7.4 Controlling Gene Expression

Not all proteins are required by all cells at all times. It would be inefficient for a cell to transcribe and translate all its genes at all times. Rather, both prokaryotic and eukaryotic cells regulate gene expression in response to their own life cycles and environmental conditions. For example, human **insulin** is only required when the glucose level in the blood is high. Similarly, the *E. coli* enzyme that facilitates the breakdown of lactose is only transcribed and translated when the *E. coli* bacteria are exposed to lactose. The optimal functioning of an organism requires that genes be turned on and off as they are needed. Even though each cell contains the entire genome of an organism, cells know which genes to express and when. Intricate systems have evolved to fine-tune gene expression in both prokaryotes and eukaryotes. These systems ensure that all cells express only the genes they require, and that gene expression responds to the cellular environment.

insulin a hormone produced in the pancreas that lowers the blood glucose level by promoting the uptake of glucose by the body cells

Some proteins are always needed in a cell, and their genes are continuously transcribed and translated. Known as housekeeping genes, these genes regulate processes such as metabolism, growth, and DNA replication and transcription. The products of other, more specialized genes may be found only in certain types of cells or under particular environmental conditions. For example, liver cells require repair enzymes to manage the toxins in the body, and genes that produce hemoglobin molecules are transcribed and translated only in the cells that give rise to red blood cells.

In this section, you will learn about the mechanisms that cells use to regulate gene expression. Not only are these systems important to the biology of all organisms, but they also provide important tools that allow molecular biologists to manipulate the expression of eukaryotic genes by bacteria, viruses, and even cancer cells.

Prokaryotic Gene Control Mechanisms

Gene expression in prokaryotes is regulated in response to the concentrations of two molecules: lactose and tryptophan. Both of these responses are examples of negative feedback control.

The *Lac* Operon

The sugar lactose, a potential source of energy for prokaryotes, must be acquired directly from the environment. To regulate the expression of the genes that are required for lactose metabolism, prokaryotes use what is known as the operon model of gene expression. The ***lac* operon** is a cluster of three genes that code for the proteins involved in the metabolism of lactose. The *lac* operon consists of a promoter (the site where DNA transcription begins), an **operator** (the sequence of bases that control transcription), and the coding regions for the various enzymes that actually metabolize the lactose. Upstream from the operon is a gene that codes for a **repressor protein**. This protein takes cues from the environment (in this case, the concentration of lactose within a cell) and regulates the production of the lactose-metabolizing proteins. For the *lac* operon, this protein is called the *lacI* protein or *lac* repressor. The genes that code for the *lac* repressor are always transcribed, so the *lac* repressor is always present within a cell. How this protein behaves, however, and the rate of synthesis of the other *lac* proteins, depends on the concentration of lactose in the cell.

***lac* operon** a cluster of genes that contains the DNA sequences to regulate the metabolism of lactose

operator the region in the operon that regulatory factors bind to

repressor protein a protein that binds to the operator to repress gene transcription

In the absence of lactose, the *lac* repressor is active and binds to the operator (**Figure 1(a)**, next page). This keeps RNA polymerase from binding to the promoter region and stops the lactose-metabolizing enzymes from being synthesized. When lactose is present within a cell, some of it binds to a site on the *lac* repressor, rendering it inactive (**Figure 1(b)**, next page). The inactive *lac* repressor is unable to bind to the operator and block transcription. With the operator free of obstruction, RNA polymerase is able to bind to the promoter region, and transcription of the *lac* genes begins. The enzymes that metabolize lactose are then synthesized and start to break

down the lactose in the cell. As the concentration of lactose in the cell decreases, the amount of inactivated *lac* repressor decreases (that is, there is less and less lactose to deactivate the *lac* repressor). Eventually, the reactivated *lac* repressor again binds to the operator, stopping transcription.

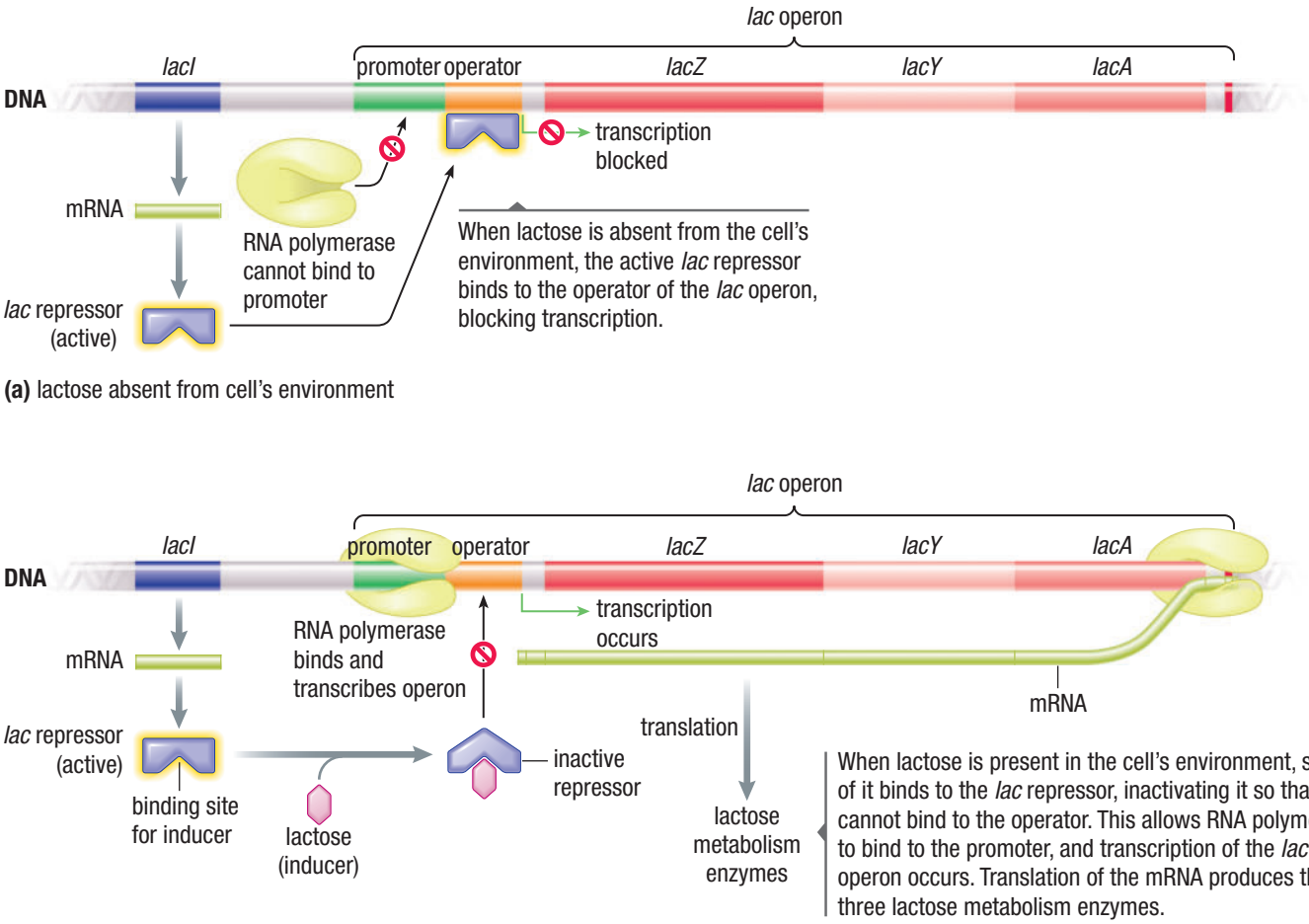


Figure 1 Regulation of the *lac* operon in prokaryotes by the (a) absence and (b) presence of lactose

Note that the lactose itself acts as a signal molecule, telling the cell when to synthesize the lactose-metabolizing enzymes. This type of signal molecule is called an **inducer**, since it serves to initiate the production of enzymes. The *lac* operon is known as an inducible operon because the inducer inactivates the repressor and allows the gene to be transcribed. There is a direct correlation between the amount of lactose in a cell and the rate at which the *lac* enzymes are synthesized. As the concentration of lactose in the cell increases or decreases, so too does the transcription of the *lac* genes. This is an important way for the cell to conserve energy, by not synthesizing proteins when they are not necessary.

inducer a signal molecule that triggers the expression of an operon's genes

The *trp* Operon

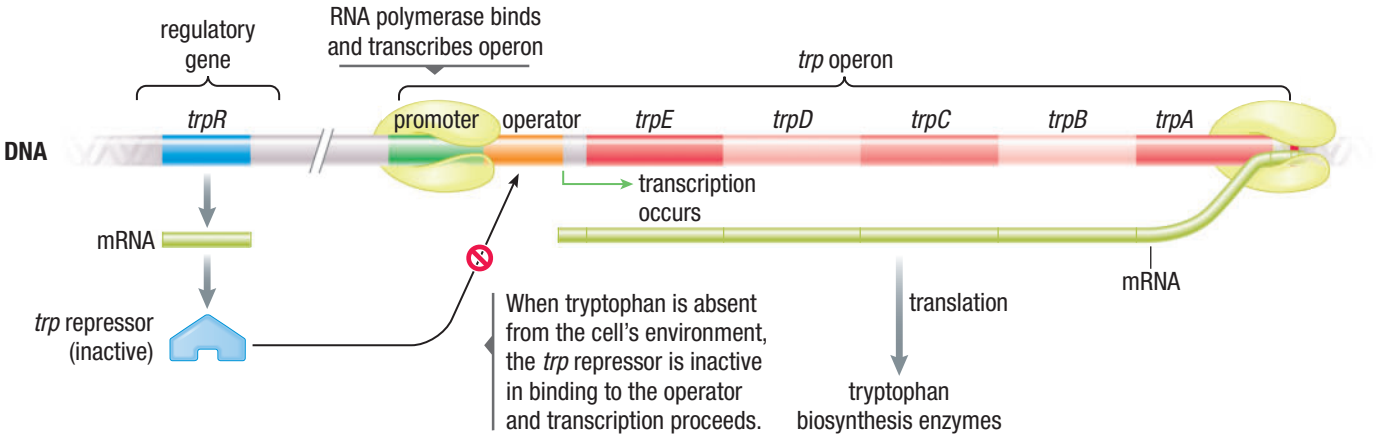
Tryptophan is an important amino acid that is used to build proteins. Most prokaryotic cells are able to synthesize tryptophan independently, but they can also take it up directly if it is available in the environment. The operon that regulates the production of tryptophan in a cell is called the *trp* operon.

The *trp* operon has the same structure as the *lac* operon: a promoter and an operator that precede the genes coding for tryptophan-synthesizing enzymes. There is also a gene that codes for a *trp* repressor protein. This repressor protein is always synthesized (like in the *lac* operon), but the difference is in how the repressor protein

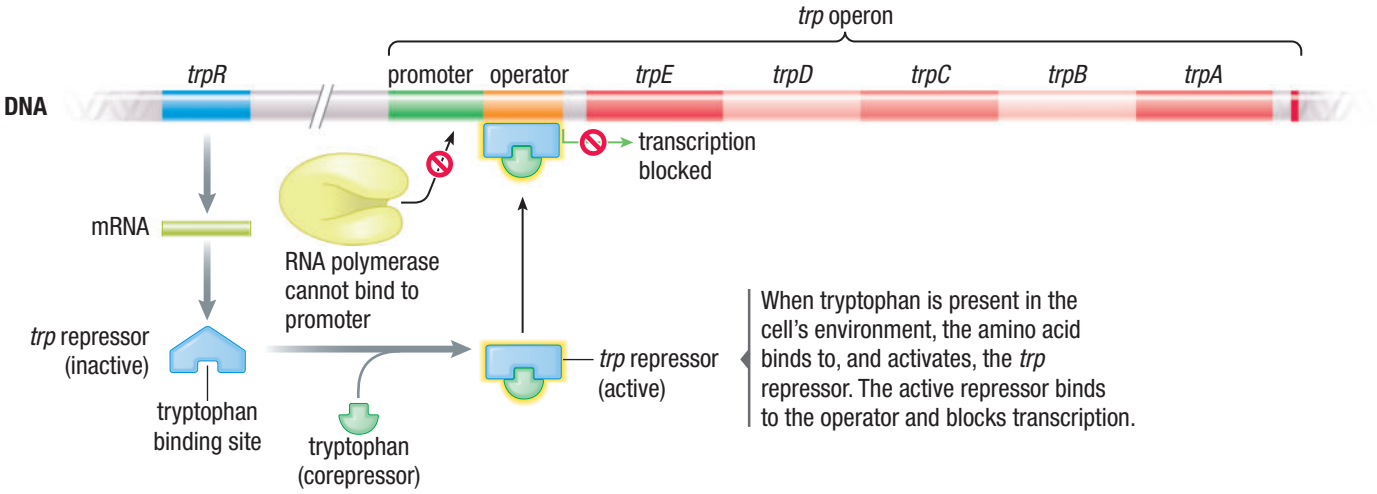
acts to regulate the expression of the tryptophan enzymes. Whereas the *lac* repressor protein is inactivated by a signal molecule, the *trp* repressor protein is activated in the presence of tryptophan.

When tryptophan is not present in the environment, the repressor protein is in an inactive state and does not bind to the operator (**Figure 2(a)**). RNA polymerase is able to bind to the promoter region, and the transcription of the genes that are responsible for the biosynthesis of tryptophan can proceed. When tryptophan is present in the environment, the cell can conserve energy by using the available tryptophan and stopping the transcription of the genes that code for the enzymes involved in the biosynthesis of the amino acid (**Figure 2(b)**). Tryptophan acts as a signal molecule and activates the repressor protein. The activated repressor protein is then able to bind to the operator and stop transcription of the tryptophan-synthesizing genes. When a signal molecule functions in this way, it is called a **corepressor**: it serves to repress (rather than induce) the expression of a set of genes. As the concentration of environmental tryptophan decreases, the repressor proteins become deactivated and the tryptophan-synthesizing genes are transcribed. This negative correlation between the amount of tryptophan in the environment and the rate of tryptophan synthesis is an example of a negative feedback mechanism.

corepressor a signal molecule that binds to a regulatory protein to reduce the expression of an operon's genes



(a) tryptophan absent from cell's environment



(b) tryptophan present in cell's environment

Figure 2 Regulation of the *trp* operon by the (a) absence and (b) presence of tryptophan

Eukaryotic Gene Control Mechanisms

Eukaryotic gene expression requires a larger number of steps. Therefore, the methods for regulating eukaryotic gene expression are more complex than the methods for regulating prokaryotic gene expression. Eukaryotes do not use the operon system just described.

Instead, the control mechanisms in eukaryotes fall into four categories:

- transcriptional (as mRNA is being synthesized)
- post-transcriptional (as mRNA is being processed)
- translational (as the protein is being synthesized)
- post-translational (after the protein has been synthesized)

Transcriptional Regulation

Although eukaryotic gene regulation occurs at multiple levels, the most common type of regulation occurs during transcription. You may remember from Chapter 6 that in eukaryotic chromatin, the DNA is wrapped around histone proteins. Therefore, the gene promoters are not accessible to the proteins that initiate transcription. This configuration keeps the gene promoters inactive. For a gene to be transcribed, the chromatin must be partially unwound to expose the promoter.

In one type of transcriptional regulation, the promoter is exposed when an activator molecule binds to a sequence that is upstream of the gene's promoter and signals a protein remodelling complex (**Figure 3**). As a result, the histone core proteins are displaced from the DNA (chromatin remodelling), exposing the promoter. In another type of transcriptional regulation, an activator molecule is bound to a regulatory sequence upstream of the gene to be transcribed. This signals an enzyme that can add an acetyl group (CH_3COO^-) to histones. The addition of the acetyl group to histones loosens their association with DNA, and the promoter becomes accessible.

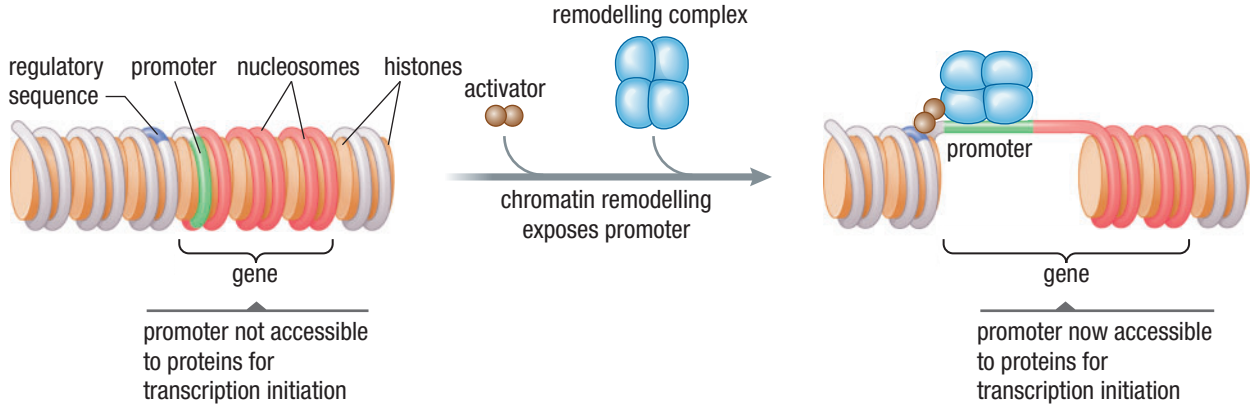


Figure 3 A chromatin remodelling complex exposes the promoter, and thus enables transcription.

To initiate transcription, a series of proteins, called general transcription factors, accumulate on the promoter. They bind to a specific region of the promoter (TATA box) and provide a substrate that the RNA polymerase can bind to and thus begin transcription (**Figure 4**, next page). Together, the general transcription factors and RNA polymerase form the transcription initiation complex. This establishes a base rate of gene transcription, which can be further altered by additional proteins called activators and repressors, depending on the needs of the cell. The activators and repressors (like the regulatory proteins in the *lac* and *trp* operons) attach themselves to the promoter to increase or decrease the rate of transcription.

Another method of eukaryotic gene regulation that is important for biotechnology applications is methylation. A methyl group ($-CH_3$) is added to the cytosine bases in the promoter of a gene, inhibiting transcription. This effect is called silencing.

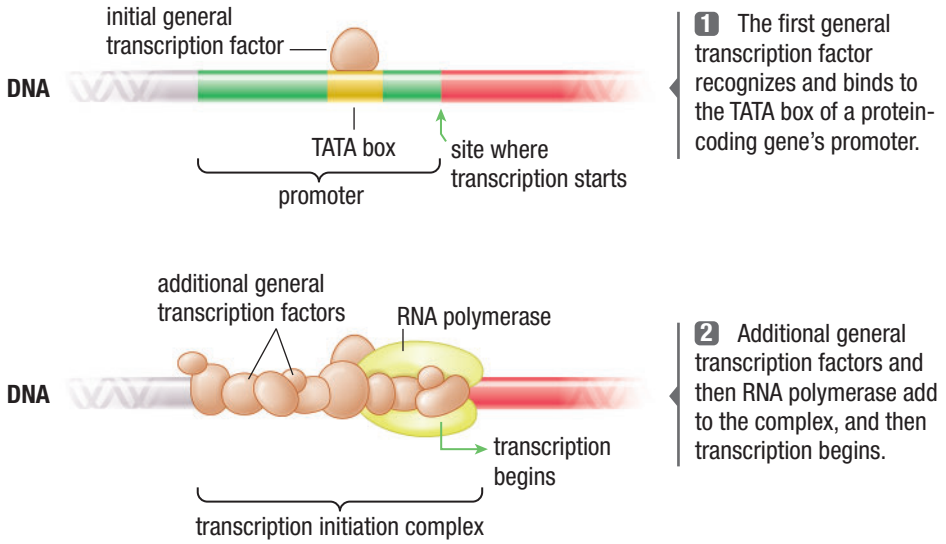


Figure 4 General transcription factors accumulate on the promoter to initiate transcription.

Methylation is another way to put genes or entire regions of chromosomes "on hold" until they are required. For example, the genes that code for the production of hemoglobin are in an inactive, methylated state in almost all the body cells. However, the cells in bone marrow, which produce red blood cells, use specific enzymes to remove the methyl groups and allow transcription.

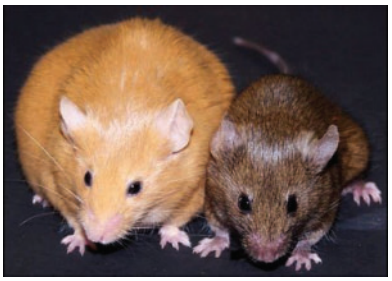


Figure 5 A mouse whose agouti genes have been turned on (left) is very different from a mouse whose agouti genes have been silenced.

An example of the power of gene methylation is seen in agouti mice. Mice whose agouti genes have been turned on can look entirely different in both colour and size, even though they are genetically identical. One mouse may be small and brown, while its twin may be obese and yellow (**Figure 5**). In normal, healthy mice, the agouti genes are kept in the "off" position (silenced). Methyl groups are attached to the corresponding regions of DNA, and transcription is prevented. In yellow and/or obese mice, however, the same genes are not methylated. Thus, these genes are expressed or "turned on." Mice whose agouti gene is "on" have a higher risk of cancer and diabetes.

A number of environmental triggers have been shown to contribute to the agouti gene promoter being methylated or not methylated. One suspected trigger is a chemical called bisphenol A, which, until recently, was found in many plastic bottles, including baby bottles. Researchers exposed pregnant mice to bisphenol A and watched as more of their offspring developed into yellow, obese mice than would normally be expected. These results supported the hypothesis that exposure to bisphenol A results in demethylation, the removal of methyl groups from DNA. However, not all the offspring grew up to be obese. Researchers concluded that bisphenol exposure did not guarantee obesity in mice, but simply increased the risk of developing obesity. The silencing effect is used by researchers who wish to test the effects of a single gene. Once the coding sequence is located, they can methylate the gene and observe the effect of its absence on the organism they are studying.

Post-transcriptional Regulation

Post-transcriptional regulation influences gene expression by several mechanisms, including changes in pre-mRNA processing and the rate at which mRNAs are degraded. Alternative splicing is one example of post-transcriptional regulation. Alternative splicing produces different mRNAs from pre-mRNA by removing different combinations of introns. The remaining exons are spliced together (Section 7.2). Depending on which protein is required by the cell, an intron in one pre-mRNA may be considered an exon in another pre-mRNA and may therefore not be spliced out of the second pre-mRNA transcript. The resulting mRNAs are translated to produce

a family of related polypeptides and their associated final proteins with various combinations of amino acid sequences. Alternative splicing is a powerful tool, which is used by cells to optimize the production of different proteins depending on cell type. Perhaps 75 % of human genes are alternatively spliced at the pre-mRNA level.

Another example of post-transcriptional regulation involves binding masking proteins to mRNA. When the mRNA is associated with a masking protein, it does not undergo protein synthesis. Masking proteins are a common form of control in many animal eggs, keeping mRNAs in an inactive form until an egg has been fertilized and embryonic development is underway. When it is time for the mRNA to be translated, other proteins remove the masking proteins.

A third example of post-transcriptional regulation changes the rate of degradation of mRNAs. A regulatory molecule, such as a hormone, will directly or indirectly affect the rate of mRNA breakdown. For example, in the mammary gland of a rat, it takes about 5 h for half of the mRNA for casein (a milk protein) to break down. In the presence of the hormone prolactin, the time increases to 92 h. Prolactin is a hormone that is synthesized in the brain and the mammary glands. When a large amount of casein is needed for milk production, there is an abundant supply of prolactin and the lifespan of the casein mRNA is extended.

Translational Regulation

Translational regulation occurs during protein synthesis by a ribosome. One important mechanism changes the length of the poly(A) tail of the mRNA molecules. Specific enzymes can add or delete repeating sequences of adenine at the ends of the mRNA molecules. This change in the length of the poly(A) tail may increase or decrease the time that is required to translate the mRNA into a protein. Scientists do not completely understand how this mechanism functions, but they think that the cell may take cues from the environment or intracellular molecules to adjust the rate at which certain mRNA molecules are translated.

Post-translational Regulation

After the mRNA is translated and proteins are synthesized, the cell can still regulate expression by limiting the availability of functional proteins. Three methods are used: processing, chemical modification, and degradation.

When proteins are initially synthesized, they are in an inactive form and must be activated by various processing mechanisms. Insulin (which regulates the level of glucose in the blood) is initially synthesized as proinsulin, an inactivated precursor. A processing mechanism removes specific sections of the protein and renders it active. The cell can regulate these types of processing mechanisms to control the availability of activated proteins and thus regulate the end product of certain genes.

During the chemical modification of a protein, certain chemical groups that are attached to the protein are added or deleted, affecting its function. The presence or absence of these chemical groups puts the protein "on hold" until it is needed. Once environmental and cellular conditions are sufficient, the groups are added or removed and the protein can carry out its function.

Like most other biosynthesized molecules, proteins are subject to constant degradation both inside and outside the cell. Some proteins are used for only a few minutes before they are broken down, while others can last the entire lifetime of an organism. This rate of degradation is under regulatory control, modifying the rate at which the products of gene expression are available. Short-lived proteins are tagged with a small protein called ubiquitin, which is recognized by the degradation mechanisms of the cell. Adding or removing these tags can either shorten or extend the functional life of a protein. The degraded proteins are broken down into their constituent amino acids and recycled to synthesize new proteins.

Table 1 summarizes the gene control mechanisms in eukaryotic cells.

Table 1 Four Levels of Control of Gene Expression in Eukaryotic Cells

Type of control	Description	Specific examples
transcriptional	regulates which genes are transcribed (DNA to mRNA) or controls the rate at which transcription occurs	• Access to promoters is provided by loosening a DNA molecule from histones. • Activator and repressor proteins bind to the promoter and enhance or decrease the rate of transcription. • Methyl groups are added to cytosine bases in the promoter. RNA polymerase cannot bind and transcribe.
post-transcriptional	controls the availability of mRNA molecules to ribosomes; pre-mRNA molecules undergo changes in the nucleus, resulting in final mRNA before translation occurs	• Alternative splicing occurs. Different combinations of introns are removed, and the remaining exons are spliced together. • Masking proteins bind to mRNA and inhibit further processing. • The rate of degradation of mRNA is dependent on the need of the cell for the gene product.
translational	controls how often and how rapidly mRNA transcripts will be translated into proteins	• Variation of the length of the poly(A) tail is related to the rate of translation.
post-translational	controls when proteins become fully functional, how long they are functional, and when they are degraded	• Processing occurs. The polypeptide is chemically modified to render it an active protein. • The presence of hormones may lengthen or shorten the length of time that a protein is functional. • Ubiquitin-tagged proteins are degraded.

Cancer

Cancer cells lack the regulatory mechanisms that keep healthy cells under control. The constant lengthening of telomeres is one way that cancer cells can grow out of control (Sections 6.5 and 6.6). This phenomenon (like all aspects of a cell's function) is controlled by gene expression. Some mutations in a cell's genome can have negative or even deadly effects on the organism. The probability that a given sequence of DNA has experienced a mutation increases over the lifespan of an organism. Every exposure to a potential mutagen (such as radiation or smoke), no matter how small, has a cumulative effect on the number of mutations in the genome. This is why many forms of cancer occur most often during old age, after the genome has accumulated mutations that can affect cellular functions. A mutation in one cell is passed on only to that cell's daughter cells, not to any other cells in the organism.

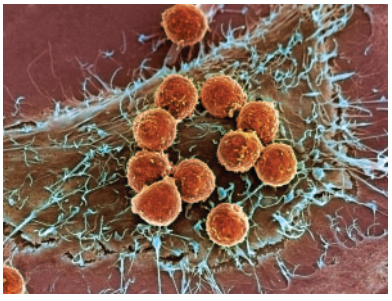


Figure 6 White blood cells attacking a cancerous cell in the human body

If a cell begins to deviate from normal cell division, it can produce a mass of undifferentiated cells called a tumour. If this mass of cells grows slowly, remains in place, and does not return once it has been removed, it is called a benign tumour. Benign tumours are usually not life-threatening. If the cells grow uncontrollably, invade surrounding tissue, and begin to affect the functions of the organism, they are called malignant tumours or cancers (**Figure 6**). Malignant tumours are more difficult to remove from the body, and measures such as chemotherapy and radiation are required.

The unchecked growth and indefinite lifespan of cancer cells are the result of changes in gene regulation. The usual mechanisms and signals that allow healthy cells to express their genes properly have little to no effect on cancerous cells, and this is reflected in certain specific mutations that arise in most tumours. Healthy cells contain a set of genes that code for various proteins that stimulate cell growth. In cancerous cells, these genes are mutated to become oncogenes. Oncogenes cause the constant and undifferentiated cell division that creates tumours.

Changes in gene regulation can arise from mutations in the promoters, mutations in the coding regions that affect the functions of the protein, or the introduction of foreign DNA from viruses (Section 7.7). The mutations that give rise to cancer cells may seem like exceptions, but it is just as exceptional that most of an organism's genes are regulated almost perfectly throughout its entire life.

7.4 Review

Summary

- A cell responds to changes in its environment by regulating the rate at which its genes are expressed.
- Prokaryotes use operons to regulate gene expression. An operon is a section of DNA that includes a promoter, an operator, a regulatory protein, and the sequence of genes that code for one or more specific proteins.
- The *lac* operon uses a signal molecule (lactose) that induces the expression of the operon's genes, whereas the *trp* operon uses a signal molecule (tryptophan) that represses the expression of the operon's genes.
- Gene regulation in eukaryotes can occur during transcription, post-transcription, translation, or post-translation.
- Transcriptional control regulates which genes are transcribed and/or the rate at which transcription occurs.
- Forms of post-transcriptional control include alternative splicing of exons and introns, masking proteins binding to mRNA, and the rate of degradation of mRNA.
- Translational control involves how often and how rapidly an mRNA is translated.
- Post-translational control regulates when proteins become fully functional, how long they are functional, and when they are degraded.
- Cancerous cells lack the regulatory mechanisms that allow healthy cells to grow and express their genes properly.

Questions

1. Define the following terms:
   (a) operon
   (b) operator
   (c) corepressor
   (d) repressor
   (e) housekeeping genes
   (f) inducer K/U
2. Why do eukaryotes have a more complex system of gene regulation than prokaryotes? Use an example to explain your reasoning. T/I
3. What features do the *lac* and *trp* operons have in common? How do these operons differ? T/I
4. How does the *lac* operon regulate the production of the enzymes needed to metabolize lactose? Summarize the key ideas and concepts in list format. K/U C
5. Describe the *trp* operon system if the level of tryptophan is low. Include the activities and states of all the major enzymes and proteins found in the system. What changes take place if tryptophan is suddenly made available as a nutrient to the bacteria? K/U
6. Give an example of a eukaryotic regulatory mechanism that occurs
   (a) during transcription
   (b) during translation
   (c) after translation K/U
7. Eukaryotic transcription is generally controlled by the binding of regulatory proteins to DNA sequences, rather than by the modification of RNA polymerases. Develop a hypothesis that explains the reason for this. T/I
8. The twin mice in Figure 5 (page 336) have one small but important difference in their agouti gene. What is the relationship between the phenotype of the agouti mouse (left in the photograph) and chemicals such as bisphenol-A? T/I
9. Discuss what happens to cells to cause them to become cancerous. K/U