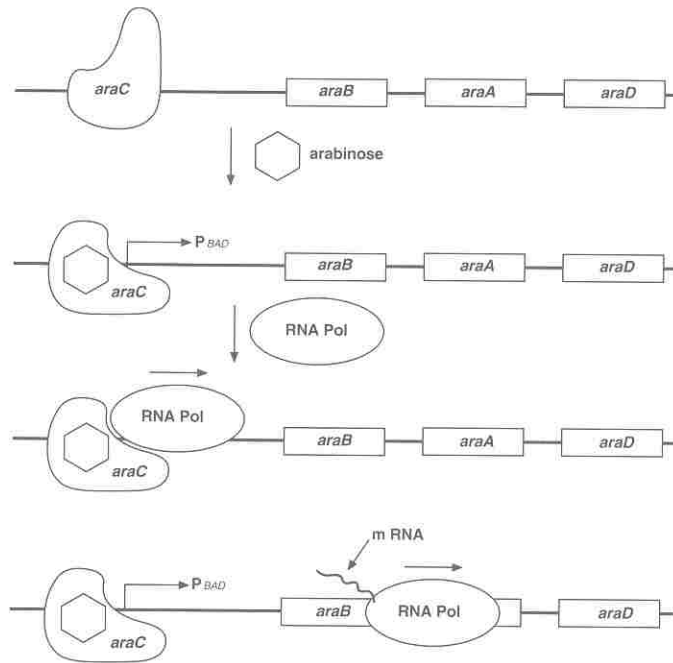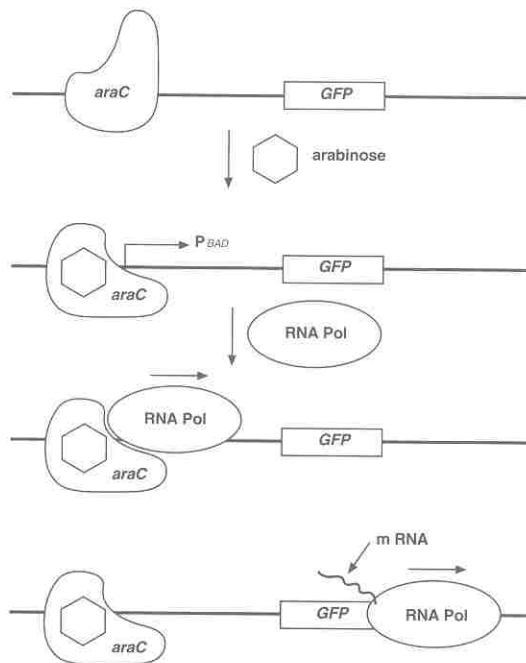# The Arabinose Operon

Expression of Green Fluorescent Protein

Appendix D Gene Regulation

Our bodies contain thousands of different proteins which perform many different jobs. Digestive enzymes are proteins; some of the hormone signals that run through our bodies and the antibodies protecting us from disease are proteins. The information for assembling a protein is carried in our DNA. The section of DNA which contains the code for making a protein is called a gene. There are over 30,000–100,000 genes in the human genome. Each gene codes for a unique protein: one gene, one protein. The gene that codes for a digestive enzyme in your mouth is different from one that codes for an antibody or the pigment that colors your eyes.

Organisms regulate expression of their genes and ultimately the amounts and kinds of proteins present within their cells for a myriad of reasons, including developmental changes, cellular specialization, and adaptation to the environment. Gene regulation not only allows for adaptation to differing conditions, but also prevents wasteful overproduction of unneeded proteins which would put the organism at a competitive disadvantage. The genes involved in the transport and breakdown (catabolism) of food are good examples of highly regulated genes. For example, the sugar arabinose is both a source of energy and a source of carbon. *E. coli* bacteria produce three enzymes (proteins) needed to digest arabinose as a food source. The genes which code for these enzymes are not expressed when arabinose is absent, but they are expressed when arabinose is present in their environment. How is this so?

Regulation of the expression of proteins often occurs at the level of transcription from DNA into RNA. This regulation takes place at a very specific location on the DNA template, called a promoter, where RNA polymerase sits down on the DNA and begins transcription of the gene. In bacteria, groups of related genes are often clustered together and transcribed into RNA from one promoter. These clusters of genes controlled by a single promoter are called operons.

The three genes (*araB, araA* and *araD*) that code for three digestive enzymes involved in the breakdown of arabinose are clustered together in what is known as the arabinose operon.[3] These three proteins are dependent on initiation of transcription from a single promoter, P_{BAD}. Transcription of these three genes requires the simultaneous presence of the DNA template (promoter and operon), RNA polymerase, a DNA binding protein called *araC* and arabinose. *araC* binds to the DNA at the binding site for the RNA polymerase (the beginning of the arabinose operon). When arabinose is present in the environment, bacteria take it up. Once inside, the arabinose interacts directly with *araC* which is bound to the DNA. The interaction causes *araC* to change its shape which in turn promotes (actually helps) the binding of RNA polymerase and the three genes *araB, A* and *D*, are transcribed. Three enzymes are produced, they break down arabinose, and eventually the arabinose runs out. In the absence of arabinose the *ara*C returns to its original shape and transcription is shut off.

The DNA code of the pGLO plasmid has been engineered to incorporate aspects of the arabinose operon. Both the promoter (P_{BAD}) and the *araC* gene are present. However, the genes which code for arabinose catabolism, *araB, A* and *D*, have been replaced by the single gene which codes for GFP. Therefore, in the presence of arabinose, *araC* protein promotes the binding of RNA polymerase and GFP is produced. Cells fluoresce brilliant green as they produce more and more GFP. In the absence of arabinose, *araC* no longer facilitates the binding of RNA polymerase and the GFP gene is not transcribed. When GFP is not made, bacteria colonies will appear to have a wild-type (natural) phenotype—of white colonies with no fluorescence.

This is an excellent example of the central molecular framework of biology in action:
DNA→RNA→PROTEIN→TRAIT.