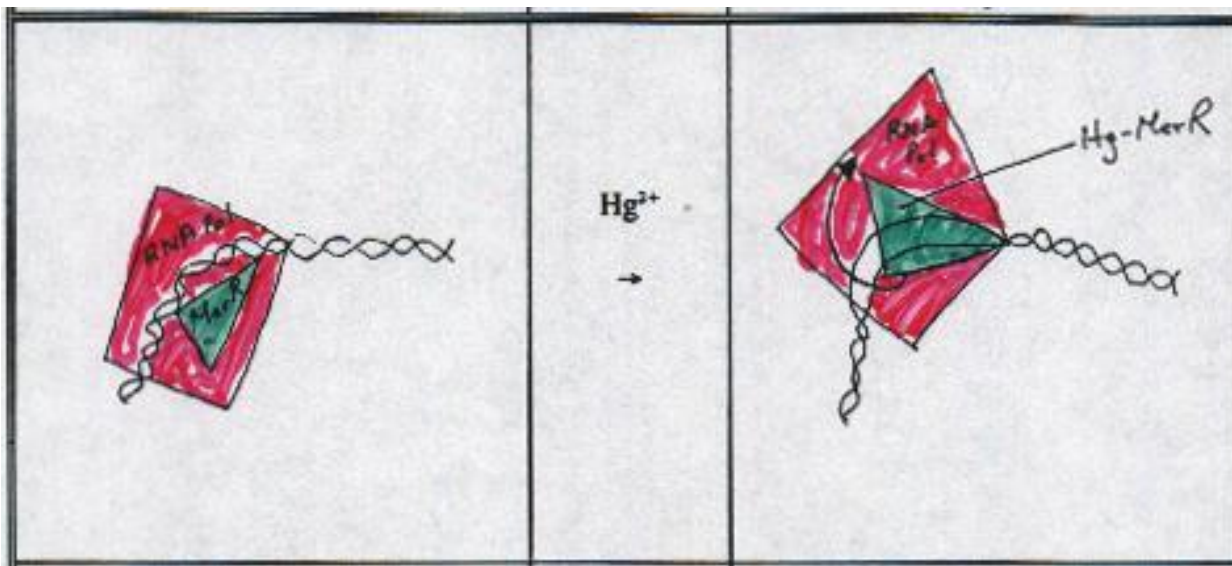
- Fe Catalytic
- Cu Metal ion receptor
- As Metal ion receptor
- Hg Metal ion receptor
- Zn Structural

# Bacterial example:

## $\text{Hg}^{2+}$ -responsive transcription activation

- Bacterial mercury resistance proteins (mer R proteins from mer genes) code for metal detoxification mechanisms:
- 1. Mer R bends DNA to inactivate transcription in the absence of  $\text{Hg}^{2+}$
  2. When  $\text{Hg}^{2+}$  is present,  $\text{Hg}^{2+}$  binds with Mer R
  3. The Hg-mer-R complex is a transcriptional switching protein that holds  $\text{Hg}^{2+}$  in a trigonal geometry, Figure, right. Trigonal geometry places the metal ion ( $\text{Hg}^{2+}$ ) in the center of a flat triangle.
  4. Hg-mer-R untwists DNA to "open it up" as a better template for RNA Pol, Figure, following page.





- In the first step, left panel, there is a repressed DNA sequence; there is no  $Hg^{2+}$  present. There is NO transcription, hence, there is no mercury resistant protein synthesized.
- Once the bacterium detects  $Hg^{2+}$ , right panel, Hg-mer-R complex untwists the DNA and increases "template availability" of the opened DNA to RNA Pol.
- This availability leads to expression of transcription and, hence, the mercury resistant protein is synthesized and the bacterium is protected from the toxic effects of  $Hg^{2+}$ .