Chapter for today: Chap. XV

Major points for the day:

1. Gene expression must be regulated
2. Regulation can occur at several levels
3. Cancer can result from inappropriate expression

Reprise

We now know that genes are expressed following the central dogma:

DNA → RNA → Protein

- DNA is the self–replicating genetic material
- It is “transcribed” into an unstable messenger nucleic acid, RNA
- RNA is “translated” into proteins on the ribosome

What we will consider is the ways in which cells regulate expression of genes. We will see that there are multiple steps in gene expression which can be regulated:

- DNA segments can be massively reduplicated by specific over–replication; the multiple gene copies thus created result in much more RNA being produced
- Initiation of transcription can be either decreased or increased
- The stability of the mRNA can be increased or decreased
- The frequency of translational initiation can be increased or decreased
- The protein product can be either modified, or selectively destroyed

Often multiple of these events occur during the expression of a single gene.

- The gene my laboratory studies shows signs of both negative and positive regulation of transcription, translational regulation, and post–translational modification
- The norm may be for genes to be multiply regulated

The need for control of gene expression

The problem of genetic control is this:

- All the cells of your body contain identical copies of your genetic instructions
- Yet these cells appear very different—muscle cells, nerve cells, and epithelial cells have little in common
- This morphological difference reflects a profound difference in expression of genes
  - At the most extreme, reticulocytes (red blood cells) develop from cells which express a single protein, the oxygen carrier hemoglobin
  - Other cells express no hemoglobin
  - Nerve cells express special proteins responsible for propagating electrical signals, which are expressed in no other cell types
  - Muscle cells express vast amounts of the proteins which make up the contractile apparatus, many of which are expressed nowhere else

This wide variety of cell structure and function reflects the need to sensitively control the time, location and extent of expression of a huge variety of genes

- As we shall see, the inability to accomplish this control may lead to disease states, including cancer

Bacteria provide a simple model for genetic control

Beginning in the mid–1950s a pair of French geneticists developed a model system which has stood as a paradigm for genetic control in all cells

- The scientists were François Jacob and Jacques Monod, working at the institute named for the most famous microbiologist of the 19th century, the Institut Pasteur in Paris
- The bacterium that they worked with is a common gut bacterium, *Escherichia coli*
- The system which they focused their work on was the metabolism of a sugar found in milk, lactose.

The problem they were interested in was how a bacterium could rapidly produce an enzymatic activity in response to the presence of an externally available chemical

- Early in their work they had found that addition of lactose to *E. coli* caused the very rapid appearance of an enzyme whose job was to cleave lactose, releasing the two sugars from which it is made, glucose and a related 6–carbon sugar, galactose.
- The enzyme was called β–galactosidase. Its only function in *E. coli* was to split lactose, and was essentially not present when lactose was unavailable
- Careful study showed that in fact a small amount of the enzyme was always present, and that adding lactose stimulated its production almost 500–fold

How was this stimulation, they called it "induction", accomplished

Early in their work they assumed that β–galactosidase proteins would be unable to fold into the proper structure to cleave lactose if no lactose were around, and that induction merely resulted from lactose activating previously synthesized protein by helping it to form the active conformation.

We know now that this is incorrect, as Jacob and Monod showed.

- It is the **expression** of β–galactosidase protein that is induced by lactose
- The level at which this control is exerted is at transcriptional initiation
- In fact, there are three enzymes whose synthesis is controlled coordinately by lactose: β–galactosidase, lactose permease, and trans–acetylase
- Permease is the protein which brings lactose into the cell (it is a sugar transporter). The function of trans–acetylase is still unknown

The three proteins are expressed from three genes, *lacZ*, *lacY*, and *lacA*, respectively

- The three genes are located in a cluster on the *E. coli* chromosome
- The three genes are transcribed by RNA polymerase into a single messenger RNA
- Thus, turning on the transcription of one would turn on the transcription of all three
- The amount of RNA present in the cell is proportional to the amount of protein expressed from each gene

  By regulating transcription initiation, the cell can coordinately regulate expression of all three genes

Such a group of coordinately regulated, linked genes is called an "**operon**" since Jacob and Monod thought that this was the "operational unit" of gene expression in *E. coli*

Two styles of transcriptional control of the *lac* operon

The *lac* operon is under both negative and positive control. The main contribution of Jacob and Monod was to provide a mechanistic explanation of how a gene could be regulated both negatively and positively; this explanation is called the "operon model".

At the 5' end of the *lac* operon is a promoter

- You remember my saying that a promoter is the site at which RNA polymerase binds to DNA to initiate transcription
- The amount of mRNA produced from a given operon depends on the efficiency of promoter recognition
- The *lac* operon regulates expression of the three *lac* genes by regulating the efficiency of that process

Negative Control of Transcription:

Immediately upstream of the three *lac* genes is a fourth gene which is responsible for negative regulation of the operon

- This gene expresses a protein called the ***lac* repressor**
- The repressor can bind to a short sequence of the *lac* promoter
- When repressor binds to that sequence it physically blocks access by RNA polymerase to the DNA
- As a result, as long as it is bound no *lac* mRNA can be initiated

The site at which the repressor binds is called an "**operator**"

- When lactose is present in the cell it binds to the repressor protein, inducing a change in its shape
- Repressor then can not bind to the operator, and RNA polymerase is no longer prevented from binding to the promoter
- As a result, *lac* mRNA is synthesized
- One of the effects of that synthesis is that expression of permease increases drastically
- This causes even more lactose to enter the cell, ensuring that the operon will remain fully active until the lactose is no longer present

Positive Control of Transcription:

Cells would prefer to use glucose as carbon source, even if alternative sources like lactose are present. Therefore the *lac* operon is **only** expressed at a high level when **glucose is not present**

- There are several carbon sources which are used only when glucose is absent: lactose, galactose, & arabinose are three examples
- *E. coli* has evolved a method to control the expression of all of these operons
- It turns out that this mechanism is a positive one, that is, when glucose is absent the transcription of each these genes (subject to the presence of the corresponding sugar substrate) can be turned on

This requires a specific **activator protein** called CAP

- The job of CAP is to facilitate initiation of transcription by RNA polymerase on any of these genes
- It also turns out that the *lac* promoter is not very efficient (its sequence is not recognized well by RNA polymerase)
- CAP protein can bind near the promoter, and facilitate binding by RNA polymerase

- CAP can only bind when a small molecule, cyclic AMP or cAMP, binds to the CAP protein first
  - This is an effect opposite to that of lactose on the repressor
  - cAMP changes the shape of CAP, making it bind its site on the DNA better

  When the concentration of cAMP is lowered, CAP can not bind the cAMP effectively, and thus can not turn on the *lac* promoter

When glucose is present the concentration of cAMP is very low, and the CAP–dependent genes tend to be transcribed very poorly

- This is a type of control which is common in *E. coli*
- Since one gene product controls the expression of multiple operons this is called "**global control**"
- Global positive control is probably the most common form of gene regulation in eukaryotes
- Global negative control also occurs commonly, but specific negative control as seen in the *lac* operon is very rare