Molecular Genetics III
Gene Regulation

The protein products of genes are not needed all of the time. Genes must be turned on or off under different conditions.

Why regulate the expression of genes?

1. Not all proteins are needed all the time. It is economical to produce proteins as needed in amounts that are matched to the need.

2. It is required for cell differentiation. Even though all the cells of an organism have the same DNA, cells of different tissues perform different functions and require different proteins to perform those functions. Some genes are turned on and others turned off in different patterns in different cells.
There are two levels where the expression of genes can be regulated

1. transcription - controlling which mRNA is produced and how fast it is produced
2. translation - controlling how proteins are made from mRNA most regulation goes on at the level of transcription

Most gene expression is regulated through controlling transcription.

Proteins are used to regulate the rate of transcription of genes Some genes are turned on by proteins and some genes are turned off by proteins

**repressor proteins** - stop the expression of a gene by blocking transcription.
**activator proteins** - increase the expression of a gene by promoting the ability of RNA polymerase to associate with a gene and begin synthesizing mRNA

The best example of a system the exhibits both types of transcriptional control is found in *Escherichia coli*. The system is called the **lac operon**.
Genetic organization in prokaryotes

Genes that have a related function are often grouped together and regulated as a unit. All are either on or off.

The collection of genes that are regulated together is called an **operon**.

In the *lac* operon of *E. coli* the three genes are involved in the uptake and metabolism of lactose - a disaccharide

- Z - codes for a protein called β-galactosidase
- Y - codes for a protein called permease
- A - codes for a protein called transacetylase

Z, Y and A are grouped together and have a single controlling region.

The logic of the *lac operon*

Glucose is the preferred sugar of *E. coli*. Only when glucose levels are low will *E. coli* switch to alternative sources of energy.

Lactose is a disaccharide (glucose-galactose). To metabolize lactose, the bond between the two sugars must be broken and then the sugars can be metabolized in glycolysis. If there is already glucose available, there is no need to synthesize a protein that will cleave lactose.

Conditions:

<table>
<thead>
<tr>
<th>Glucose</th>
<th>Lactose</th>
<th>Z,Y,A transcription</th>
</tr>
</thead>
<tbody>
<tr>
<td>present</td>
<td>absent</td>
<td>NO</td>
</tr>
<tr>
<td>present</td>
<td>present</td>
<td>NO</td>
</tr>
<tr>
<td>absent</td>
<td>absent</td>
<td>NO</td>
</tr>
<tr>
<td>absent</td>
<td>present</td>
<td>YES</td>
</tr>
</tbody>
</table>
Two switches are involved in the regulation of the expression of the lac operon. One involves a repressor protein that is sensitive to the presence or absence of lactose. The other involves an activator protein that is sensitive to an indicator of glucose levels in the cell.

When lactose is absent, the repressor protein can bind to a region just upstream from the Z, Y and A genes, called the operator, and block the association of RNA polymerase.

When lactose is present, the lactose associates with the repressor protein and blocks association of the repressor with the operator. The absence of a repressor at the operator potentially allows RNA polymerase to begin transcription.
The other switch involves an activator protein, the **catabolite activator protein** (CAP).

CAP is a general activator protein. It activates the expression of many genes in the cell when ATP levels are low.

When ATP levels are low, ADP levels are high, and some of the ADP is converted to cyclic AMP (cAMP) by adenyl cyclase.

High levels of cAMP is a general indicator of low energy levels in the cell. This means glucose levels are low.

When cAMP is present, it binds to CAP and then CAP associates with the *lac* promoter - which allows RNA polymerase to bind and potentially begin transcription.

If glucose levels are high, cAMP is absent, CAP can’t associate with the promoter, and RNA polymerase can’t begin transcription.
\[ lac \text{ Operon of } E. \text{ coli} \]

![Diagram of the lac operon showing the control system, regulatory region, and coding region.]

\[ \text{Regulation of } lac \text{ Operon} \]

<table>
<thead>
<tr>
<th>Glucose</th>
<th>Lactose</th>
<th>CAP binding site</th>
<th>RNA-polymerase binding site (promoter)</th>
<th>Operator</th>
<th>( lacZ ) gene</th>
</tr>
</thead>
<tbody>
<tr>
<td>++</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>+ -</td>
<td></td>
<td>Repressor</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- -</td>
<td></td>
<td>RNA polymerase</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- +</td>
<td></td>
<td>RNA polymerase</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>+ ++</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- Operon OFF because CAP is not bound
- Operon OFF both because \( lac \) repressor is bound and CAP is not
- Operon OFF because \( lac \) repressor is bound
- Operon ON because CAP is bound and \( lac \) repressor is not
The *lac* operon is a regulated set of genes for the catabolism of lactose. Other operons are involved in anabolic processes.

The *trp operon* is a set of genes involved in the synthesis of tryptophan. When tryptophan is rare the genes for the synthesis of tryptophan are turned on. When tryptophan is common the genes for the synthesis of tryptophan are turned off.

A repressor protein sensitive to the presence of tryptophan is involved in the regulation of the *trp* operon genes.
Gene regulation in eukaryotes is more complex. There are many more conditions that must be satisfied for a gene to be turned on under the proper circumstances in eukaryotes.

Many conditions must be true for a cell to begin synthesizing a specific protein:

If A is true, and if B is true, and if C is true, etc. then the gene should be turned on.

A simple system like the *lac* operon or the *trp* operon with a repressor that binds and blocks transcription and an activator that stimulates transcription would not allow the flexibility needed for eukaryotic cells. There is simply not enough room for all the activator proteins to associate with the DNA in the vicinity of the promoter.

The eukaryotic system allows a gene to be controlled by regulatory sequences that are far from the gene. These are **enhancer** sequences.

An enhancer sequence provides a site for an activator to bind the to the DNA. The enhancer-activator complex then can interact with a transcription factor at a gene’s promoter and allow transcription to begin. DNA looping is often involved in the interaction.
The distant enhancer control system, combined with repressors, and other proteins (co-activators, basal factors) allows many regulatory sequences and regulatory genes to simultaneously exert control over the expression of a single gene.

Chromosome structure also plays a role in gene regulation

The chromatin fibers of chromosomes can exist in a highly coiled state (heterochromatin) or in a relaxed state (euchromatin).

Genes found within heterochromatin are not expressed. The Barr body is an example of a chromosome that is entirely in the heterochromatin form. In other cases, smaller regions of a chromosome may be heterochromatin in some tissues and euchromatin in others.
Transcription can be prevented by DNA methylation. The methylation of cytosine bases in DNA stops RNA polymerase from transcribing a gene. It ensures that genes that are turned off remain turned off.

Although most gene regulation occurs before transcription. A gene’s expression can be regulated even after transcription.

The mRNA must be modified through the addition of a 5’ cap and a poly-A tail, and introns must be removed before it can leave the nucleus.
After mRNA leaves the nucleus regulation is still possible