Gene Regulation -- The Lac Operon

Specific proteins are present in different tissues and some appear only at certain times during development.

All cells of a higher organism have the full set of genes:

Carrots regenerated from phloem cells in the 1950s

Whole animals can be made from differentiated cells:
Dolly, a sheep created when a nucleus from a mammary gland cell was injected into an enucleated egg, several species of mammals have been cloned.

Here at TAMU, a bull Second Chance
www.reporternews.com/1999/texas/clone0902.html a cat (CopyCat) and though not yet successful, the "Missyplicity" project www.missyplicity.com/ clearly show that all the genes are still present in mature cells, even if they are not transcribed in all tissues.

Regulation can occur at all levels:
- multiple genes
- promoter efficiency
- mRNA stability
- translation
- post translational modification
- protein stability,

Regulation of transcription: is especially effective because
mRNA typically has a short half life (1.8 minutes in E. coli) so stopping mRNA synthesis leads to rapid changes in protein synthesis.
it takes lots of energy to make mRNAs (and proteins) so making them when they are not needed is inefficient.

The Lac Operon has to do with the ability of E. coli to utilize the sugar lactose. Lactose is a 12 Carbon sugar made of 2 simpler 6 carbon sugars, glucose and galactose. Glucose is a very efficient carbon source; it can enter directly into the metabolic paths that provide both energy and substrates for making more complex compounds. If lactose is provided as the carbon source, it must first be broken down into the two component sugars before it can be used.

The enzyme for breaking down lactose in E. coli is called β-galactosidase. The following observations demonstrated that the gene that codes for β-galactosidase in E. coli is regulated:

E. coli grown in glucose as the sole carbon source have about 3 copies of the enzyme β-galactosidase/cell.
*E. coli* grown in lactose as the sole carbon source have about 3,000 copies of the enzyme β-galactosidase/cell.

The system of regulation seen here is called "induction" since synthesis of the enzyme is "turned on" only when needed. Induction typically is used to regulate "breakdown" (catabolic) pathways as opposed to "synthetic" (anabolic) pathways.

Francios Jacob and Jaque Monod won a Nobel prize for their work in describing how the lac-operon functions. They used a genetic approach to address the problem, by identifying mutants that did not have normal regulation of β-galactosidase. We will first look at the model they derived, and then see how the behavior of mutants led to the model.

The lac-operon is actually a series of adjacent genes and regulatory elements in one small part of the *E. coli* circular chromosome.
**Lac-Operon components**

![Diagram of lac operon components]

Definitions: P strands for **promoter**; it is the site where RNA polymerase attaches in order to transcribe mRNA.

Although all promoters have the same function and share similar sequences that are recognized by RNA polymerase, they differ enough so that some are very strong (leading to high levels of transcription) and others are weak (rarely transcribed). Thus, one level of regulating gene expression comes as a consequence of the strength of the promoter at the beginning of the gene.

The **I** gene is called a **regulator gene**; it is transcribed to make a mRNA which is translated to a **repressor protein**. There is a termination signal at the end of the **I** gene.

**O** stands for **Operator**; it is a short sequence of bases that acts like a switch that can be recognized by repressor protein.

**Z, Y** and **A** are all "structural genes (genes that code for polypeptides)

**Z** codes for β-galactosidase; **Y** codes for lactose permease, a protein that functions to actively bring lactose from outside to cell to the inside, even against a concentration gradient. **A** codes for transacetylase, an enzyme that is also needed to breakdown many sugars related to lactose.

One long mRNA is made for the **Z, Y and A** genes; this is the basis for the system being called an operon. All 3 genes that code for enzymes needed to use β-galactoside molecules as a source of carbon and energy are adjacent and are coordinately turned on or off by regulating transcription. Operons are only found in prokaryotes; in eukaryotes, each structural gene has its own promoter and regulatory elements.
**Lac-operon function when only glucose is present;** that is when we expect it to be turned off (numbers indicate steps in the description):

1. The Promoter for the I gene is always "on", but is very weak, so it is transcribed only rarely. A gene that is not regulated, other than by the strength of its promoter, is said to be "constitutive".

2. The I mRNA is translated into a polypeptide; 4 copies make one repressor protein. A typical cell will have only about 10 copies of this protein.

3. In the absence of lactose, the repressor protein binds to the operator, preventing transcription from the second promoter. Almost no ZYA mRNA is made. (The operator is split in 2 parts each with 28 of 35 bases in a palindrome; when the repressor binds it "folds" the DNA so that the promoter is not accessible).

**When only lactose is present the model works as follows:**

1. transcription
2. translation
3. repressor protein
4. mRNA
5. b-gal'ase, permease, TA'ase

lactose (●)
Stepwise: 1. The Promoter for the I gene occasionally is bound by an RNA polymerase to initiate transcription.

2. The I mRNA is translated into the repressor protein.

3. Lactose (actually one stereo-isomer called allolactose which is a minor product of β-gal'ase function) binds to the repressor very efficiently and converts the repressor into an inactive state, where it can't bind the Operator. The process is reversed when all the lactose is digested, so the system again will turn off.

4. When the very strong Promoter for making Z-Y-A mRNA is not blocked, many copies of the mRNA are made. The small amount of lactose that diffuses in is able to initiate induction of transcription of the Z-Y-A mRNA. Even as the message is being made, translation begins and the 3 proteins are made.

5. Translation begins at the 5' end of the mRNA and makes β-galactosidase from the Z gene. There is a stop codon, followed immediately by another AUG start, so many, but not all, ribosomes read on through and make permease from the Y gene. The same process allows some A gene product to also be made.

Mutations that define the Lac-Operon model:

Jacob and Monod found mutants that did not show the normal regulation, that is, synthesis of Z, Y, and A proteins only when lactose was present.

Two kinds of mutants gave constitutive (continuous) synthesis of β-galactosidase, permease and TAase, no matter what carbon source was present

a) Mutants in the I gene (i or i') mutants all were mapped (we will take up mapping later in the course) to a similar location. Flac strains were used to make recipient cells that were partial diploids with two copies of the lac-operon. When the copies created heterozyous cells (I+, I') normal regulation was observed. This indicated that the I gene codes for something that can move and interact with the operators of both copies of the lac-operon present in these cells.
b) Mutants in the promoter that change the base sequence so that it is no longer recognized by the repressor protein are also constitutive for Z, Y and A expression. In this case however, the mutation is dominant in partial diploids:

Since the operator in the upper strand cannot be bound by active repressor, β-galactosidase and TA'ase will always be made. In this cell, permease will show normal regulation since it is only made by the lower copy of the lac-operon. These results told Jacob and Monod that the Operator regulated transcription only of gene on the same DNA molecule.

Mutations in the I gene that prevent the repressor protein from interacting with lactose, but still bind to the operator and permanently turn off all transcription. These are called \(i^S\) mutations for superrepressors.
The partial diploids told Jacob and Monod that the I gene coded for something that could move about the cell and interact with any DNA. Now such gene products are referred to as "trans-acting factors". Some mutations are an example of a trans-dominant effect.

The operator mutations only prevented transcription from the same DNA strand; thus they are "cis-elements" and in this case, act as cis-dominant mutations.

Some technical details:

If both glucose and lactose are present, cells use up the glucose before turning on the lac-operon. When energy begins to become limiting, a signal molecule (cAMP = cyclic-AMP) builds up, binds to a catabolite activating protein (CAP); and the complex in turn binds to a site between the promoter segments of the lac operon. Binding of the cAMP/CAP complex opens the promoter for RNA polymerase binding (providing the lac-repressor isn't bound also).

Mutations in the Z and Y gene can lower the level of allolactose inside the cell and thus affect regulation. This potentially confusing affect is avoided by using IPTG, an artificial inducer that can enter the cell and bind repressor protein without functional Z or Y.

Some mutations in Z and Y have a polar effect. For example, a frameshift in Z will also mean that Y and A won't be made since the ribosomes will remain out-of-phase.

In general, when a trans-acting factor binds to DNA to prevent transcription, the term "negative regulation" is used. Binding that increases transcription (such as the cAMP/CAP complex) is called positive regulation.

From this simple model it is easy to see how a cascade of events could lead to complex regulatory schemes; the product of one reaction could induce the next pathway etc. until development is accomplished.

Repression of genes in biosynthetic pathways

If we look at the pathway for the synthesis of the amino acid histidine, several levels of regulation can be seen. Since E. coli cells spend part of their life cycle in a colon, they may well have plenty of histidine available from digested proteins. On the other hand, if no histidine is available they can make it "from scratch". The pathway for histidine biosynthesis starts with ATP+ PRPP and after 11 enzyme catalyzed reaction histidine is made. Given that a bacterium that wastes the resources and energy used to make histidine when it is not needed would not compete well in nature, it is not surprising that several levels of regulation have evolved. One very common mechanism for biosynthetic (anabolic) pathways is feedback or end-product inhibition. In end-product inhibition, the final product of the pathway binds to the first enzyme unique to the pathway to inactivate it, thus slowing the rate of synthesis. As is the case with lactose interaction with the lac-repressor, histidine does not bind to the active site of the enzyme but to a different (allosteric site) which
causes the protein to assume an inactive configuration. The reaction is also reversible, so that when the endproduct concentration becomes low, the enzyme will regain its active conformation.

If there is really an excess of histidine, it would also be costly to make a long mRNA and 11 unneeded enzymes only to shut off the first enzyme. In this case, a system much like the lac-operon takes over. In his case, since the function is to **turn off** genes when a compound is present, the model is an example of **repression**.

P Reg P O E1 E2 E3 E4 E5 E6 E7 E8 E9 E10 E11

In the case of the histidine operon, the regulator gene product is called a co-repressor. By itself, it cannot bind to the Operator unless it is first bound to histidine. Thus, a relatively subtle difference accounts for regulation in the inducible lac-operon and the repressible his-operon.