

17 Control of Gene Expression in Eukaryotes



Anopheles mosquitoes transmit *Plasmodium falciparum*, one of the protozoan parasites that causes malaria, to humans. The parasite survives in humans by altering the expression of its surface proteins through gene regulation, thereby evading the human immune defenses. [Dr. Dennis Kunkel/Visuals Unlimited.]

HOW A PARASITE CHANGES ITS SPOTS

Malaria is one of the most pervasive and destructive of all infectious diseases, infecting about 300 million people worldwide and killing upward of a million people every year. The causative agents of malaria are protozoan parasites of the genus *Plasmodium*; the most deadly forms of the disease are caused by *P. falciparum* and *P. vivax*, which are widespread in many tropical regions of the world. These parasites are transmitted from person to person with deadly precision by mosquitoes. In humans, *Plasmodium* takes up residence in red blood cells, where they produce proteins (antigens) that appear on the surfaces of the host cells (Figure 17.1).

Humans are blessed with a powerful immune system, which is capable of detecting foreign antigens and destroying cells that carry them. So why does a malaria parasite survive in its human host, reproduce, and eventually cause malaria? The malaria parasite survives because it continually alters the expression of its surface proteins in a process called antigenic variation and therefore avoids destruction by the immune system.

The surface antigens expressed by *Plasmodium* are encoded by a family of genes that exist in multiple copies. In *P. falciparum*, the antigens are encoded by the *var* gene family, which consists of approximately 60 genes that vary slightly in DNA sequence and encode different surface antigens. In an amazing feat of deception, only one gene of the family is expressed at any one time, and the gene that is expressed changes continuously. As the immune system detects and mounts an attack on the surface antigens of parasitized cells, the antigen that is expressed changes; so that the parasite stays one step ahead of the immune system and is able to survive within the host.

The key to *Plasmodium*'s ability to avoid immune surveillance is gene regulation, the selective and changing expression of its *var* genes. As will be discussed in this chapter, gene regulation in eukaryotic cells can potentially be accomplished by a number of different mechanisms; *Plasmodium* regulates its *var* genes by altering chromatin structure.

The *var* genes of *Plasmodium* are located at the ends of its chromosomes, near the telomeres. These genes are normally silenced in *Plasmodium*, because chromatin in the telomeric regions is highly condensed, with the DNA tightly bound to histone proteins and inaccessible to proteins required for transcription. In 2006, researchers at the Pasteur Institute in Paris showed that individual *var* genes are expressed when chromatin structure



17.1 *Plasmodium falciparum* infects human red blood cells. A cross section through an infected human red blood cell, showing single-celled malaria parasites (green). [Omikron/Photo Researchers.]

is disrupted by chemical changes in the histone proteins. Enzymes that modify the histones are recruited to the promoter of a specific gene, where they bring about chemical changes in histone proteins, loosening their grip on DNA and providing access to transcription factors, RNA polymerase, and other proteins required for transcription. The mystery of how gene expression is transferred from one *var* gene to another is not yet known but may entail some type of physical interaction between genes.

This chapter is about gene regulation in eukaryotic cells, which typically takes place at multiple levels. We begin by considering how gene expression is influenced by changes in chromatin structure, which can be altered by several different mechanisms, including the modification of histones as used by *Plasmodium* in controlling the expression of its *var* genes. We next consider the initiation of transcription in eukaryotes, which is controlled by a complex interaction of several types of proteins and the DNA regulatory elements to which they bind. We examine several ways in which gene expression is controlled through the processing, degradation, and translation of mRNA. We end by revisiting some of the similarities in gene regulation in bacteria and eukaryotes.

17.1 Eukaryotic Cells and Bacteria Have Many Features of Gene Regulation in Common, but They Differ in Several Important Ways

As discussed in Chapter 16, many features of gene regulation are common to both bacterial and eukaryotic cells. For example, in both types of cells, DNA-binding proteins influence the ability of RNA polymerase to initiate transcription. However, there are also some differences. First, most eukaryotic genes are not organized into operons and are rarely transcribed together into a single mRNA molecule, although some operon-like gene clusters have been discovered in eukaryotes. In eukaryotic cells, each structural gene typically has its own promoter and is transcribed separately. Second, chromatin structure affects gene expression in eukaryotic cells; DNA must unwind from the histone proteins before transcription can take place. Third, the presence of the nuclear membrane in eukaryotic cells separates transcription and translation in time and space. Therefore, the regulation of gene expression in eukaryotic cells is characterized by a greater diversity of mechanisms that act at different points in the transfer of information from DNA to protein.

Eukaryotic gene regulation is less well understood than bacterial regulation, partly owing to the larger genomes in

eukaryotes, their greater sequence complexity, and the difficulty of isolating and manipulating mutations that can be used in the study of gene regulation. Nevertheless, great advances in our understanding of the regulation of eukaryotic genes have been made in recent years, and eukaryotic regulation continues to be a cutting-edge area of research in genetics.

17.2 Changes in Chromatin Structure Affect the Expression of Genes

One type of gene control in eukaryotic cells is accomplished through the modification of chromatin structure. In the nucleus, histone proteins associate to form octamers, around which helical DNA tightly coils to create chromatin (see Figure 11.4). In a general sense, this chromatin structure represses gene expression. For a gene to be transcribed, transcription factors, activators, and RNA polymerase must bind to the DNA. How can these events take place with DNA wrapped tightly around histone proteins? The answer is that, before transcription, chromatin structure changes and the DNA becomes more accessible to the transcriptional machinery.

DNase I Hypersensitivity

Several types of changes are observed in chromatin structure when genes become transcriptionally *active*. As genes become transcriptionally active, regions around the genes become highly sensitive to the action of DNase I (see Chapter 11). These regions, called **DNase I hypersensitive sites**, frequently develop about 1000 nucleotides upstream of the start site of transcription, suggesting that the chromatin in these regions adopts a more open configuration during transcription. This relaxation of the chromatin structure allows regulatory proteins access to binding sites on the DNA. Indeed, many DNase I hypersensitive sites correspond to known binding sites for regulatory proteins. At least three different processes affect gene regulation by altering chromatin structure: (1) the modification of histone proteins;

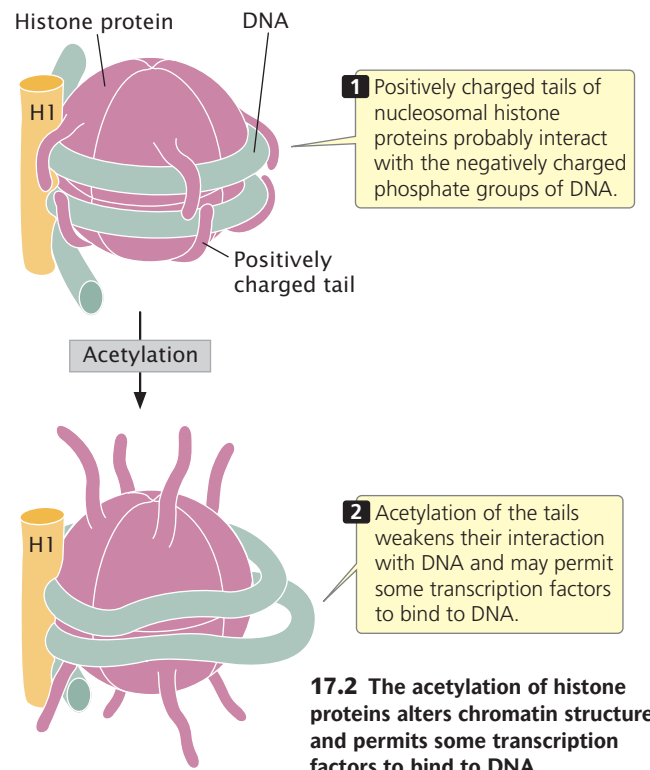
(2) chromatin remodeling; and (3) DNA methylation. Each of these mechanisms will be discussed in the sections that follow.

Histone Modification

Histones in the octamer core of the nucleosome have two domains: (1) a globular domain that associates with other histones and the DNA and (2) a positively charged tail domain that probably interacts with the negatively charged phosphate groups on the backbone of DNA. The tails of histone proteins are often modified by the addition or removal of phosphate groups, methyl groups, or acetyl groups. These modifications have sometimes been called the **histone code**, because they encode information that affects how genes are expressed.

Methylation of histones One type of histone modification is the addition of methyl groups to the tails of histone proteins. These modifications can bring about either the activation or the repression of transcription, depending on which particular amino acids in the histone tail are methylated. A common modification is the addition of three methyl groups to lysine 4 in the tail of the H3 histone protein, abbreviated H3K4me3 (K is the abbreviation for lysine). The H3K4me3 modification is frequently found in promoters of transcriptionally active genes in eukaryotes. Recent studies have identified proteins that recognize and bind to H3K4me3, including a protein called nucleosome remodeling factor (NURF). NURF and other proteins that recognize H3K4me3 have a common protein-binding domain that binds to the H3 histone tail and then alters chromatin packing, allowing transcription to take place. Research has also demonstrated that some transcription factors, which are necessary for the initiation of transcription (see Chapter 14 and Section 17.4), directly bind to H3K4me3.

Acetylation of histones Another type of histone modification that affects chromatin structure is acetylation, the addition of acetyl groups (CH_3CO) to histone proteins (**Figure 17.2**). The acetylation of histones usually stimulates transcription. For example, the addition of a single acetyl group to lysine 16 in the tail of the H4 histone prevents the formation of the 30-nm chromatin fiber (see Figure 11.4), causing the chromatin to be in an open configuration and available for transcription. In general, acetyl groups destabilize chromatin structure, allowing transcription to take place. Acetyl groups are added to histone proteins by acetyltransferase enzymes; other enzymes called deacetylases strip acetyl groups from histones and restore chromatin structure, which represses transcription. Certain transcription factors (see Chapter 13) and other proteins that regulate transcription either have acetyltransferase activity or attract acetyltransferases to DNA.

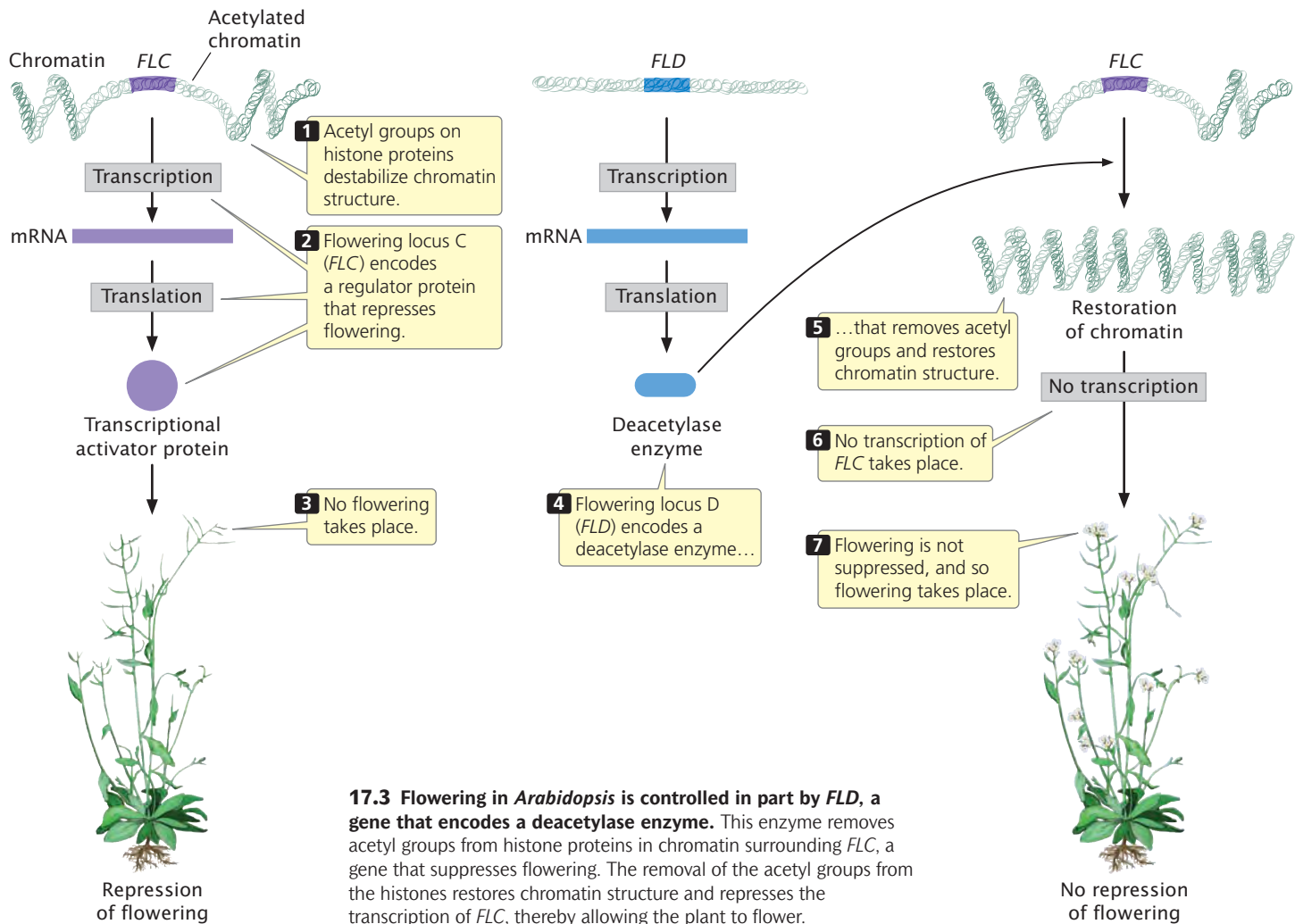


17.2 The acetylation of histone proteins alters chromatin structure and permits some transcription factors to bind to DNA.

Acetylation of histones controls flowering in *Arabidopsis* The importance of histone acetylation in gene regulation is demonstrated by the control of flowering in *Arabidopsis*, a plant with a number of characteristics that make it an excellent genetic model for plant systems (see material on the model genetic organism *Arabidopsis thaliana* in the Reference Guide to Model Genetic Organisms at the end of the book). The time at which flowering takes place is critical to the life of a plant; if flowering is initiated at the wrong time of year, pollinators may not be available to fertilize the flowers or environmental conditions may be unsuitable for the survival and germination of the seeds. Consequently, flowering time in most plants is carefully regulated in response to multiple internal and external cues, such as plant size, photoperiod, and temperature.

Among the many genes that control flowering in *Arabidopsis* is *flowering locus C* (*FLC*), which plays an important role in suppressing flowering until after an extended period of coldness (a process called vernalization). The *FLC* gene encodes a regulator protein that represses the activity of other genes that affect flowering (**Figure 17.3**). As long as *FLC* is active, flowering remains suppressed. The activity of *FLC* is controlled by another locus called *flowering locus D* (*FLD*), the key role of which is to stimulate flowering by repressing the action of *FLC*. In essence, flowering is stimulated because *FLD* represses the repressor. How does *FLD* repress *FLC*?

FLD encodes a deacetylase enzyme, which removes acetyl groups from histone proteins in the chromatin surrounding *FLC* (see Figure 17.3). The removal of acetyl groups from



histones alters chromatin structure and inhibits transcription. The inhibition of transcription prevents *FLC* from being transcribed and removes its repression on flowering. In short, *FLD* stimulates flowering in *Arabidopsis* by deacetylating the chromatin that surrounds *FLC*, thereby removing its inhibitory effect on flowering. [TRY PROBLEM 19](#) →

Chromatin Remodeling

The changes to chromatin structure discussed so far have been through alteration of the structure of the histone proteins. Some transcription factors and other regulatory proteins alter chromatin structure without altering the chemical structure of the histones directly. These proteins are called **chromatin-remodeling complexes**. They bind directly to particular sites on DNA and reposition the nucleosomes, allowing transcription factors to bind to promoters and initiate transcription.

One of the best-studied examples of a chromatin-remodeling complex is SWI–SNF, which is found in yeast, humans, *Drosophila*, and other organisms. This complex uti-

lizes energy derived from the hydrolysis of ATP to reposition nucleosomes, exposing promoters in the DNA to the action of other regulatory proteins and RNA polymerase. Research has shown that SWI–SNF is able to change the position of nucleosomes along a DNA molecule.

Evidence suggests at least two mechanisms by which remodeling complexes reposition nucleosomes. First, some remodeling complexes cause the nucleosome to slide along the DNA, allowing DNA that was wrapped around the nucleosome to occupy a position in between nucleosomes, where it is more accessible to proteins affecting gene expression. Second, some complexes cause conformational changes in the DNA, in nucleosomes, or in both so that DNA that is bound to the nucleosome assumes a more exposed configuration.

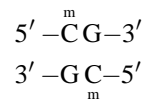
Chromatin-remodeling complexes are targeted to specific DNA sequences by transcriptional activators or repressors that attach to a remodeling complex and then bind to the promoters of specific genes. There is also evidence that chromatin-remodeling complexes work together with enzymes that alter histones, such as acetyltransferase

enzymes, to change chromatin structure and expose DNA for transcription.

DNA Methylation

Another change in chromatin structure associated with transcription is the methylation of cytosine bases, which yields 5-methylcytosine (see Figure 10.18). The methylation of cytosine in DNA is distinct from the methylation of histone proteins mentioned earlier. Heavily methylated DNA is associated with the repression of transcription in vertebrates and plants, whereas transcriptionally active DNA is usually unmethylated in these organisms. Abnormal patterns of methylation are also associated with some types of cancer.

DNA methylation is most common on cytosine bases adjacent to guanine nucleotides (CpG, where p represents the phosphate group in the DNA backbone); so two methylated cytosines sit diagonally across from each other on opposing strands:



DNA regions with many CpG sequences are called **CpG islands** and are commonly found near transcription start sites. While genes are not being transcribed, these CpG islands are often methylated, but the methyl groups are removed before the initiation of transcription. CpG methylation is also associated with long-term gene repression, such as on the inactivated X chromosome of female mammals (see Chapter 4).

Evidence indicates that an association exists between DNA methylation and the deacetylation of histones, both of which repress transcription. Certain proteins that bind tightly to methylated CpG sequences form complexes with other proteins that act as histone deacetylases. In other words, methylation appears to attract deacetylases, which remove acetyl groups from the histone tails, stabilizing the nucleosome structure and repressing transcription. The demethylation of DNA allows acetyltransferases to add acetyl groups, disrupting nucleosome structure and permitting transcription. **TRY PROBLEM 20** →

CONCEPTS

Sensitivity to DNase I digestion indicates that transcribed DNA assumes an open configuration before transcription. Chromatin structure can be altered by modifications of histone proteins, by chromatin-remodeling complexes that reposition nucleosomes, and by the methylation of DNA.

✓ CONCEPT CHECK 1

What are some of different processes that affect gene regulation by altering chromatin structure?

17.3 Epigenetic Effects Often Result from Alterations in Chromatin Structure

Changes in chromatin structure that affect gene expression, such as the mechanisms just described, are examples of the phenomena of epigenetics. In Chapters 5 and 11, we considered the concept of epigenetics, which was defined as alterations to DNA and chromatin structure that affect traits and are passed on to other cells or future generations but are not caused by changes in the DNA base sequence.

In the past, the term epigenetics was applied to various genetic phenomena that could not be easily explained by traditional Mendelian principles. For example, in the 1950s, Alexander Brink described **paramutation** in corn, in which one allele of a genotype altered the expression of another allele. In corn, alleles at the *R* locus help to determine pigmentation; the *R^f* allele is normally dominant and encodes purple kernels. The *Rst* allele encodes spotted kernels. Brink observed that, when the *R^f* allele was present in a genotype with *Rst*, the effect of the *R^f* allele was altered so that, in later generations, it encoded reduced levels of purple pigment. In this example of paramutation, the *Rst* allele altered the expression of the *R^f* allele; this diminished effect of the *R^f* allele on pigmentation persisted for several generations, even in the absence of the *Rst* allele. Paramutation violates Mendel's principle of segregation, which states that, when gametes are formed, each allele separates and is transmitted independently to the next generation; paramutation violates this rule because the effect of the *R^f* allele remains associated with the *Rst* allele even after the two alleles have separated. Epigenetics has been called "inheritance, but not as we know it."

In recent years, scientists have discovered that many of these types of phenomena result from changes in chromatin structure that affect the expression of genetic information. Epigenetic changes are heritable in the sense that they are passed from cell to cell and sometimes across generations, but they differ from genetic traits in that they are not encoded by the DNA base sequence and are reversible.

Epigenetic Effects

A number of examples of epigenetic effects have been documented. In some cases, the molecular basis of these effects is at least partly understood.

Epigenetic changes induced by maternal behavior

A fascinating example of epigenetics is seen in the long-lasting effects of maternal behavior in rats. Mother rats lick and groom their offspring, usually while the mother arches her back and nurses the offspring. The offspring of mothers who display more licking and grooming behavior are, as adults, less fearful and show reduced hormonal responses to stress compared with the offspring of mothers who lick and groom less. These long-lasting differences in the offspring

are not due to genetic differences inherited from their mothers—at least not genetic differences in the base sequences of the DNA. Research has demonstrated that offspring exposed to more licking and grooming from the mother develop a different pattern of DNA methylation compared with offspring exposed to less licking and grooming. These differences in DNA methylation affect the acetylation of histone proteins (discussed earlier) that persist into adulthood and alter the expression of the glucocorticoid receptor gene, a gene that plays a role in hormonal responses to stress. The expression of other stress-response genes also is affected.

To demonstrate the effect of altered chromatin structure on the stress response of the offspring, researchers infused the brains of young rats with a deacetylase inhibitor, which prevents the removal of acetyl groups from the histone proteins. After infusion of the deacetylase inhibitor, differences in DNA methylation and histone acetylation associated with grooming behavior disappeared, as did the difference in responses to fear and stress in the adults. These results demonstrate that the mother rat's behavior (licking and grooming) brings about epigenetic changes in chromatin of the offspring, which cause long-lasting differences in their behavior.

Epigenetic effects caused by prenatal exposure In another study, researchers found that the exposure of embryonic rats to the fungicide vinclozolin, which reduces sperm production, led to reduced sperm production not only in the treated animals (when they reached puberty), but also in several subsequent generations. Increased DNA methylation was seen in sperm of the males that were exposed to vinclozolin and these patterns of methylation were inherited. This study and others have raised concerns that, through epigenetic changes, environmental exposure to some chemicals might have long-lasting effects on the health of future generations.

Epigenetic effects in monozygotic twins Monozygotic (identical) twins develop from a single egg fertilized by a single sperm that divides and gives rise to two zygotes (see Chapter 6). Monozygotic twins are genetically identical, in the sense that they possess identical DNA sequences, but they often differ somewhat in appearance, health, and behavior. The nature of these differences in the phenotypes of identical twins is poorly understood, but recent evidence suggests that at least some of these differences may be due to epigenetic changes. In one study, Mario Fraga and his colleagues examined 80 identical twins and compared the degree and location of their DNA methylation and histone acetylation. They found that DNA methylation and histone acetylation of the two members of an identical twin pair were similar early in life, but older twins had remarkable differences in their overall content and distribution of DNA methylation and histone acetylation. Furthermore, these dif-

ferences affected gene expression in the twins. This research suggests that identical twins do differ epigenetically and that phenotypic differences between them may be caused by differential gene expression. Additional examples of epigenetic effects include genomic imprinting (discussed in Chapter 5) and X-chromosome inactivation (discussed in Chapter 4).

Molecular Mechanisms of Epigenetic Changes

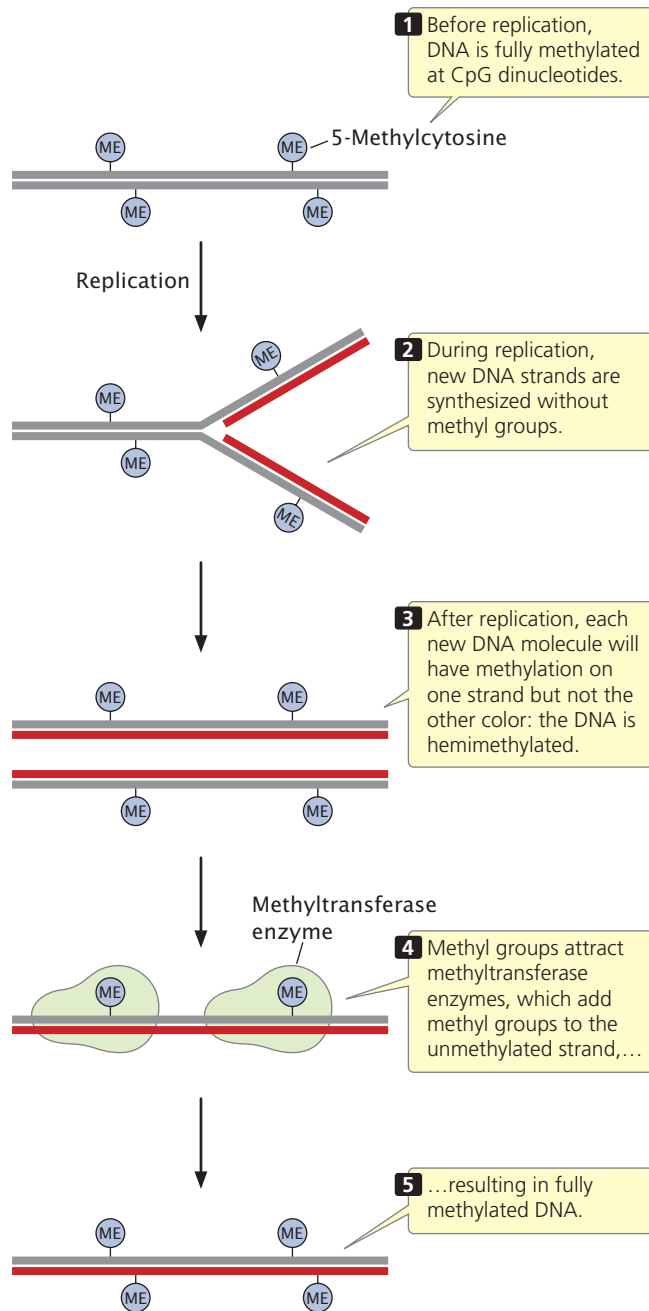
Epigenetics alters gene expression, alteration that is stable enough to be transmitted through mitosis (and sometimes meiosis) but that can also be changed. Most evidence suggests that epigenetic effects are brought about by physical changes in chromatin structure. Earlier, we considered a number of chemical changes in DNA and histone proteins that potentially alter chromatin structure, including DNA methylation, modification of histone proteins, and repositioning of nucleosomes. There is also growing evidence that small RNA molecules (siRNAs, miRNAs, and piRNAs; see Section 17.6) take part in some types of epigenetic control of gene expression.

The fact that epigenetic marks are passed on to other cells and (sometimes) to future generations means that changes in chromatin structure associated with epigenetic phenotypes must be faithfully maintained when chromosomes replicate. How are epigenetic changes maintained through the process of cell division? Evidence suggests that DNA methylation is often stably maintained through replication. Recall that cytosine bases of CpG sequences are often methylated, which means that two methylated cytosine bases sit diagonally across from each other on opposing strands (see section on DNA methylation). Before replication, cytosine bases on both strands are methylated (**Figure 17.4**). Immediately after semiconservative replication, the cytosine base on one strand (the template strand) will be methylated, but the cytosine base on the other strand (the newly replicated strand) will be unmethylated (**Figure 17.4**). Special methyltransferase enzymes recognize the hemimethylated state of CpG dinucleotides and add methyl groups to the unmethylated cytosine bases, creating two new DNA molecules that are fully methylated.

How histone modifications, nucleosome positioning, and other types of epigenetic marks might be maintained across replication is less clear. One possibility, discussed in Chapter 12, is that, after replication, the epigenetic marks remain on the original histones, and these marks recruit enzymes that make similar changes to the new histones.

The Epigenome

The overall pattern of chromatin modifications possessed by each individual organism has been termed the **epigenome**. The epigenome helps to explain stable patterns of gene expression. For example, all cells in the human body are



17.4 DNA methylation is stably maintained through DNA replication.

genetically identical, and yet different cell types exhibit remarkably different phenotypes—a nerve cell is quite distinct in its shape, size, and function from an intestinal cell. These differences in phenotypes are stable and passed from one cell to another, in spite of the fact that the DNA sequences of all the cells are identical.

Embryonic stem cells are undifferentiated cells that are capable of forming every type of cell in an organism, a property referred to as **pluripotency**. As a stem cell divides and gives rise to a more specialized type of cell, the gene-expression program becomes progressively fixed so that each particular cell type expresses only those genes necessary to carry out the functions of that cell type. Although the control of these cell-specific expression programs is still poorly understood, changes in DNA methylation and chromatin structure clearly play important roles in silencing some genes and activating others.

Recently, researchers have determined the genome-wide distribution of DNA methylation. In mammals, most methylation is at cytosine bases, converting cytosine into 5-methylcytosine. The researchers identified the location of 5-methylcytosine across the entire genome in two cell types: (1) an undifferentiated human stem cell; and (2) a fibroblast, a fully differentiated connective tissue cell. The researchers found widespread differences in the methylation patterns of these two types of cells. The finding that patterns of DNA methylation vary among cell types has led to speculation that cytosine methylation may provide the means by which cell types stably maintain their differences during development.

CONCEPTS

Epigenetic effects—stable alterations to gene expression that are not caused by changes in the DNA base sequence—are frequently associated with changes in chromatin structure. The epigenome is the genomewide pattern of epigenetic changes possessed by an individual organism.

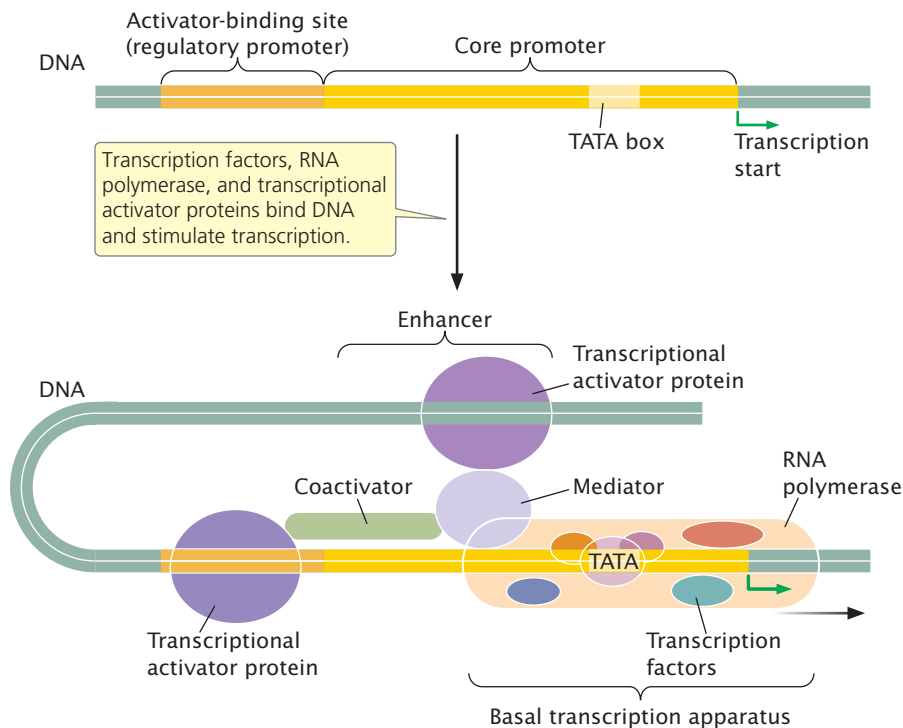
✓ CONCEPT CHECK 2

Which is *not* a mechanism of epigenetic change?

- DNA methylation
- Alteration of a DNA base sequence in a promoter
- Histone acetylation
- Nucleosome repositioning

17.4 The Initiation of Transcription Is Regulated by Transcription Factors and Transcriptional Regulator Proteins

We just considered one level at which gene expression is controlled—the alteration of chromatin and DNA structure. We now turn to another important level of control—that is, control through the binding of proteins to DNA sequences that affect transcription. Transcription is an important level of control in eukaryotic cells, and this control requires



17.5 Transcriptional activator proteins bind to sites on DNA and stimulate transcription. Most act by stimulating or stabilizing the assembly of the basal transcription apparatus.

a number of different types of proteins and regulatory elements. The initiation of eukaryotic transcription was discussed in detail in Chapter 13. Recall that general transcription factors and RNA polymerase assemble into a *basal transcription apparatus*, which binds to a *core promoter* located immediately upstream of a gene. The basal transcription apparatus is capable of minimal levels of transcription; *transcriptional regulator proteins* are required to bring about normal levels of transcription. These proteins bind to a regulatory promoter, which is located upstream of the core promoter (Figure 17.5), and to *enhancers*, which may be located some distance from the gene. Some transcriptional regulator proteins are activators, stimulating transcription. Others are repressors, inhibiting transcription.

Transcriptional Activators and Coactivators

Transcriptional activator proteins stimulate and stabilize the basal transcription apparatus at the core promoter. The activators may interact directly with the basal transcription apparatus or indirectly through protein coactivators. Some activators and coactivators, as well as the general transcription factors, also have acetyltransferase activity and so further stimulate transcription by altering chromatin structure.

Transcriptional activator proteins have two distinct functions (see Figure 17.5). First, they are capable of binding DNA at a specific base sequence, usually a consensus sequence in a regulatory promoter or enhancer; for this function, most transcriptional activator proteins contain one or more DNA-binding motifs, such as the

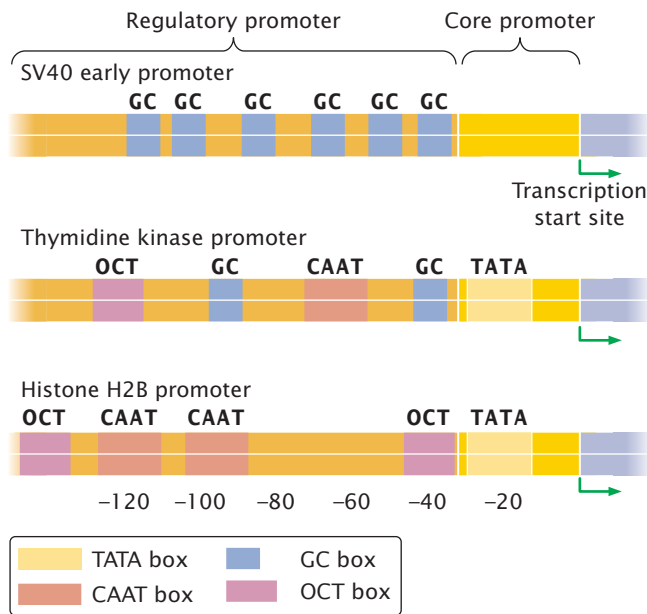
helix-turn-helix, zinc finger, and leucine zipper (see Chapter 16). A second function is the ability to interact with other components of the transcriptional apparatus and influence the rate of transcription.

Within the regulatory promoter are typically several different consensus sequences to which different transcriptional activators can bind. Among different promoters, activator-binding sites are mixed and matched in different combinations (Figure 17.6), and so each promoter is regulated by a unique combination of transcriptional activator proteins.

Transcriptional activator proteins bind to the consensus sequences in the regulatory promoter and affect the assembly or stability of the basal transcription apparatus at the core promoter. One of the components of the basal transcription apparatus is a complex of proteins called the **mediator** (see Figure 17.5). Transcriptional activator proteins binding to sequences in the regulatory promoter (or enhancer, see next section) make contact with the mediator and affect the rate at which transcription is initiated. Some regulatory promoters also contain sequences that are bound by proteins that lower the rate of transcription through inhibitory interactions with the mediator.

Regulation of galactose metabolism through GAL4

An example of a transcriptional activator protein is GAL4, which regulates the transcription of several yeast genes whose products metabolize galactose. Like the genes in the *lac* operon, the genes that control galactose metabolism are inducible: when galactose is absent, these genes are not



17.6 The consensus sequences in the promoters of three eukaryotic genes illustrate the principle that different sequences can be mixed and matched in different combinations. A different transcriptional activator protein binds to each consensus sequence, and so each promoter responds to a unique combination of activator proteins.

transcribed and the proteins that break down galactose are not produced; when galactose is present, the genes are transcribed and the enzymes are synthesized. GAL4 contains several zinc fingers and binds to a DNA sequence called UAS_G (upstream activating sequence for GAL4). UAS_G exhibits the

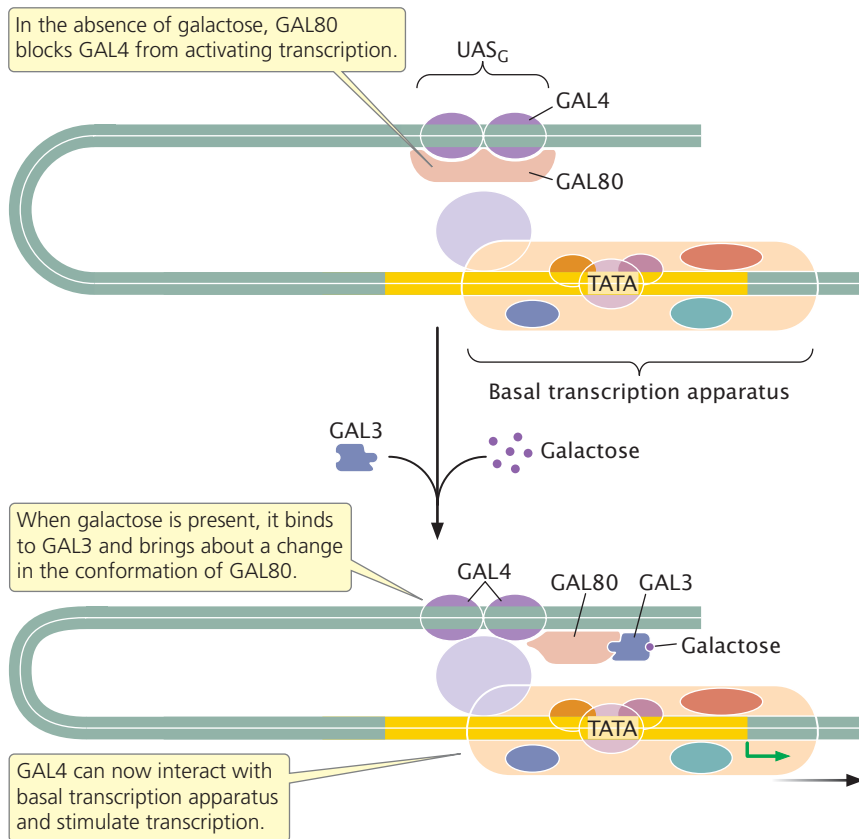
properties of an enhancer—a regulatory sequence that may be some distance from the regulated gene and is independent of the gene in position and orientation (see Chapter 13). When bound to UAS_G, GAL4 activates the transcription of yeast genes needed for metabolizing galactose.

A particular region of GAL4 binds another protein called GAL80, which regulates the activity of GAL4 in the presence of galactose. When galactose is absent, GAL80 binds to GAL4, preventing GAL4 from activating transcription (Figure 17.7). When galactose is present, however, it binds to another protein called GAL3, which interacts with GAL80, causing a conformational change in GAL80 so that it can no longer bind GAL4. The GAL4 protein is then free to activate the transcription of the genes, whose products metabolize galactose.

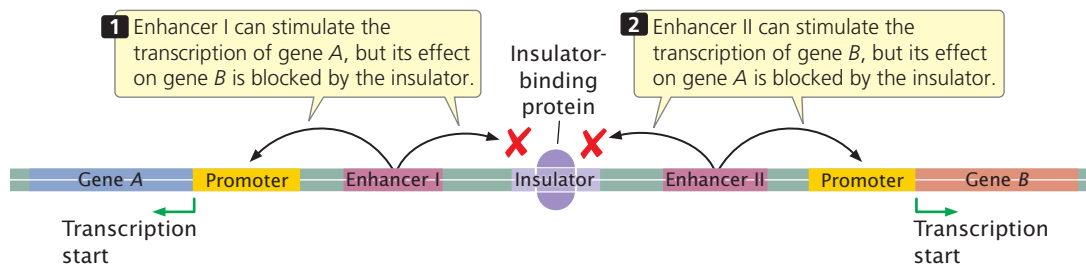
GAL4 and a number of other transcriptional activator proteins contain multiple amino acids with negative charges that form an *acidic activation domain*. These acidic activators stimulate transcription by enhancing the assembly of the basal transcription apparatus.

Transcriptional Repressors

Some regulatory proteins in eukaryotic cells act as repressors, inhibiting transcription. These repressors bind to sequences in the regulatory promoter or to distant sequences called *silencers*, which, like enhancers, are position and orientation



17.7 Transcription is activated by GAL4 in response to galactose. GAL4 binds to the UAS_G site and controls the transcription of genes in galactose metabolism.



17.8 An insulator blocks the action of an enhancer on a promoter when the insulator lies between the enhancer and the promoter.

independent. Unlike repressors in bacteria, most eukaryotic repressors do not directly block RNA polymerase. These repressors may compete with activators for DNA binding sites: when a site is occupied by an activator, transcription is activated, but, if a repressor occupies that site, there is no activation. Alternatively, a repressor may bind to sites near an activator site and prevent the activator from contacting the basal transcription apparatus. A third possible mechanism of repressor action is direct interference with the assembly of the basal transcription apparatus, thereby blocking the initiation of transcription.

CONCEPTS

Transcriptional regulatory proteins in eukaryotic cells can influence the initiation of transcription by affecting the stability or assembly of the basal transcription apparatus. Some regulatory proteins are activators and stimulate transcription; others are repressors and inhibit initiation of transcription.

✓ CONCEPT CHECK 3

Most transcriptional activator proteins affect transcription by interacting with

- a. introns.
- b. the basal transcription apparatus.
- c. DNA polymerase.
- d. nucleosomes.

Enhancers and Insulators

Enhancers are capable of affecting transcription at distant promoters. For example, an enhancer that regulates the gene encoding the alpha chain of the T-cell receptor is located 69,000 bp downstream of the gene's promoter. Furthermore, the exact position and orientation of an enhancer relative to the promoter can vary. How can an enhancer affect the initiation of transcription taking place at a promoter that is tens of thousands of base pairs away? In many cases, regulator proteins bind to the enhancer and cause the DNA between the enhancer and the promoter to loop out, bringing the promoter and enhancer close to each other, and so the transcriptional regulator proteins are able to directly interact with the basal transcription apparatus at the core promoter (see Figure 17.5). Some enhancers may be attracted to promoters by proteins that bind to sequences in the regulatory promoter and “tether” the enhancer close to the core promoter. A typical enhancer is some 500 bp in length and contains 10 binding sites for proteins that regulate transcription.

Most enhancers are capable of stimulating any promoter in their vicinities. Their effects are limited, however, by **insulators** (also called boundary elements), which are DNA sequences that block or insulate the effect of enhancers in a position-dependent manner. If the insulator lies between the enhancer and the promoter, it blocks the action of the enhancer; but, if the insulator lies outside the region between the two, it has no effect (**Figure 17.8**). Specific proteins bind to insulators and play a role in their blocking activity. Some insulators also limit the spread of changes in chromatin structure that affect transcription. **TRY PROBLEM 24** →

CONCEPTS

Some regulatory proteins bind to enhancers, which are regulatory elements that are distant from the gene whose transcription they stimulate. Insulators are DNA sequences that block the action of enhancers.

✓ CONCEPT CHECK 4

How does the binding of regulatory proteins to enhancers affect transcription at genes that are thousands of base pairs away?

Regulation of Transcriptional Stalling and Elongation

Transcription in eukaryotes is often regulated through factors that affect the initiation of transcription, including changes in chromatin structure, transcription factors, and transcriptional regulatory proteins. Research indicates that transcription may also be controlled through factors that affect stalling and elongation of RNA polymerase after transcription has been initiated.

The basal transcription apparatus—consisting of RNA polymerase, transcription factors, and other proteins—assembles at the core promoter. When the initiation of transcription has taken place, RNA polymerase moves downstream, transcribing the structural gene and producing an RNA product. At some genes, RNA polymerase initiates transcription and transcribes from 24 to 50 nucleotides of RNA but then pauses or stalls. For example, stalling is observed at genes that encode **heat-shock proteins** in *Drosophila*—proteins that help to prevent damage from stressing agents such as extreme heat. Heat-shock proteins are produced by a large number of different genes. During

Table 17.1 A few response elements found in eukaryotic cells

Response Element	Responds to	Consensus Sequence
Heat-shock element	Heat and other stress	CNNGAANNTCCNNG
Glucocorticoid response element	Glucocorticoids	TGGTACAAATGTTCT
Phorbol ester response element	Phorbol esters	TGACTCA
Serum response element	Serum	CCATATTAGG

Source: After B. Lewin, *Genes IV* (Oxford University Press, 1994), p. 880.

times of environmental stress, the transcription of all the heat-shock genes is greatly elevated. RNA polymerase initiates transcription at heat-shock genes in *Drosophila* but, in the absence of stress, stalls downstream of the transcription initiation site. Stalled polymerases are released when stress is encountered, allowing rapid transcription of the genes and the production of heat-shock proteins that facilitate adaptation to the stressful environment.

Stalling was formerly thought to take place at only a small number of genes, but recent research indicates that stalling is widespread and common throughout eukaryotic genomes. For example, stalled RNA polymerases were found at hundreds of genes in *Drosophila*. Factors that promote stalling have been identified; a protein called negative elongation factor (NELF) causes RNA polymerase to stall after initiation. Reducing the amount of NELF in cultured cells stimulates the elongation of RNA molecules at many genes with stalled RNA polymerases.

CONCEPTS

At some genes, RNA polymerase may pause or stall downstream of the promoter. Regulatory factors affect stalling and the elongation of transcription.

Coordinated Gene Regulation

Although most eukaryotic cells do not possess operons, several eukaryotic genes may be activated by the same stimulus. Groups of bacterial genes are often coordinately expressed (turned on and off together) because they are physically clustered as an operon and have the same promoter, but coordinately expressed genes in eukaryotic cells are not clustered. How, then, is the transcription of eukaryotic genes coordinately controlled if they are not organized into an operon?

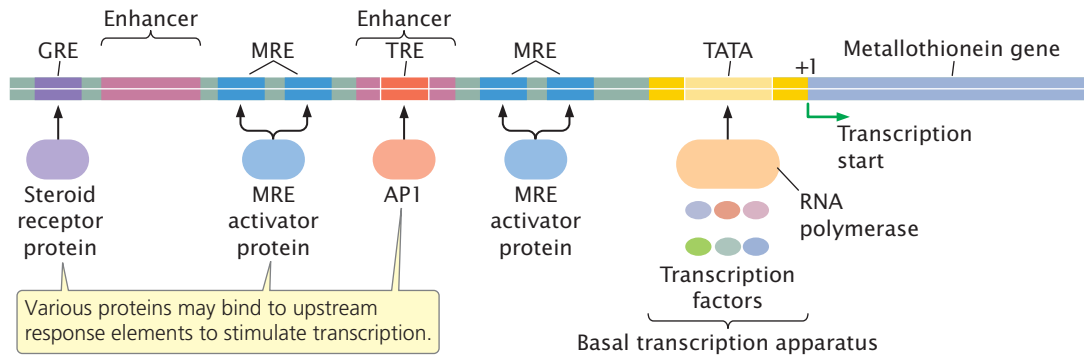
Genes that are coordinately expressed in eukaryotic cells are able to respond to the same stimulus because they have short regulatory sequences in common in their promoters or enhancers. For example, different eukaryotic heat-shock genes possess a common regulatory element upstream of their start sites. Such DNA regulatory sequences are called

response elements; they are short sequences that typically contain consensus sequences (Table 17.1) at varying distances from the gene being regulated. The response elements are binding sites for transcriptional activators. A transcriptional activator protein binds to the response element and elevates transcription. The same response element may be present in different genes, allowing multiple genes to be activated by the same stimulus.

A single eukaryotic gene may be regulated by several different response elements. For example, the metallothionein gene protects cells from the toxicity of heavy metals by encoding a protein that binds to heavy metals and removes them from cells. The basal transcription apparatus assembles around the TATA box, just upstream of the transcription start site for the metallothionein gene, but the apparatus alone is capable of only low rates of transcription.

Other response elements found upstream of the metallothionein gene contribute to increasing its rate of transcription. For example, several copies of a metal response element (MRE) are upstream of the metallothionein gene (Figure 17.9). Heavy metals stimulate the binding of activator proteins to MREs, which elevates the rate of transcription of the metallothionein gene. Because there are multiple copies of the MRE, high rates of transcription are induced by metals. Two enhancers also are located in the upstream region of the metallothionein gene; one enhancer contains a response element known as TRE, which stimulates transcription in the presence of an activated protein called AP1. A third response element called GRE is located approximately 250 nucleotides upstream of the metallothionein gene and stimulates transcription in response to certain hormones.

This example illustrates a common feature of eukaryotic transcriptional control: a single gene may be activated by several different response elements found in both promoters and enhancers. Multiple response elements allow the same gene to be activated by different stimuli. At the same time, the presence of the same response element in different genes allows a single stimulus to activate multiple genes. In this way, response elements allow complex biochemical responses in eukaryotic cells.



17.9 Multiple response elements (MREs) are found in the upstream region of the metallothionein gene. The basal transcription apparatus binds near the TATA box. In response to heavy metals, activator proteins bind to several MREs and stimulate transcription. The TRE response element is the binding site for transcription factor AP1, which is stimulated by phorbol esters. In response to glucocorticoid hormones, steroid-receptor proteins bind to the GRE response element located approximately 250 nucleotides upstream of the metallothionein gene and stimulate transcription.

17.5 Some Genes Are Regulated by RNA Processing and Degradation

In bacteria, transcription and translation take place simultaneously. In eukaryotes, transcription takes place in the nucleus and the pre-mRNAs are then processed before moving to the cytoplasm for translation, allowing opportunities for gene control after transcription. Consequently, post-transcriptional gene regulation assumes an important role in eukaryotic cells. A common level of gene regulation in eukaryotes is RNA processing and degradation.

Gene Regulation Through RNA Splicing

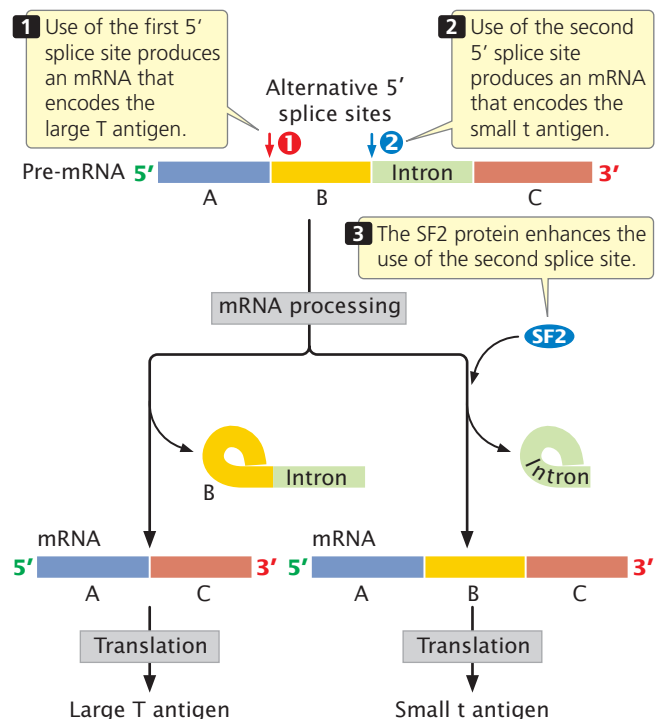
Alternative splicing allows a pre-mRNA to be spliced in multiple ways, generating different proteins in different tissues or at different times in development (see Chapter 14). Many eukaryotic genes undergo alternative splicing, and the regulation of splicing is probably an important means of controlling gene expression in eukaryotic cells.

Alternative splicing in the T-antigen gene The T-antigen gene of the mammalian virus SV40 is a well-studied example of alternative splicing. This gene is capable of encoding two different proteins, the large T and small t antigens. Which of the two proteins is produced depends on which of two alternative 5' splice sites is used in RNA splicing (Figure 17.10). The use of one 5' splice site produces mRNA that encodes the large T antigen, whereas the use of the other 5' splice site (which is farther downstream) produces an mRNA encoding the small t antigen.

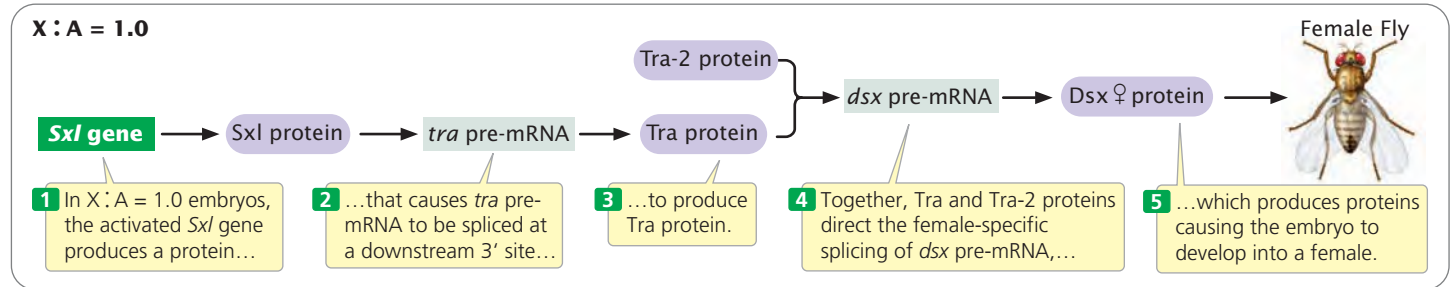
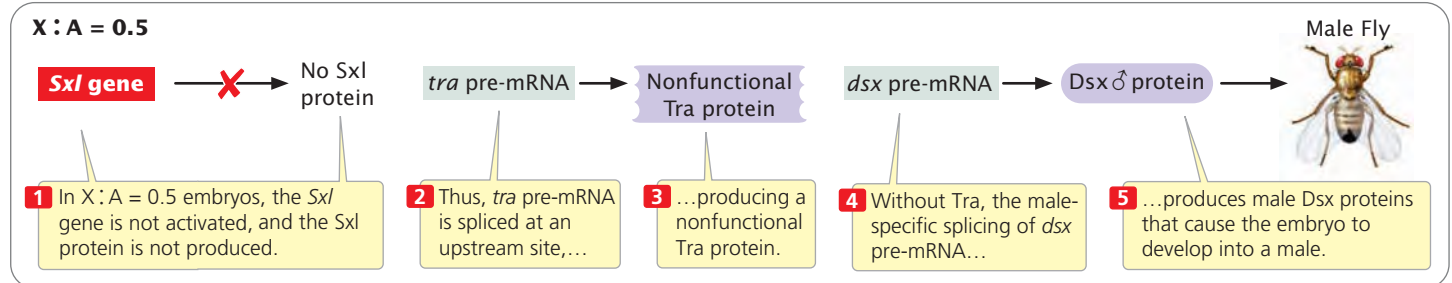
A protein called splicing factor 2 (SF2) enhances the production of mRNA encoding the small t antigen (see Figure 17.10). Splicing factor 2 has two binding domains: one domain is an RNA-binding region and the other has alternating serine and arginine amino acids. These two domains are typical of SR (serine- and arginine-rich) proteins, which often play a role in regulating splicing. Splicing factor 2

stimulates the binding of small nuclear ribonucleoproteins (snRNPs) to the 5' splice site, one of the earliest steps in RNA splicing (see Chapter 14). The precise mechanism by which SR proteins influence the choice of splice sites is poorly understood. One model suggests that SR proteins bind to specific splice sites on mRNA and stimulate the attachment of snRNPs, which then commit the site to splicing.

Alternative splicing in *Drosophila* sexual development Another example of alternative mRNA splicing that regulates the expression of genes controls whether a fruit fly



17.10 Alternative splicing leads to the production of the small t antigen and the large T antigen in the mammalian virus SV40.

XX genotype**XY genotype****17.11** Alternative splicing controls sex determination in *Drosophila*.

develops as male or female. Sex differentiation in *Drosophila* arises from a cascade of gene regulation. When the ratio of X chromosomes to the number of haploid sets of autosomes (the X : A ratio; see Chapter 4) is 1, a female-specific promoter is activated early in development and stimulates the transcription of the *sex-lethal* (*Sxl*) gene (Figure 17.11). The protein encoded by *Sxl* regulates the splicing of the pre-mRNA transcribed from another gene called *transformer* (*tra*). The splicing of *tra* pre-mRNA results in the production of the Tra protein (see Figure 17.11). Together with another protein (Tra-2), Tra stimulates the female-specific splicing of pre-mRNA from yet another gene called *double-sex* (*dsx*). This event produces a female-specific Dsx protein, which causes the embryo to develop female characteristics.

In male embryos, which have an X : A ratio of 0.5 (see Figure 17.11), the promoter that transcribes the *Sxl* gene in females is inactive; so no Sxl protein is produced. In the absence of Sxl protein, *tra* pre-mRNA is spliced at a different 3' splice site to produce a nonfunctional form of Tra protein (Figure 17.12). In turn, the presence of this nonfunctional Tra in males causes *dsx* pre-mRNAs to be spliced differently from that in females, and a male-specific Dsx protein is produced (see Figure 17.11). This event causes the development of male-specific traits.

In summary, the Tra, Tra-2, and Sxl proteins regulate alternative splicing that produces male and female phenotypes in *Drosophila*. Exactly how these proteins regulate alternative splicing is not yet known, but the Sxl protein (produced only in females) possibly blocks the upstream splice site on the *tra* pre-mRNA. This blockage would force the spliceosome to use the downstream 3' splice site, which

causes the production of Tra protein and eventually results in female traits (see Figure 17.12). **TRY PROBLEM 25** →

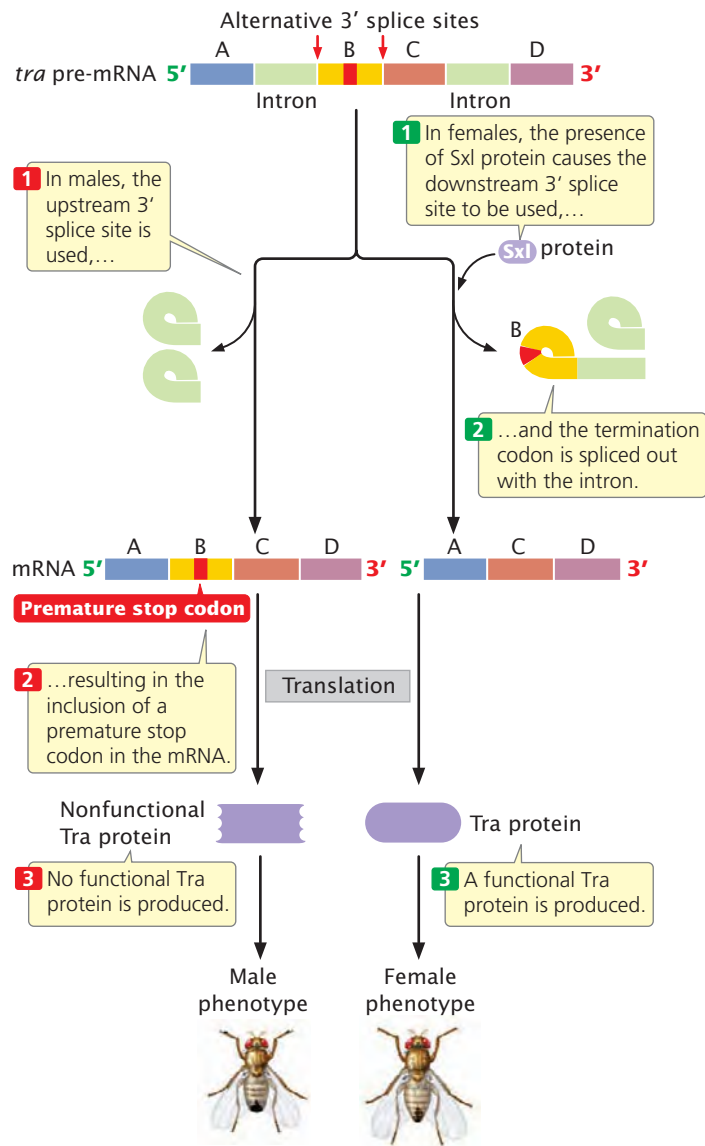
CONCEPTS

Eukaryotic genes can be regulated through the control of mRNA processing. The selection of alternative splice sites leads to the production of different proteins.

The Degradation of RNA

The amount of a protein that is synthesized depends on the amount of corresponding mRNA available for translation. The amount of available mRNA, in turn, depends on both the rate of mRNA synthesis and the rate of mRNA degradation. Eukaryotic mRNAs are generally more stable than bacterial mRNAs, which typically last only a few minutes before being degraded. Nonetheless, there is great variability in the stability of eukaryotic mRNA: some mRNAs persist for only a few minutes; others last for hours, days, or even months. These variations can result in large differences in the amount of protein that is synthesized.

Cellular RNA is degraded by ribonucleases, enzymes that specifically break down RNA. Most eukaryotic cells contain 10 or more types of ribonucleases, and there are several different pathways of mRNA degradation. In one pathway, the 5' cap is first removed, followed by 5'→3' removal of nucleotides. A second pathway begins at the 3' end of the mRNA and removes nucleotides in the 3'→5' direction. In a third pathway, the mRNA is cleaved at internal sites.



17.12 Alternative splicing of *tra* pre-mRNA. Two alternative 3' splice sites are present.

Messenger RNA degradation from the 5' end is most common and begins with the removal of the 5' cap. This pathway is usually preceded by the shortening of the poly(A) tail. Poly(A)-binding proteins (PABPs) normally bind to the poly(A) tail and contribute to its stability-enhancing effect. The presence of these proteins at the 3' end of the mRNA protects the 5' cap. When the poly(A) tail has been shortened below a critical limit, the 5' cap is removed, and nucleases then degrade the mRNA by removing nucleotides from the 5' end. These observations suggest that the 5' cap and the 3' poly(A) tail of eukaryotic mRNA physically interact with each other, most likely by the poly(A) tail bending around so that the PABPs make contact with the 5' cap (see Figure 15.18).

Much of RNA degradation takes place in specialized complexes called P bodies. However, P bodies appear to be

more than simply destruction sites for RNA. Recent evidence suggests that P bodies can temporarily store mRNA molecules, which may later be released and translated. Thus, P bodies help control the expression of genes by regulating which RNA molecules are degraded and which are sequestered for later release. RNA degradation facilitated by small interfering RNAs (siRNAs) also may take place within P bodies (see next section).

Other parts of eukaryotic mRNA, including sequences in the 5' untranslated region (5' UTR), the coding region, and the 3' UTR, also affect mRNA stability. Some short-lived eukaryotic mRNAs have one or more copies of a consensus sequence consisting of 5'-AUUUAUAA-3', referred to as the AU-rich element, in the 3' UTR. Messenger RNAs containing AU-rich elements are degraded by a mechanism in which microRNAs take part (see next section). **TRY PROBLEM 26** →

CONCEPTS

The stability of mRNA influences gene expression by affecting the amount of mRNA available to be translated. The stability of mRNA is affected by the 5' cap, the poly(A) tail, the 5' UTR, the coding section, and sequences in 3' UTR.

✓ CONCEPT CHECK 5

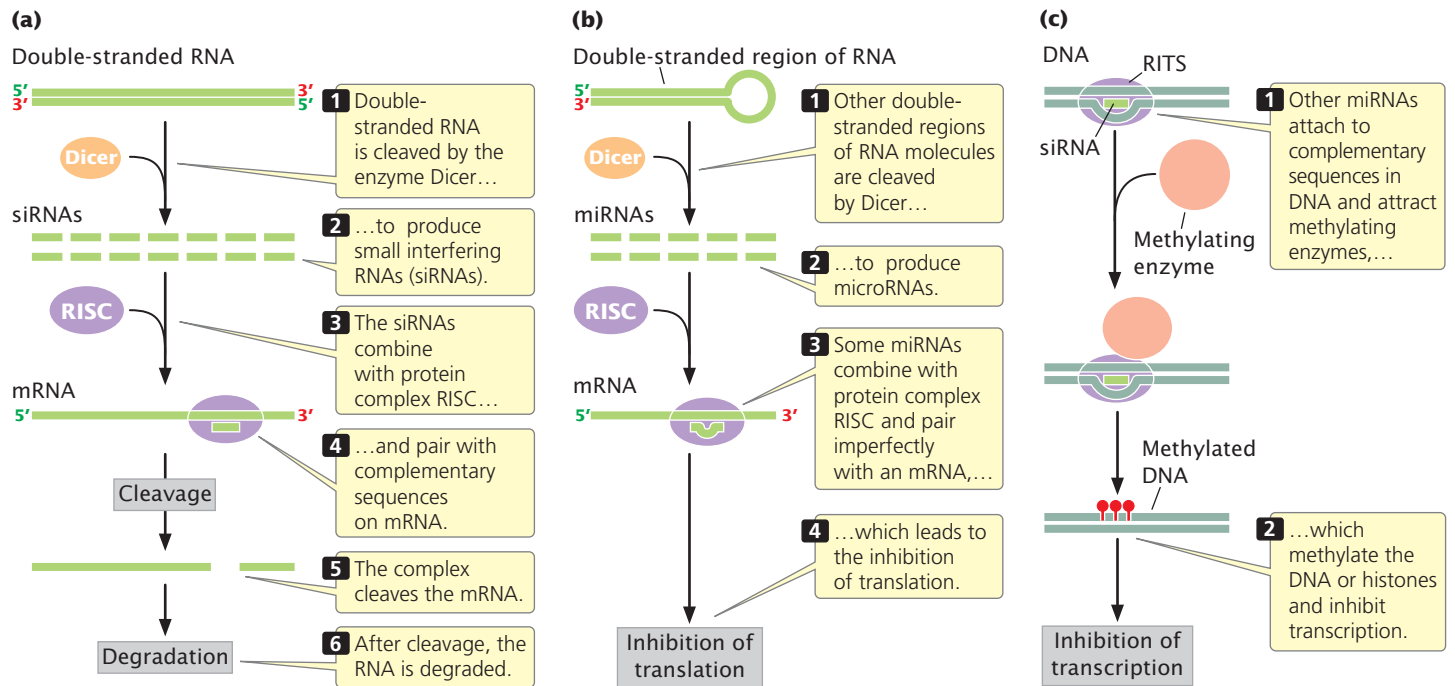
How does the poly(A) tail affect stability?

17.6 RNA Interference Is an Important Mechanism of Gene Regulation

The expression of a number of eukaryotic genes is controlled through RNA interference, also known as RNA silencing and posttranscriptional gene silencing (see Chapter 14). Recent research suggests that as much as 30% of human genes are regulated by RNA interference. Although many of the details of this mechanism are still being investigated, RNA interference is widespread in eukaryotes, existing in fungi, plants, and animals. This mechanism is also widely used as a powerful technique for artificially regulating gene expression in genetically engineered organisms (see Chapter 19).

Small Interfering RNAs and MicroRNAs

RNA interference is triggered by microRNAs (miRNAs) and small interfering RNAs (siRNAs), depending on their origin and mode of action (see Chapter 14). An enzyme called Dicer cleaves and processes double-stranded RNA to produce siRNAs or miRNAs that are from 21 to 25 nucleotides in length (**Figure 17.13**) and pair with proteins to form an RNA-induced silencing complex (RISC). The RNA component of RISC then pairs with complementary base sequences of specific mRNA molecules, most often with sequences in the 3' UTR of the mRNA. Small interfering RNAs tend to



17.13 RNA silencing leads to the degradation of mRNA or to the inhibition of translation or transcription. (a) Small interfering RNAs (siRNAs) degrade mRNA by cleavage. (b) MicroRNAs (miRNAs) lead to the inhibition of translation. (c) Some small interfering RNAs (siRNAs) methylate histone proteins or DNA, inhibiting transcription.

base pair perfectly with the mRNAs, whereas miRNAs often form less-than-perfect pairings.

Mechanisms of Gene Regulation by RNA Interference

Small interfering RNAs and microRNAs regulate gene expression through at least four distinct mechanisms: (1) cleavage of mRNA, (2) inhibition of translation, (3) transcriptional silencing, or (4) degradation of mRNA.

RNA cleavage RISCs that contain an siRNA (and some that contain a miRNA) pair with mRNA molecules and cleave the mRNA near the middle of the bound siRNA (see Figure 17.13a). This cleavage is sometimes referred to as “Slicer activity.” After cleavage, the mRNA is further degraded. Thus, the presence of siRNAs and miRNAs increase the rate at which mRNAs are broken down and decrease the amount of protein produced.

Inhibition of translation Some miRNAs regulate genes by inhibiting the translation of their complementary mRNAs (see Figure 17.13b). For example, an important gene in flower development in *Arabidopsis thaliana* is *APETALA2*. The expression of this gene is regulated by an miRNA that base pairs with nucleotides in the coding region of *APETALA2* mRNA and inhibits its translation.

The exact mechanism by which miRNAs repress translation is not known, but some research suggests that it can

inhibit both the initiation step of translation and steps after translation initiation such as those causing premature termination. Many mRNAs have multiple miRNA-binding sites, and translation is most efficiently inhibited when multiple miRNAs are bound to the mRNA.

Transcriptional silencing Other siRNAs silence transcription by altering chromatin structure. These siRNAs combine with proteins to form a complex called RITS (for RNA transcriptional silencing; see Figure 17.13c), which is analogous to RISC. The siRNA component of RITS then binds to its complementary sequence in DNA or an RNA molecule in the process of being transcribed and represses transcription by attracting enzymes that methylate the tails of histone proteins. The addition of methyl groups to the histones causes them to bind DNA more tightly, restricting the access of proteins and enzymes necessary to carry out transcription (see earlier section on histone modification). Some miRNAs bind to complementary sequences in DNA and attract enzymes that methylate the DNA directly, which also leads to the suppression of transcription (see earlier section on DNA methylation).

Slicer-independent degradation of mRNA A final mechanism by which miRNAs regulate gene expression is by triggering the decay of mRNA in a process that does not require Slicer activity. For example, a short-lived mRNA with an AU-rich element in its 3' UTR is

degraded by an RNA-silencing mechanism. Researchers have identified an miRNA with a sequence that is complementary to the consensus sequence in the AU-rich element. This miRNA binds to the AU-rich element and, in a way that is not yet fully understood, brings about the degradation of the mRNA in a process that requires Dicer and RISC.

The Control of Development by RNA Interference

Much of development in multicellular eukaryotes is through gene regulation: different genes are turned on and off at specific times (see Chapter 22). In fact, when miRNAs were first discovered, researchers noticed that a mutation in an miRNA in *C. elegans* caused a developmental defect. Research now demonstrates that miRNA molecules are key factors in controlling development in plants, animals, and humans. For example, the vertebrate heart develops through the programmed differentiation and proliferation of cardiomyocytes, which are controlled by a specific miRNA termed miR-1-1.

Recent studies demonstrate that, through their effects on gene expression, miRNAs play important roles in many diseases and disorders. For example, a genetic form of hearing loss has been associated with a mutation in the gene that encodes an miRNA. Other miRNAs are associated with heart disease. One miRNA called miR-1-2 is highly expressed in heart muscle. Mice genetically engineered to express only 50% of the normal amount of miR-1-2 frequently have holes in the wall that separates their left and right ventricles, a common congenital heart defect seen in newborn humans. Overexpression of another miRNA called miR-1 in the hearts of adult mice causes cardiac arrhythmia—irregular electrical activity of the heart that can be life-threatening in humans. Changes in the expression of other miRNAs have been associated with cancer.

CONCEPTS

RNA silencing is initiated by double-stranded RNA molecules that are cleaved and processed. The resulting siRNAs or miRNAs combine with proteins to form complexes that bind to complementary sequences in mRNA or DNA. The siRNAs and miRNAs affect gene expression by cleaving mRNA, inhibiting translation, altering chromatin structure, or triggering RNA degradation.

✓ CONCEPT CHECK 6

In RNA silencing, siRNAs and miRNAs usually bind to which part of the mRNA molecules that they control?

- 5' UTR
- Segment that encodes amino acids
- 3' poly(A) tail
- 3' UTR

17.7 Some Genes Are Regulated by Processes That Affect Translation or by Modifications of Proteins

Ribosomes, aminoacyl tRNAs, initiation factors, and elongation factors are all required for the translation of mRNA molecules. The availability of these components affects the rate of translation and therefore influences gene expression. For example, the activation of T lymphocytes (T cells) is critical to the development of immune responses to viruses (see Chapter 22). T cells are normally in the G_0 stage of the cell cycle and not actively dividing. On exposure to viral antigens, however, specific T cells become activated and undergo rapid proliferation (**Figure 17.14**). Activation includes a 7- to 10-fold increase in protein synthesis that causes cells to enter the cell cycle and proliferate. This burst of protein synthesis does not require an increase in mRNA synthesis. Instead, a global increase in protein synthesis is due to the increased availability of initiation factors taking part in translation—initiation factors that allow ribosomes to bind to mRNA and begin translation. Similarly, insulin stimulates the initiation of overall protein synthesis by increasing the availability of initiation factors. Initiation factors exist in inactive forms and, in response to various cell signals, can be activated by chemical modifications of their structure, such as phosphorylation.

Mechanisms also exist for the regulation of translation of specific mRNAs. The initiation of translation in some mRNAs is regulated by proteins that bind to an mRNA's 5' UTR and inhibit the binding of ribosomes, in a fashion similar to the way in which repressor proteins bind to operators and prevent the transcription of structural genes. The translation of some mRNAs is affected by the binding of proteins to sequences in the 3' UTR.

Many eukaryotic proteins are extensively modified after translation by the selective cleavage and trimming of amino acids from the ends, by acetylation, or by the addition of phosphate groups, carboxyl groups, methyl groups, or carbohydrates to the protein. These modifications affect the transport, function, and activity of the proteins and have the capacity to affect gene expression.

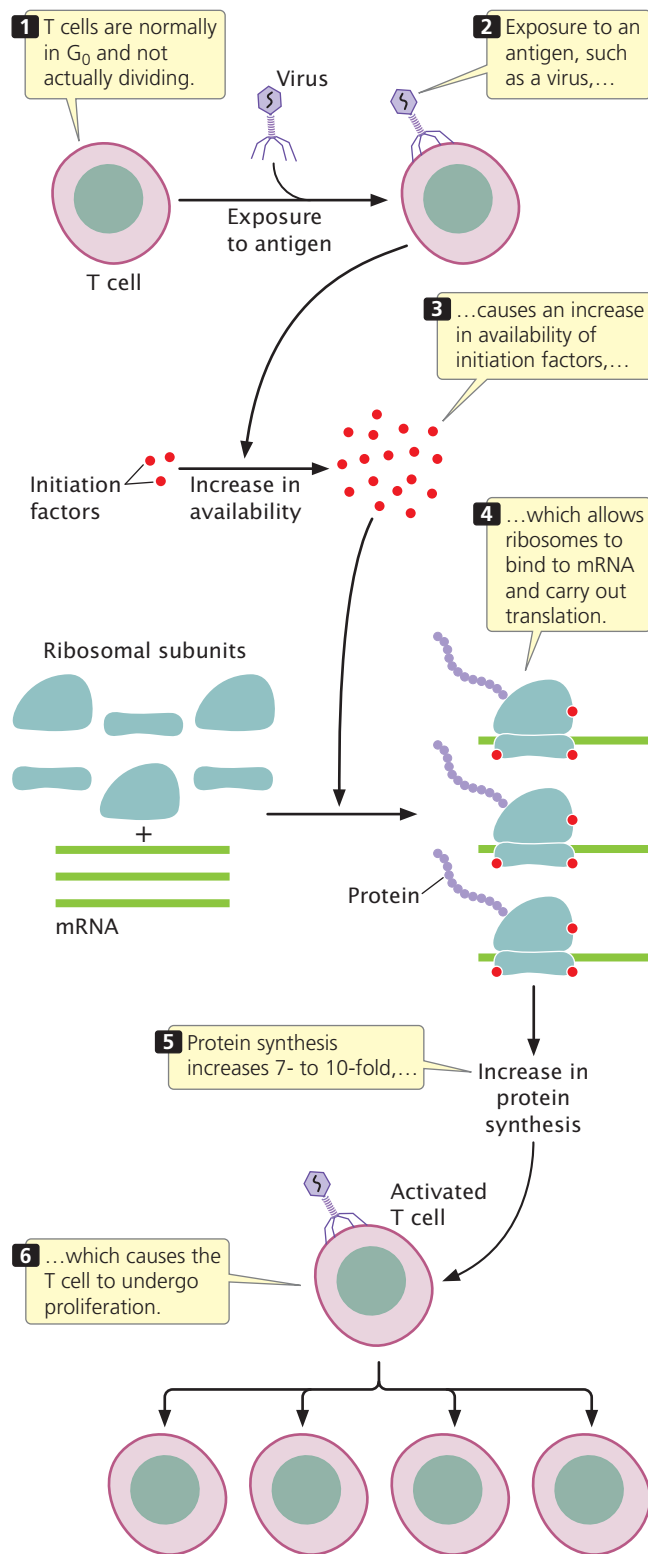
CONCEPTS

The initiation of translation may be affected by proteins that bind to specific sequences at the 5' end of mRNA. The availability of ribosomes, tRNAs, initiation and elongation factors, and other components of the translational apparatus may affect the rate of translation.

CONNECTING CONCEPTS

A Comparison of Bacterial and Eukaryotic Gene Control

Now that we have considered the major types of gene regulation in bacteria (see Chapter 16) and eukaryotes (this chapter), let's pause



17.14 The expression of some eukaryotic genes is regulated by the availability of components required for translation. In this example, exposure to an antigen stimulates an increased availability of initiation factors and a subsequent increase in protein synthesis, leading to T-cell proliferation.

to consider some of the similarities and differences in bacterial and eukaryotic gene control.

1. Much of gene regulation in bacterial cells is at the level of transcription (although it does exist at other levels). Gene

regulation in eukaryotic cells takes place at multiple levels, including chromatin structure, transcription, mRNA processing, RNA stability, RNA interference, and posttranslational control.

2. Complex biochemical and developmental events in bacterial and eukaryotic cells may require a cascade of gene regulation, in which the activation of one set of genes stimulates the activation of another set.
3. Much of gene regulation in both bacterial and eukaryotic cells is accomplished through proteins that bind to specific sequences in DNA. Regulatory proteins come in a variety of types, but most can be characterized according to a small set of DNA-binding motifs.
4. Chromatin structure plays a role in eukaryotic (but not bacterial) gene regulation. In general, condensed chromatin represses gene expression; chromatin structure must be altered before transcription can take place. Chromatin structure is altered by the modification of histone proteins, chromatin-remodeling proteins, and DNA methylation.
5. Modifications to chromatin structure in eukaryotes may lead to epigenetic changes, which are changes that affect gene expression and are passed on to other cells or future generations.
6. In bacterial cells, genes are often clustered in operons and are coordinately expressed by transcription into a single mRNA molecule. In contrast, most eukaryotic genes typically have their own promoters and are transcribed independently. Coordinate regulation in eukaryotic cells takes place through common response elements, present in the promoters and enhancers of the genes. Different genes that have the same response element in common are influenced by the same regulatory protein.
7. Regulatory proteins that affect transcription exhibit two basic types of control: *repressors* inhibit transcription (negative control); *activators* stimulate transcription (positive control). Both negative control and positive control are found in bacterial and eukaryotic cells.
8. The initiation of transcription is a relatively simple process in bacterial cells, and regulatory proteins function by blocking or stimulating the binding of RNA polymerase to DNA. In contrast, eukaryotic transcription requires complex machinery that includes RNA polymerase, general transcription factors, and transcriptional activators, which allows transcription to be influenced by multiple factors.
9. Some eukaryotic transcriptional regulatory proteins function at a distance from the gene by binding to enhancers, causing the formation of a loop in the DNA, which brings the promoter and enhancer into close proximity. Some distant-acting sequences analogous to enhancers have been described in bacterial cells, but they appear to be less common.
10. The greater time lag between transcription and translation in eukaryotic cells compared with that in bacterial cells allows mRNA stability and mRNA processing to play larger roles in eukaryotic gene regulation.
11. RNA molecules (antisense RNA) may act as regulators of gene expression in bacteria. Regulation by siRNAs and miRNAs, which is extensive in eukaryotes, is absent from bacterial cells.