LECTURE 28

GENE REGULATION IN PROKARYOTES

A. General considerations:

1. Most microorganisms (cells, actually) are *totipotent* and can produce many different kinds of gene products and in large quantities. Most microorganisms (cells) *don't* produce *all* the gene products of which they are capable for the logical reason of energy conservation. For example, most microorganisms can catabolize (break down) a wide variety of sugars for energy but "prefer" to utilize glucose. Thus, when the organisms are growing on glucose, the gene products for catabolism of other sugars are not synthesized. This led initially to the concept of *structural genes* and *regulatory genes*.

2. There are different levels at which gene expression can be regulated: these levels include transcription, mRNA processing, mRNA stability, translation, and post-translation.

3. There are genes that are expressed continually (*constitutive* genes). These often are referred to as "housekeeping" genes, i.e., those genes whose products are essential components of most living cells. Examples include genes for rRNAs, tRNAs, ribosomal proteins, and enzymes in "general" metabolic processes (e.g., glycolysis).

    a) Note that there are different levels of constitutivity, i.e., constitutivity itself is regulated

4. Regulated "genes" of two general types:

    a) *Inducible* genes $\rightarrow$ "turned on" when needed [catabolic pathways]
    b) *Repressible* genes $\rightarrow$ "turned off" when not needed [anabolic pathways]

5. Operons -- coordinately regulated units of gene expression:

    a) Operons usually consist of a minimum of three *regulatory* "genes" and some number of *structural* genes. The regulatory genes are *repressors, promoters, and operators*; the structural genes generally are contiguous on the chromosome and produce enzymes that function in contiguous steps in a biochemical pathway. Usually, a single, *multigenic* mRNA transcript is produced, facilitating coordinate control of structural gene expression.

    b) Inducible operons usually involve control of structural genes involved in catabolic pathways.

        (i) The typical regulatory logic is for a *repressor* (usually a protein) to bind to an *operator*, preventing transcription of the *structural genes*. The operator is located between the promoter and the structural genes. The *inducer* (usually the catabolite or substance to be broken down) binds to the repressor, preventing its binding to the operator. The RNA polymerase at the promoter is thus free to transcribe the

structural genes. When the inducer is no longer present, the repressor binds to operator and "turns off" the operon.

c) Repressible operons usually involve control of structural genes involved in anabolic pathways.

(i) The typical regulatory logic is that a *repressor* (usually a protein) cannot bind to the *operator* without a *co-repressor* (usually the end product of the anabolic pathway). When co-repressor (end product) is present, a repressor/co-repressor complex binds to the operator and prevents transcription of the structural genes. When the co-repressor (end product) is no longer available, the repressor alone cannot bind to the operator, thus turning on the operon to make more end product (which is the co-repressor).

6. Negative and positive control:

a) Negative control -- regulatory elements "turn off" an operon
b) Positive control -- regulatory elements "turn on" an operon

B. THE LACTOSE OPERON

1. The lac operon is comprised of three regulatory genes and three structural genes. It is an inducible operon under negative control.

| P_I | *lac*I |……..| P_{lac} | *lac*O | *lac*Z | *lac*Y | *lac*A |

a) The regulatory genes include the *repressor* (*lac*I gene), *promoter* (P_{lac} gene), and *operator* (*lac*O gene).

b) The structural genes include *lac*Z (coding for the enzyme β-galactosidase that cleaves lactose into glucose and galactose), *lac*Y (coding for the lac permease, a protein that transports lactose across the cell membrane), and *lac*A (coding for a transacetylase whose function is unknown).

Notes: (i) lactose is a disaccharide of glucose and galactose
(ii) glucose is metabolized by glycolysis
(iii) galactose "induces" the galactose operon, where galactose is converted into glucose

2. Operation of the lac operon is straightforward.

a) A derivative of lactose [*allolactose*] is the *inducer*.

b) The operon is *repressed* in absence of lactose, and *derepressed* in the presence of lactose.

c) Normally, a few molecules of β-galactosidase and lac permease are produced per cell per generation, even in the absence of lactose. This is called *escape synthesis* and occurs during DNA replication when the repressor is temporarily "off" the operator. This escape synthesis is necessary, as β-galactosidase is needed to convert lactose into allolactose (the inducer) and lac permease is needed to transport lactose into the cell.

d) The operon system was elucidated through the use of constitutive mutations in the I (repressor) and O (operator) genes.

   (i) Mutants at the I gene (I^- mutants) produced defective repressor that could not bind to operator.

   (ii) Mutants at the O gene (O^C mutants) were defective operators that could not recognize and bind "normal" repressor.

e) Jacob and Monod used combinations of I and O gene mutations in partial diploids (through sexduction). They demonstrated the following:

   (i) The I^+ allele was *trans* dominant to the I^- allele, (i.e., an I^+ allele on an episome and with a *lac*Z mutant could "regulate" a normal *lac*Z^+ allele on the chromosome with an I^- allele). This was compatible with the hypothesis that the I gene product produced a substance that was diffusable through the cytoplasm (a protein, in this case).

   (ii) The O^C mutation was *cis* dominant but <u>not</u> *trans* dominant (i.e., an O^C allele only regulated *lac*Z^+ alleles that were contiguous or on the same chromosome). This was compatible with the hypothesis that the O gene was a <u>site</u> and did not produce a substance that was diffusable through the cytoplasm.

f) Other aspects of the lac operon are:

   (i) A low efficiency promoter for the I gene "regulates" the regulatory gene.

   (ii) There are a number of interesting mutants at the lacI gene (e.g., I^S).

   (iii) The *lac* promoter (*lac*P gene) also is *cis* dominant (as might be expected).

   (iv) The lac promoter contains two separate components: a binding site for the RNA polymerase (plus σ), and a binding site for the CAP protein.

C. CATABOLITE REPRESSION [GLUCOSE EFFECT]

1. This is a mechanism to insure the preferential utilization of glucose as an energy source. It involves glucose, CAP (*Catabolite Activator Protein)* protein, and cyclic AMP.

2. Cyclic AMP is "sensitive" to intracellular concentration of glucose. When glucose is present, levels of cyclic AMP are low; whereas when glucose is absent, levels of cyclic AMP are high.

3. Under conditions of lactose but no glucose, a cyclic AMP + CAP protein complex binds to the CAP site in the *lac*P gene and increases the efficiency of RNAP (σ) binding.

4. However, when both lactose and glucose present, cyclic AMP levels are low and RNAP (σ) binding at *lac*P is very inefficient.

5. The same is true for almost all sugar-catabolizing operons. In this case, the regulatory elements (cyclic AMP + CAP protein) are required to "turn on" (actually, increase binding efficiency of) the operon. This constitutes an example of *positive control*.