LECTURE 29

GENE REGULATION IN PROKARYOTES [CONTINUED]

A. The tryptophan (trp) operon

trpR → *trp* repressor protein
trpP_1 → *trp* promoter (high efficiency)
trpO → *trp* operator
trpL → *trp* attenuator
trpE → ε polypeptide of anthranilate synthetase
trpD → δ polypeptide of anthranilate synthetase
trpP_2 → internal promoter (low efficiency)
trpC → indole gylcerophosphate synthetase
trpB → β polypeptide of tryptophan synthetase
trpA → α polypeptide of tryptophan synthetase

B. There are two major modes of regulation of the trp operon: *Repression* and *Attenuation*.

1. Repression follows the regulatory logic of a typical repressible system.

   a) A repressor – co-repressor (tryptophan) complex binds to the operator and represses operon by impeding RNA polymerase (bound at P_1).

   b) In the absence of tryptophan, no repressor is bound to operator and the operon is derepressed.

   c) A low efficiency promoter (P_2) leads to low levels of products of structural genes trpC, trpB, and trpA.

   d) In trpR mutants where the repressor protein is non-functional (cannot bind to co-repressor), addition of tryptophan to medium *still* reduces trp structural gene products by at least tenfold. This was the first clue that another type of regulation [*attenuation*] was operating.

2. Attenuation: a second method of regulation that is independent of regulation by "repression".

   a) Involves the trpL gene, a sequence of 162 nucleotides that produces an mRNA leader sequence; near the 3' end of this leader sequence (trpL gene) is a region called the *attenuator*.

      Attenuator: A nucleotide sequence almost identical to the transcription termination

signals found in genes with *rho*-independent termination, i.e., a GC-rich tract with dyad symmetry that permits a "hairpin" (stem-and-loop) configuration, followed by a poly A (poly U) tract.

b) Attenuation occurs via the "premature" termination of transcription at the attenuator. The "premature" mRNA is only 140 nucleotide long, and is produced only when a tRNA carrying tryptophan ($tRNA^{Trp}$) is available. This suggests that either tryptophan, a "charged" $tRNA^{Trp}$, or both is (are) somehow involved in attenuation control of the tryptophan operon.

c) The trpL gene actually contains four different sequence tracts that can base pair in various combinations to form "hairpin" (stem-and-loop) structures; sequences of these four are such that "hairpins" can form between…

Region 1 & Region 2 Region 2 & Region 3 Region 3 & Region 4

d) Because Region 2 can only form a "hairpin" with either Region 1 or Region 3 (at the same time), there are *only* two possible secondary structures for the trpL mRNA leader sequence:

(i) Region 1 paired with Region 2, and Region 3 paired with Region 4, or…

(ii) Region 2 paired with Region 3, leaving Region 1 and Region 4 unpaired

e) Regions 3 and 4 are the *attenuator*, so when the secondary structure is Region1 paired with Region 2 (a Region 1 – Region 2 hairpin), transcription is terminated prematurely as the RNA polymerase will "recognize" what is a "normal" transcription-termination signal. If the alternate secondary structure (a Region 2 – Region 3 hairpin) occurs, there is no transcription-termination signal and transcription proceeds into the trp E gene (the first "structural" gene.

f) If transcription proceeds through the trpL gene, the 162 mRNA leader sequence (product of the trpL gene) contains an AUG start codon near (but not at) the 5' end of the mRNA, followed by 13 codons for (13) amino acids and a UGA stop codon. The nucleotide-sequence tract that is Region 1 occurs just *before* the UGA stop codon (and hence is included *within* the 13 codons) but just *after* two codons (UGG) that call before the amino acid *tryptophan*.

g) In the absence (or low levels) of tryptophan (and "charged" $tRNA^{Trp}$), the ribosomes "stall" during translation but at a location that *prevents* Region 1 from forming a "hairpin" with Region 2. The net result is that Region 2 pairs with Region 3, preventing formation of the transcription-termination signal formed by pairing of Region 3 with Region 4. The consequence of this is that transcription *continues* past the attenuator and into the trpE gene.

h) In the presence of tryptophan (and "charged" $tRNA^{Trp}$), the two trp codons are "read", translation stops at the UGA codon (theoretically producing a 14 amino acid peptide), and Region 1 is available to pair with Region 2, leaving Region 3 available to pair with Region 4. The latter forms a transcription-termination signal, transcription is terminated at the attenuator, and the tryptophan structural genes are *not* transcribed.

3. Three final points about attenuation:

   a) Attenuation control accentuates (is in fact dependent upon) the close coupling of transcription and translation in prokaryotes.

   b) Attenuation control can regulate mRNA production (and tryptophan production) even when the regulatory genes are mutated.

   c) Regulation by attenuation is not unique to the tryptophan operon, and is known for at least five other amino acid operons in *E. coli*. In one of these (the histidine operon), attenuation is thought to be the only mechanism of transcriptional control.

THE ARABINOSE OPERON: AN EXAMPLE OF POSITIVE AND NEGATIVECONTROL

A. The arabinose operon is quite complex and not fully understood. Only the basics are given below.

   1. Arabinose is a sugar that is broken down (catabolized) by three structural genes: *ara*B, *ara*A, and *ara*D.

   2. The regulator protein, *ara*C, can both repress (negative control) and induce (positive control) transcription of arabinose structural genes.

   3. In the absence of arabinose, the *ara*C protein binds to a site called *ara*O_2 (an operator) and a site called *ara*I. This is thought to result in a change in the secondary structure of P_{BAD} (the promoter for the structural genes) that prevents or interferes with RNA polymerase binding at P_{BAD}.

   4. In the presence of arabinose and cyclic AMP, the *ara*C protein becomes an activator of transcription. In this case, there is involvement of an arabinose-*ara*C protein complex and a cyclic AMP-CAP protein complex, both of which have binding sites at the *araI* site. The "model" is that the two complexes positively affect binding of RNA polymerase to P_{BAD}.

TRANSLATIONAL AND POST-TRANSLATIONAL REGULATION

A. Translational control:

1. In prokaryotes, multigenic mRNAs mean that equal amounts of mRNA exists for each "gene" in the message. Direct measurements, however, have demonstrated that many protein products from multigenic mRNAs are not produced in equivalent amounts. For example, in *E. coli* growing under optimal conditions, the lac operon structural-gene proteins are produced in different amounts, *viz*.,

β-galactosidase	3,000 molecules
lac permease	1,500 molecules
transacetylase	600 molecules

2. There are four postulated mechanisms:

   a) Unequal efficiencies of translation initiation;

   b) decreased efficiency of ribosome binding as translation moves downstream through intergenic regions; hairpins" (stem-and-loop) regions have been implicated;

   c) differential rates of mRNA degradation; and

   d) negative self-regulation, where translation is inhibited by one of the translated products.

B. Post-translational control:

1. The best known mechanism is where the end product of a biosynthetic pathway binds to and inhibits an early (usually the initial) enzyme in the pathway. This is called *feedback* or *end-product inhibition*.

2. One example is tryptophan which can feedback inhibit the enzyme anthranilate synthetase, the first step in the tryptophan biosynthetic pathway.