

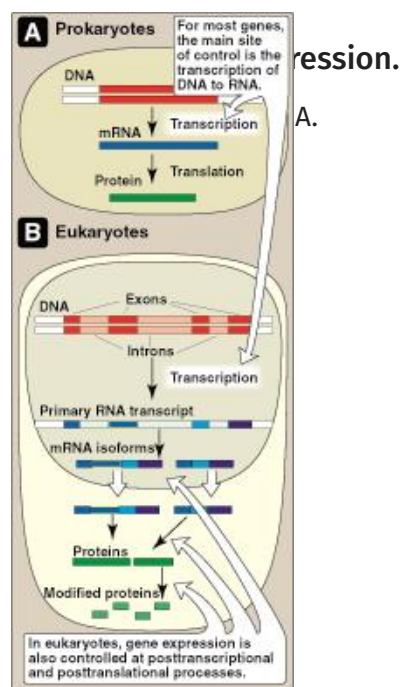
33: Regulation of Gene Expression

Overview



Gene expression refers to the multistep process that ultimately results in the production of a functional gene product, either ribonucleic acid (RNA) or protein. The first step in gene expression, the use of deoxyribonucleic acid (DNA) for the synthesis of RNA (transcription), is the primary site of regulation in both prokaryotes and eukaryotes. In eukaryotes, however, gene expression also involves extensive posttranscriptional and posttranslational processes as well as actions that influence access to particular regions of the DNA. Each of these steps can be regulated to provide additional control over the kinds and amounts of functional products that are produced.

Not all genes are tightly regulated. For example, genes described as ‘constitutive’ encode products required for basic cellular functions and so are expressed at essentially a constant level. They are also known as “housekeeping” genes. Regulated genes, however, are expressed only under certain conditions. They may be expressed in all cells of the body or in only a subset of cells, for example, the gene for fibrinogen alpha chain, which is expressed only in hepatocytes. The ability to regulate gene expression (i.e., to determine if, how much, and when particular gene products will be made) gives the cell control over structure and function. It is the basis for cellular differentiation, morphogenesis, and adaptability of any organism. Control of gene expression is best understood in prokaryotes, but many themes are repeated in eukaryotes. [Figure 33.1](#) shows some of the sites where gene expression can be controlled.

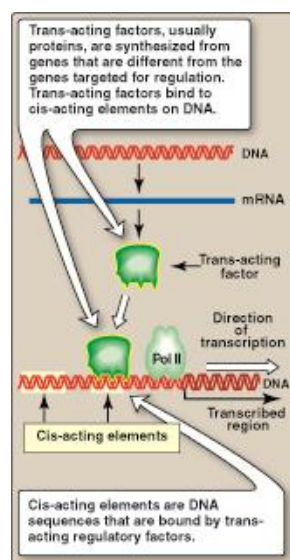
FIGURE 33.1

Regulatory Sequences and Molecules



Regulation of transcription, the initial step in all gene expression, is controlled by regulatory sequences of DNA that are usually embedded in the noncoding regions of the genome. The interaction between these DNA sequences and regulatory molecules, such as transcription factors, can induce or repress the transcriptional machinery, influencing the kinds and amounts of products that are produced. The regulatory DNA sequences are called *cis*-acting because they influence expression of genes on the same chromosome as the regulatory sequence (see p. 488). The regulatory molecules are called *trans*-acting because they can diffuse (transit) through the cell from their site of synthesis to their DNA-binding sites ([Fig. 33.2](#)). For example, a protein transcription factor (a *trans*-acting molecule) that regulates a gene on chromosome 6 might itself have been encoded by a gene on chromosome 11. The binding of proteins to DNA is through structural motifs such as the zinc finger ([Fig. 33.3](#)), leucine zipper, or helix-turn-helix in the protein.

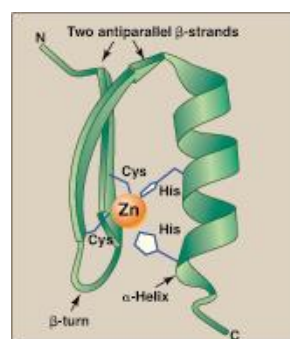
FIGURE 33.2



and trans-acting factors.

A; Pol II = RNA polymerase II.

FIGURE 33.3



common motif in proteins that bind DNA.

histidine.

Regulation of Prokaryotic Gene Expression



In prokaryotes such as the bacterium *Escherichia coli* (*E. coli*), regulation of gene expression occurs primarily at the level of transcription and, in general, is mediated by the binding of trans-acting proteins to cis-acting regulatory elements on their single DNA molecule (chromosome). (Note: Regulating the first step in the expression of a gene is an efficient approach, insofar as energy is not wasted making unneeded gene products.) Transcriptional control in prokaryotes can involve the initiation or premature termination of transcription.

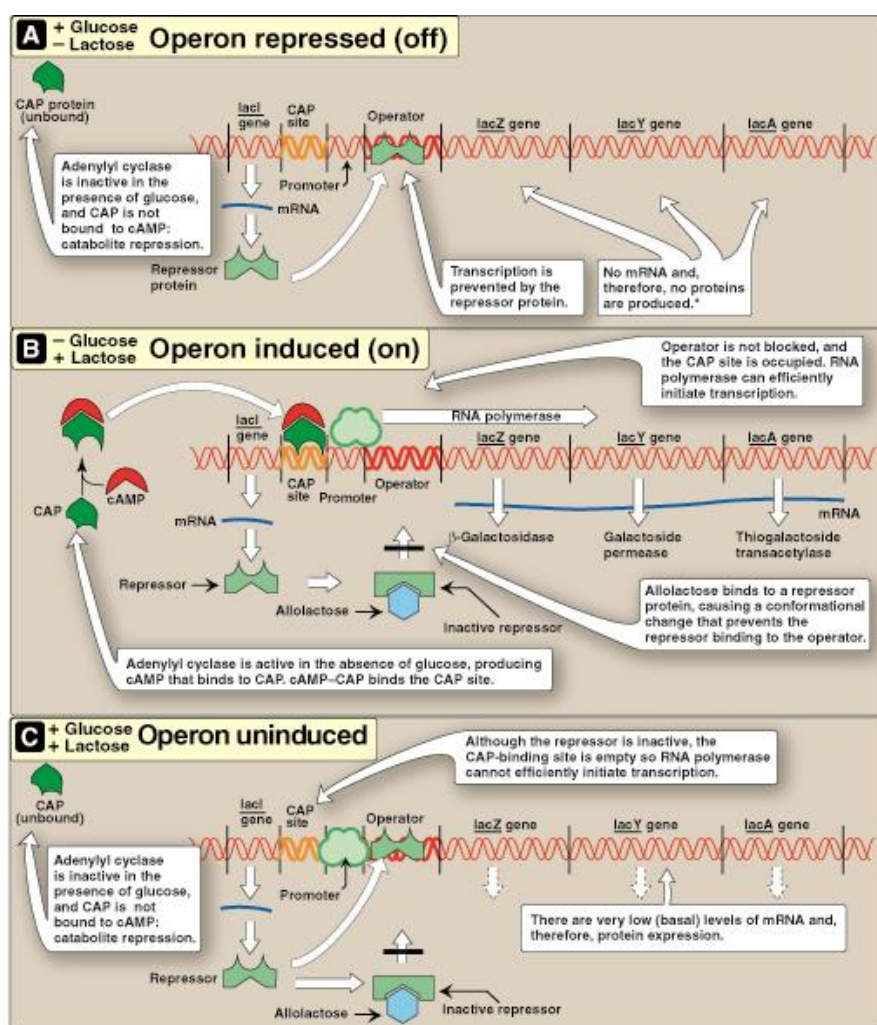
Messenger RNA transcription from bacterial operons

In bacteria, the structural genes that encode proteins involved in a particular metabolic pathway are often found sequentially grouped on the chromosome along with the cis-acting elements that regulate the transcription of these genes. The transcription product is a single polycistronic messenger RNA ([mRNA], see p. 483). The genes are, thus, coordinately regulated (i.e., turned on or off as a unit). This entire package is referred to as an operon.

Operators in bacterial operons

Bacterial operons contain an operator, a segment of DNA that regulates the activity of the structural genes of the operon by reversibly binding a protein known as the repressor. If the operator is not bound by the repressor, RNA polymerase (RNA pol) binds the promoter, passes over the operator, and reaches the protein-coding genes that it transcribes to mRNA. If the repressor is bound to the operator, the polymerase is blocked and does not produce mRNA. As long as the repressor is bound to the operator, no mRNA (and, therefore, no proteins) are made. However, when an inducer molecule is present, it binds to the repressor, causing the repressor to change shape so that it no longer binds the operator. When this happens, RNA pol can initiate transcription. One of the best-understood examples is the inducible lactose (*lac*) operon of *E. coli* that illustrates both positive and negative regulation (Fig. 33.4).

FIGURE 33.4



/ glucose, (B) only lactose, and (C)

ntly dissociates from the operator at a molecules of permease (and β -ecome unavailable.) CAP = catabolite ssenger RNA.

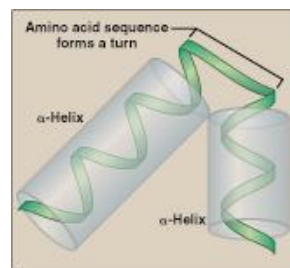
Lactose operon

The *lac* operon contains the genes that code for three proteins involved in the catabolism of the disaccharide lactose: the *lacZ* gene codes for β -galactosidase, which hydrolyzes lactose to galactose and glucose; the *lacY* gene codes for a permease, which facilitates the movement of lactose into the cell; and the *lacA* gene codes for thiogalactosidetransacetylase, which acetylates lactose. (Note: The physiologic function of this acetylation is unknown.) All of these proteins are maximally produced only when lactose is available to the cell and glucose is not. (Note: Bacteria use glucose, if available, as a fuel in preference to any other sugar.) The regulatory portion of the operon is upstream of the three structural genes and consists of the promoter region where RNA pol binds and two additional sites, the operator (O) and the catabolite activator protein (CAP) sites, where regulatory proteins bind. The *lacZ*, *lacY*, and *lacA* genes are maximally expressed only when the O site is empty and the CAP site is bound by a complex of cyclic adenosine monophosphate ([cAMP], see p. 103) and the CAP, sometimes called the cAMP regulatory protein (CRP). A regulatory gene, the *lacI* gene, codes for the repressor protein (a trans-acting factor) that binds to the O site with high affinity. (Note: The *lacI* gene has its own promoter and is not part of the *lac* operon.)

When only glucose is available

In this case, the *lac* operon is repressed (turned off). Repression is mediated by the repressor protein binding via a helix-turn-helix motif (Fig. 33.5) to the O site, which is downstream of the promoter (see Fig. 33.4A). Binding of the repressor interferes with the binding of RNA pol to the promoter, thereby inhibiting transcription of the structural genes. This is an example of negative regulation.

FIGURE 33.5



motif of the *lac* repressor protein.

When only lactose is available

In this case, the *lac* operon is induced (maximally expressed, or turned on). A small amount of lactose is converted to an isomer, allolactose. This compound is an inducer that binds to the repressor protein, changing its conformation so that it can no longer bind to the O site. In the absence of glucose, adenyl cyclase is active, and cAMP is made and binds to the CAP. The cAMP–CAP trans-acting complex binds to the CAP site, causing RNA pol to initiate transcription with high efficiency at the promoter site (see Fig. 33.4B). This is an example of positive regulation. The transcript is a single polycistronic mRNA molecule that contains three sets of start and stop codons. Translation of the mRNA produces the three proteins that allow lactose to be used for energy production by the cell. (Note: In contrast to the inducible *lacZ*, *lacY*, and *lacA* genes, whose expression is regulated, the *lacI* gene is constitutive. Its gene product, the repressor protein, is always made and is active unless the inducer is present.)

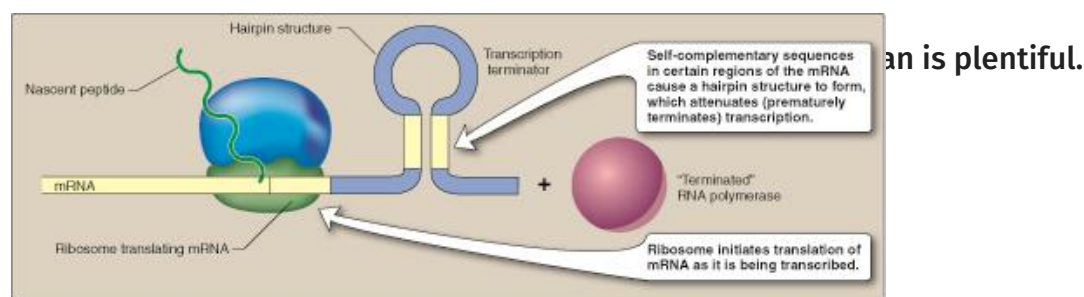
When both glucose and lactose are available

In this case, the *lac* operon is uninduced, and transcription is negligible, even if lactose is present at a high concentration. Adenylyl cyclase is inhibited in the presence of glucose (a process known as catabolite repression) so no cAMP–CAP complex forms, and the CAP site remains empty. Therefore, the RNA pol is unable to effectively initiate transcription, even though the repressor is not bound to the O site. Consequently, the three structural genes of the operon are expressed only at a very low (basal) level (see Fig. 33.4C). (Note: Induction causes a 50-fold enhancement over basal expression.)

Tryptophan operon

The tryptophan (*trp*) operon contains five structural genes that code for enzymes required for the synthesis of the amino acid tryptophan (Trp). As with the *lac* operon, the *trp* operon is subject to negative control. However, for the repressible *trp* operon, negative control includes Trp itself binding to a repressor protein and facilitating the binding of the repressor to the operator: Trp is a corepressor. Because repression by Trp is not always complete, the *trp* operon, unlike the *lac* operon, is also regulated by a process known as attenuation. With attenuation, transcription is initiated but is terminated well before completion (Fig. 33.6). If Trp is plentiful, transcription initiation that escaped repression by Trp is attenuated (stopped) by the formation of an attenuator, a hairpin (stem-loop) structure in the mRNA similar to that seen in rho-independent termination (see p. 486). (Note: Because transcription and translation are temporally linked in prokaryotes [see p. 503], attenuation also results in the formation of a truncated, nonfunctional peptide product that is rapidly degraded.) If Trp becomes scarce, the operon is expressed. The 5' end of the mRNA contains two adjacent codons for Trp. The lack of Trp causes ribosomes to stall at these codons, covering regions of the mRNA required for formation of the attenuation hairpin. This prevents attenuation and allows transcription to continue.

FIGURE 33.6



Transcriptional attenuation can occur in prokaryotes because translation of an mRNA begins before its synthesis is complete. This does not occur in eukaryotes because the presence of a membrane-bound nucleus spatially and temporally separates transcription and translation.

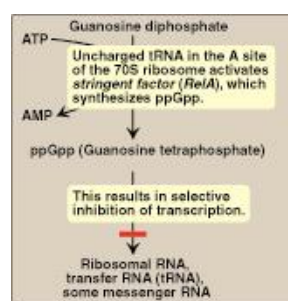
Coordination of transcription and translation

Although transcriptional regulation of mRNA production is primary in bacteria, regulation of ribosomal RNA (rRNA) and protein synthesis plays an important role in adaptation to environmental stress.

Stringent response

E. coli has seven operons that synthesize the rRNA needed for ribosome assembly, and each is regulated in response to changes in environmental conditions. Regulation in response to amino acid starvation is known as the stringent response. The binding of an uncharged transfer RNA (tRNA) to the A site of a ribosome (see p. 501) triggers a series of events that leads to the production of the alarmone, guanosine 5'-diphosphate, 3'-diphosphate (ppGpp). The synthesis of this unusual derivative of guanosine diphosphate (GDP) is catalyzed by stringent factor (RelA), an enzyme physically associated with ribosomes. Elevated levels of ppGpp result in inhibition of rRNA synthesis (Fig. 33.7). ppGpp binds RNA pol and alters promoter selection through the use of different sigma factors for the polymerase (see p. 484). In addition to rRNA synthesis, tRNA synthesis and some mRNA synthesis (e.g., for ribosomal proteins [r-proteins]) are also inhibited. However, synthesis of mRNA for enzymes required for amino acid biosynthesis is not inhibited. The stringent response prevents the wasteful production of more ribosomes and promotes the production of needed amino acids when amino acids are scarce.

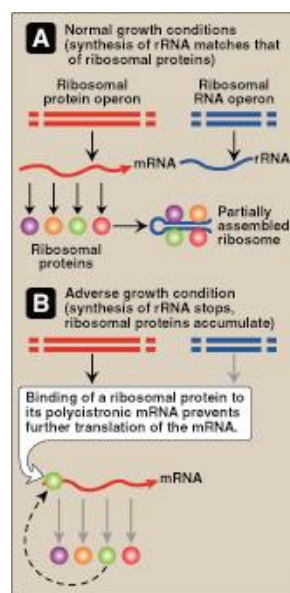
FIGURE 33.7



cription by the stringent response to amino acid starvation.

Regulatory ribosomal proteins

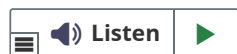
Operons for r-proteins can be inhibited by an excess of their own protein products. For each operon, one specific r-protein functions in the repression of translation of the polycistronic mRNA from that operon (Fig. 33.8). The r-protein does so by binding to the Shine–Dalgarno (SD) sequence located on the mRNA just upstream of the first initiating AUG codon (see p. 497) and acting as a physical impediment to the binding of the small ribosomal subunit to the SD sequence. Thus, one r-protein inhibits synthesis of all the r-proteins of the operon. This same r-protein also binds to rRNA and with a higher affinity than for mRNA. If the concentration of rRNA falls, the r-protein then is available to bind its own mRNA and inhibit its translation. This coordinated regulation keeps the synthesis of r-proteins in balance with the transcription of rRNA, so that each is present in appropriate amounts for the formation of ribosomes.

FIGURE 33.8

ation by an excess of ribosomal proteins.

A; rRNA = ribosomal RNA.

Regulation of Eukaryotic Gene Expression



The higher degree of complexity of eukaryotic genomes, as well as the presence of a nuclear membrane, necessitates a wider range of regulatory processes. As with the prokaryotes, transcription is the primary site of regulation. Again, the theme of trans-acting factors binding to cis-acting elements is seen. Operons, however, are not found in eukaryotes, which must use alternate strategies to solve the problem of how to coordinately regulate all the genes required for a specific response. In eukaryotes, gene expression is also regulated at multiple levels other than transcription. For example, the major modes of posttranscriptional regulation at the mRNA level are alternative mRNA splicing and polyadenylation, control of mRNA stability, and control of translational efficiency. Additional regulation at the protein level occurs by mechanisms that modulate stability, processing, or targeting of the protein.

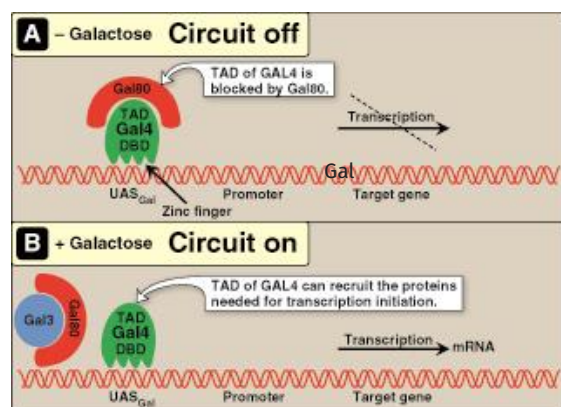
Coordinate regulation

The need to coordinately regulate a group of genes to cause a particular response is of key importance in organisms with more than one chromosome. An underlying theme occurs repeatedly: A trans-acting protein functions as a specific transcription factor (STF) that binds to a cis-acting regulatory consensus sequence (see p. 463) on each of the genes in the group even if they are on different chromosomes. (Note: The STF has a DNA-binding domain [DBD] and a transcription-activation domain [TAD]. The TAD recruits coactivators, such as histone acetyltransferases [see p. 487], and the general transcription factors [see p. 488] that, along with RNA pol, are required for formation of the transcription initiation complex at the promoter. Although the TAD recruits a variety of proteins, the specific effect of any one of them is dependent upon the protein composition of the complex. This is known as combinatorial control.) Examples of coordinate regulation in eukaryotes include the galactose circuit and the hormone response system.

Galactose circuit

This regulatory scheme allows for the use of galactose when glucose is not available. In yeast, a unicellular organism, the genes required to metabolize galactose are on different chromosomes. Coordinated expression is mediated by the protein Gal4 (Gal = galactose), a STF that binds to a short regulatory DNA sequence upstream of each of the genes. The sequence is called the upstream activating sequence Gal (UAS_{Gal}). Binding of Gal4 to UAS_{Gal} through zinc fingers in its DBD occurs in both the absence and presence of galactose. When the sugar is absent, the regulatory protein Gal80 binds Gal4 at its TAD, thereby inhibiting gene transcription (Fig. 33.9A). When present, galactose activates the Gal3 protein. Gal3 binds Gal80, thereby allowing Gal4 to activate transcription (Fig. 33.9B). (Note: Glucose prevents the use of galactose by inhibiting expression of Gal4 protein.)

FIGURE 33.9



in the (A) absence and (B) presence of galactose.

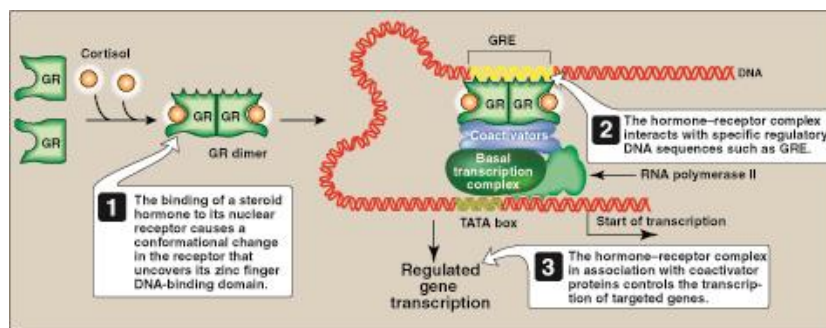
a different chromosome, each have an upstream activating
tion-activation domain; DBD = DNA-binding domain; mRNA =

Hormone response system

Hormone response elements (HREs) are DNA sequences that bind trans-acting proteins and regulate gene expression in response to hormonal signals in multicellular organisms. Hormones bind to either intracellular (nuclear) receptors (e.g., steroid hormones; see Figure 18.28) or cell-surface receptors (e.g., the peptide hormone glucagon; see Figure 23.12).

Intracellular receptors

Members of the nuclear receptor superfamily, which includes the steroid hormone (glucocorticoids, mineralocorticoids, androgens, and estrogens), vitamin D, retinoic acid, and thyroid hormone receptors, function as STF. In addition to domains for DNA-binding and transcriptional activation, these receptors also contain a ligand-binding domain. For example, the steroid hormone cortisol (a glucocorticoid) binds intracellular receptors at the ligand-binding domain (Fig. 33.10). Binding causes a conformational change in the receptor that activates it. The receptor–hormone complex enters the nucleus, dimerizes, and binds via a zinc finger motif to DNA at a regulatory element, the glucocorticoid response element (GRE) that is an example of a HRE. Binding allows recruitment of coactivators to the TAD of the receptor and results in expression of cortisol-responsive genes, each of which is under the control of its own GRE. Binding of the receptor–hormone complex to the GRE allows coordinate expression of a group of target genes, even though these genes are on different chromosomes. The GRE can be located upstream or downstream of the genes it regulates and at great distances from them. The GRE, then, can function as a true enhancer (see p. 489). (Note: If associated with repressors, hormone–receptor complexes inhibit transcription.)

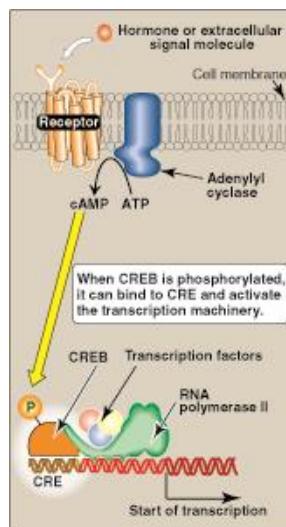
FIGURE 33.10

e receptors.

cor.

Cell-surface receptors

These receptors include those for insulin, epinephrine, and glucagon. Glucagon, for example, is a peptide hormone that binds its G protein-coupled plasma membrane receptor on glucagon-responsive cells. This extracellular signal is then transduced to intracellular cAMP, a second messenger (Fig. 33.11; also see Fig. 8.7), which can affect protein expression (and activity) through protein kinase A-mediated phosphorylation. In response to a rise in cAMP, a trans-acting factor (cAMP response element-binding [CREB] protein) is phosphorylated and activated. Active CREB protein binds via a leucine zipper motif to a cis-acting regulatory element, the cAMP response element (CRE) resulting in transcription of target genes with CRE in their promoters. (Note: The genes for phosphoenolpyruvate carboxykinase and glucose 6-phosphatase, key enzymes of gluconeogenesis [see p. 133] are examples of genes upregulated by the cAMP/CRE/CREB system.)

FIGURE 33.11

lation by receptors located in the cell membrane.

e monophosphate [cAMP] activates protein kinase A that phosphorylates cAMP response [CREB] protein.) CRE = cAMP response element.

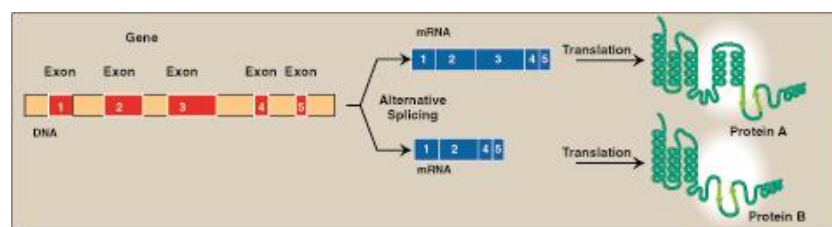
Messenger RNA processing and use

Eukaryotic mRNA undergoes several processing events before it is exported from the nucleus to the cytoplasm for use in protein synthesis. Capping at the 5' end (see p. 490), polyadenylation at the 3' end (see p. 491), and splicing (see p. 491) are essential for the production of a functional eukaryotic messenger from most pre-mRNA. Variations in splicing and polyadenylation can affect gene expression. In addition, messenger stability also affects gene expression.

Alternative splicing

Tissue-specific protein isoforms can be made from the same pre-mRNA through alternative splicing, which can involve exon skipping (loss), intron retention, and use of alternative splice donor or acceptor sites (Fig. 33.12). For example, the pre-mRNA for tropomyosin (TM) undergoes tissue-specific alternative splicing to yield a number of TM isoforms (see p. 492). (Note: Over 90% of all human genes undergo alternative splicing.)

FIGURE 33.12



proteins, or isoforms, from a single gene.

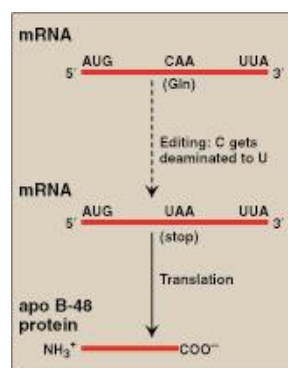
Alternative polyadenylation

Some pre-mRNA transcripts have more than one site for cleavage and polyadenylation. Alternative polyadenylation (APA) generates mRNA with different 3' ends, altering the untranslated region (UTR) or the coding (translated) sequence. (Note: APA is involved in the production of the membrane-bound and secreted forms of immunoglobulin M.)

The use of alternative splicing and polyadenylation sites, as well as alternative transcription start sites explains, at least in part, how the ~20,000 to 25,000 genes in the human genome can give rise to well over 100,000 proteins.

Messenger RNA editing

Even after mRNA has been fully processed, it may undergo an additional posttranscriptional modification in which a base in the mRNA is altered. This is known as RNA editing. An important example in humans occurs with the transcript for apolipoprotein (apo) B, an essential component of chylomicrons (see p. 254) and very-low-density lipoproteins (VLDLs), see p. 256). Apo B mRNA is made in the liver and the small intestine. However, in the intestine only, the cytosine (C) base in the CAA codon for glutamine is enzymatically deaminated to uracil (U), changing the sense codon to the nonsense or stop codon UAA, as shown in Figure 33.13. This results in a shorter protein (apo B-48, representing 48% of the message) being made in the intestine (and incorporated into chylomicrons) than is made in the liver (apo B-100, full-length, incorporated into VLDL).

FIGURE 33.13

protein (apo) B mRNA in the intestine and generation of the apo B-48 protein
protein synthesis.

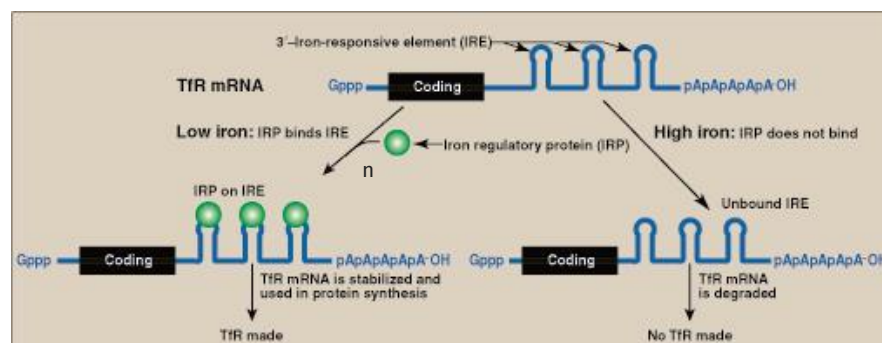
A = messenger RNA; A = adenine; C = cytosine; G = guanine; U = uracil.

Messenger RNA stability

How long an mRNA remains in the cytosol before it is degraded influences how much protein product can be produced from it. Regulation of iron metabolism and the gene-silencing process of RNA interference (RNAi) illustrate the importance of mRNA stability in the regulation of gene expression.

Iron metabolism

Transferrin (Tf) is a plasma protein that transports iron. Tf binds to cell-surface receptors (transferrin receptors [TfRs]) that get internalized and provide cells, such as erythroblasts, with iron. The mRNA for the TfR has several cis-acting iron-responsive elements (IREs) in its 3'-UTR. IREs have a short stem-loop structure that can be bound by trans-acting iron regulatory proteins (IRPs), as shown in [Figure 33.14](#). When the iron concentration in the cell is low, the IRPs bind to the 3'-IRE and stabilize the mRNA for TfR, allowing TfR synthesis. When intracellular iron levels are high, the IRPs dissociate. The lack of IRP bound to the mRNA hastens its destruction, resulting in decreased TfR synthesis. (Note: The mRNA for ferritin, an intracellular protein of iron storage, has a single IRE in its 5'-UTR. When iron levels in the cell are low, IRPs bind the 5'-IRE and prevent the use of the mRNA, and less ferritin is made. When iron accumulates in the cell, the IRPs dissociate, allowing synthesis of ferritin molecules to store the excess iron. Aminolevulinic acid synthase 2, the regulated enzyme of heme synthesis [see p. 309] in erythroblasts, also contains a 5'-IRE.) (See [Chapter 21](#) for a discussion of heme synthesis.)

FIGURE 33.14

messenger RNA [mRNA].) Gppp = 7-

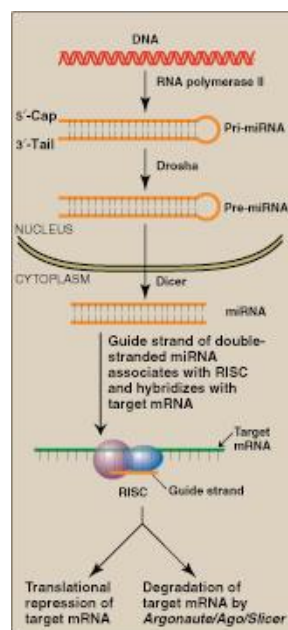
RNA interference

RNAi is a mechanism of gene silencing through decreased expression of mRNA, either by repression of translation or by increased degradation. It plays a key role in such fundamental processes as cell proliferation, differentiation, and apoptosis. RNAi is mediated by short (~22 nucleotides), noncoding RNA called microRNA (miRNA). The miRNA arise from far longer, genomically encoded nuclear transcripts, primary miRNA (pri-miRNA) that are partially processed in the nucleus to pre-miRNA by an endonuclease (Drosha), then transported to the cytoplasm. There, another endonuclease (Dicer) completes the processing and generates short, double-stranded miRNA. A single strand (the guide or antisense strand) of the miRNA associates with a cytosolic protein complex known as the RNA-induced silencing complex (RISC). The guide strand hybridizes with a complementary sequence in the 3'-UTR of a full-length target mRNA, bringing RISC to the mRNA. This can result in repression of translation of the mRNA or its degradation by an endonuclease (Argonaute/Ago/Slicer) of the RISC. The extent of complementarity appears to be the determining factor (Fig. 33.15). RNAi can also be triggered by the introduction of exogenous double-stranded short interfering RNA (siRNA) into a cell, a process that has enormous therapeutic potential.

RNA interference-based therapeutics

In 2018, the first RNAi-based therapy was approved to treat peripheral nerve disease (polyneuropathy) in patients with hereditary transthyretin-mediated amyloidosis (hATTR) caused by a mutation in the gene encoding transthyretin (TTR). The siRNA-based drug, patisiran, prevents the production of abnormal TTR protein and reduces the buildup of amyloid deposits containing TTR that form in peripheral nerves and in the heart. Several other RNAi therapeutics are undergoing clinical trials.

FIGURE 33.15



Pathways of microRNA (miRNA).

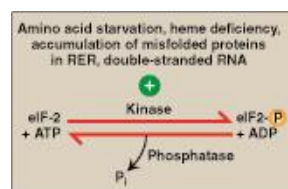
The extent of complementarity between the target messenger RNA [mRNA] and the miRNA determines the outcome. Perfect complementarity results in mRNA degradation. (Pri = primary; RISC = RNA-induced silencing complex.)

Messenger RNA translation

Regulation of gene expression can also occur at the level of mRNA translation. One mechanism by which translation is regulated is through phosphorylation of the eukaryotic translation initiation factor, eIF-2 (Fig. 33.16). Phosphorylation of eIF-2 inhibits its function and so inhibits translation at the initiation step (see p. 508). (Note: Phosphorylation of eIF-2 prevents its reactivation by inhibiting GDP–GTP exchange.)

Phosphorylation is catalyzed by kinases that are activated in response to environmental conditions, such as amino acid starvation, heme deficiency in erythroblasts, the presence of double-stranded RNA (signaling viral infection), and the accumulation of misfolded proteins in the rough endoplasmic reticulum (see p. 509).

FIGURE 33.16



Regulation of translation initiation in eukaryotes by phosphorylation of eukaryotic translation

2.

RER = rough endoplasmic reticulum; ADP = adenosine diphosphate; P_i = inorganic phosphate;  = phosphate.

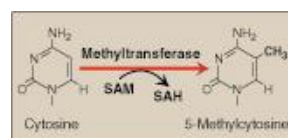
Regulation through variations in DNA

Gene expression in eukaryotes is also influenced by the accessibility of DNA to the transcriptional apparatus, the number of copies of genes, and the arrangement of DNA. (Note: Localized transitions between the B and Z forms of DNA [see p. 462] can also affect gene expression.)

Access to DNA

In eukaryotes, DNA is found complexed with histone and nonhistone proteins to form chromatin (see p. 473). Transcriptionally active, decondensed chromatin (euchromatin) differs from the more condensed, inactive form (heterochromatin) in a number of ways. Active chromatin contains histone proteins that have been covalently modified at their amino terminal ends by reversible methylation, acetylation, or phosphorylation (see p. 487 for a discussion of histone acetylation/deacetylation by histone acetyltransferase and histone deacetylase). Such modifications decrease the positive charge of these basic proteins, thereby decreasing the strength of their association with negatively charged DNA. This relaxes the nucleosome (see p. 473), allowing transcription factors access to specific regions on the DNA. Nucleosomes can also be repositioned, an ATP-requiring process that is part of chromatin remodeling. Another difference between transcriptionally active and inactive chromatin is the extent of methylation of C bases in CG-rich regions (CpG islands) in the promoter region of many genes. Methylation is by methyltransferases that use S-adenosylmethionine as the methyl donor (Fig. 33.17). Transcriptionally active genes are less methylated (hypomethylated) than their inactive counterparts, suggesting that DNA hypermethylation silences gene expression. Modification of histones and methylation of DNA are epigenetic in that they are heritable changes in DNA that alter gene expression without altering the base sequence.

FIGURE 33.17



Methylation of cytosine in eukaryotic DNA.

ionine; SAH = S-adenosylhomocysteine.

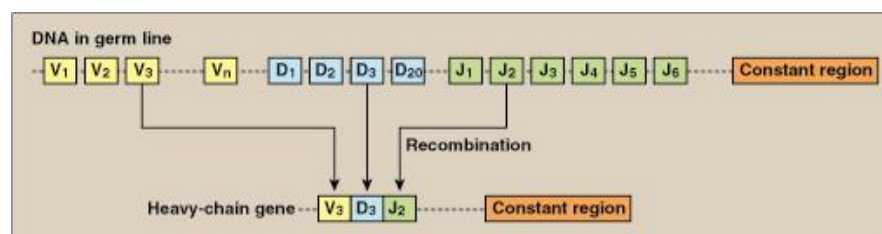
Gene copy number

A change up or down in the number of copies of a gene can affect the amount of gene product produced. An increase in copy number (gene amplification) has contributed to increased genomic complexity and is still a normal developmental process in certain nonmammalian species. In mammals, however, gene amplification is associated with some diseases and is involved in the mechanism by which cells develop resistance to particular chemotherapeutic drugs. One example is methotrexate, an inhibitor of the enzyme dihydrofolate reductase (DHFR), required for the synthesis of thymidine triphosphate (TTP) in the pyrimidine biosynthetic pathway (see p. 336 and [Figure 28.2](#)). TTP is essential for DNA synthesis. The amplification of the DHFR gene results in the expression of more DHFR enzyme, which enables the cells exposed to methotrexate to survive because TTP production can continue in the presence of the drug.

Arrangement of DNA

The process by which immunoglobulins (antibodies) are produced by B lymphocytes involves permanent rearrangements of the DNA in these cells. The immunoglobulins (e.g., IgG) consist of two light and two heavy chains, with each chain containing regions of variable and constant amino acid sequence. The variable region is the result of somatic recombination of segments within both the light- and the heavy-chain genes. During B-lymphocyte development, single variable (V), diversity (D), and joining (J) gene segments are randomly selected and brought together through gene rearrangement to form a unique variable region ([Fig. 33.18](#)). This process allows the generation of 10^9 to 10^{11} different immunoglobulins from a single gene, providing the diversity needed for the recognition of an enormous number of antigens. (Note: Pathologic DNA rearrangement is seen with translocation, a process by which two different chromosomes exchange DNA segments.)

FIGURE 33.18



Mobile DNA elements

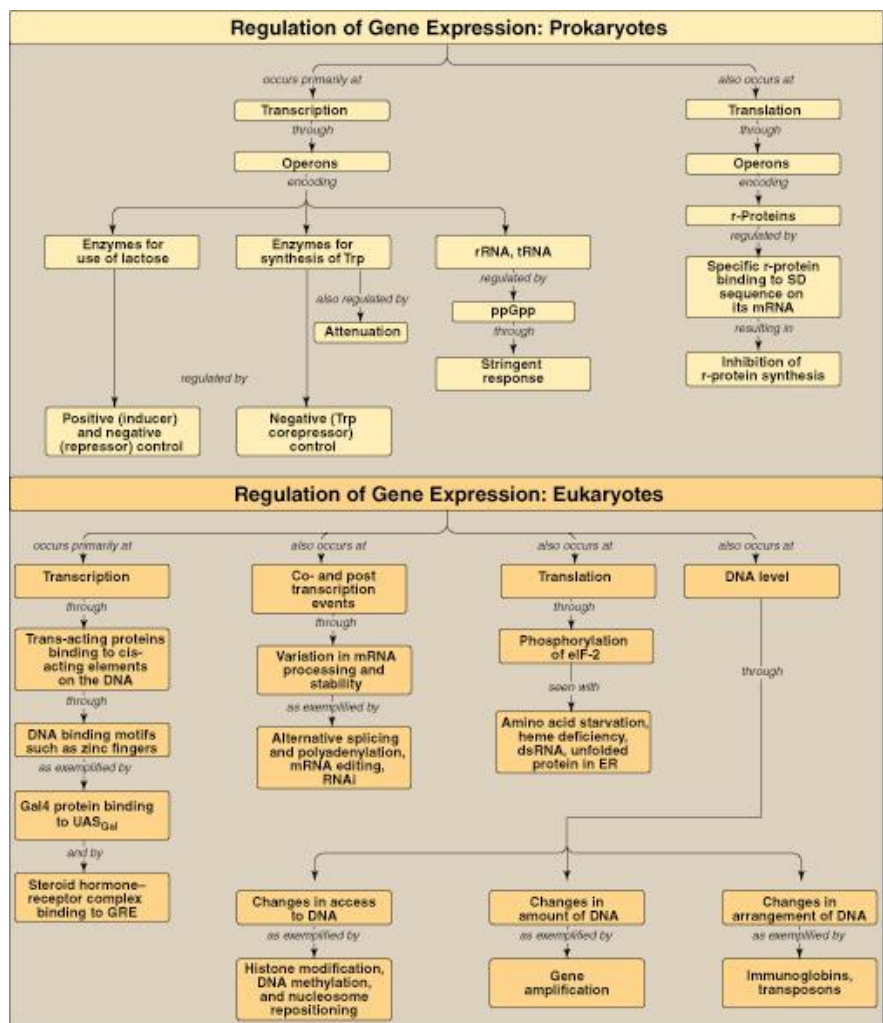
Transposons (Tns) are mobile segments of DNA that move in an essentially random manner from one site to another on the same or a different chromosome. Movement is mediated by transposase, an enzyme encoded by the Tn itself. Movement can be direct, in which transposase cuts out and then inserts the Tn at a new site, or replicative, in which the Tn is copied and the copy inserted elsewhere while the original remains in place. In eukaryotes, including humans, replicative transposition frequently involves an RNA intermediate made by a reverse transcriptase (see p. 472), in which case the Tn is called a retrotransposon. Transposition has contributed to structural variation in the genome but also has the potential to alter gene expression and even to cause disease. Tns comprise ~50% of the human genome, with retrotransposons accounting for 90% of Tns. Although the vast majority of these retrotransposons have lost the ability to move, some are still active. Their transposition is thought to be the basis for some rare cases of hemophilia A and Duchenne muscular dystrophy. (Note: The growing problem of antibiotic-resistant bacteria is a consequence, at least in part, of the exchange of plasmids among bacterial cells. If the plasmids contain Tn-carrying antibiotic resistance genes, then these genes can move from the plasmid to the bacterial chromosome so that the bacterium is resistant to one or more antimicrobial drugs even if the plasmid is lost from the cell.)

Chapter Summary



- **Gene expression** produces a functional gene product (either RNA or protein).
- **Genes** can be either **constitutive** (always expressed) or **regulated** (expressed only under certain conditions).
- Regulation of gene expression occurs primarily at **transcription** in both prokaryotes and eukaryotes and is mediated through **trans-acting proteins** binding to **cis-acting regulatory DNA elements** (Fig. 33.19).

FIGURE 33.19



- In **eukaryotes**, regulation also occurs through **DNA modifications** and through **posttranscriptional** and **posttranslational processing**.
- In **prokaryotes**, the coordinate regulation of genes whose protein products are required for a particular process is achieved through **operons** (groups of functionally related genes sequentially arranged on the chromosome along with the regulatory elements that determine their transcription). Examples from *E. coli* are the **lac operon** containing the Z, Y, and A structural genes involved in the catabolism of lactose, and the **trp operon**, which contains genes needed for the synthesis of Trp. The trp operon is also regulated by attenuation, in which mRNA synthesis that escaped repression by Trp is terminated before completion.
- In prokaryotes, transcription of **rRNA** and **tRNA** is selectively inhibited by the **stringent response** to amino acid starvation. **Translation** is also a site of prokaryotic gene regulation: Excess r-proteins bind the **SD sequence** on their own polycistronic mRNA, preventing ribosomes from binding.

- In eukaryotes, hormones coordinate the expression of groups of genes by binding to an intracellular receptor that acts as a trans-acting protein (as with steroid hormones) or to a cell surface receptor that initiates **second messenger** signaling to activate a trans-acting protein (as with peptide hormones). In each case, the protein recognizes a specific response element and binds to the DNA sequence using structural motifs such as a **zinc finger** or a **leucine zipper**.
- **Co-** and **posttranscriptional regulation** is also seen in eukaryotes and includes **alternative mRNA splicing** and **polyadenylation**, mRNA **editing**, and variations in mRNA **stability**. **Transferrin receptor** synthesis is enhanced by mRNA stability when iron concentrations are low. **RNA interference** is used to control mRNA stability and translation and is the basis for a new class of therapeutic agents.
- Regulation at the **translational level** can be caused by the **phosphorylation** and inhibition of **eukaryotic initiation factor-2**. Gene expression in eukaryotes is also influenced by **accessibility** of DNA to the transcriptional apparatus (as seen with **epigenetic** changes to histone proteins), the gene copy number, and the **arrangement** of the DNA.

Study Questions



Choose the **ONE** best answer.

33.1. Which of the following mutations is most likely to result in reduced expression of the *lac* operon?

- A. cya^- (no adenyl cyclase made)
- B. i^- (no repressor protein made)
- C. O^c (operator cannot bind repressor protein)
- D. One resulting in impaired glucose uptake
- E. $relA^-$ (no stringent response occurs)

Correct answer = A. In the absence of glucose, adenyl cyclase makes cyclic adenosine monophosphate (cAMP), which forms a complex with the catabolite activator protein (CAP). The cAMP–CAP complex binds the CAP site on the DNA, causing RNA polymerase to bind more efficiently to the *lac* operon promoter, thereby increasing expression of the operon. With cya^- mutations, adenyl cyclase is not made, and so the operon is unable to be maximally expressed even when glucose is absent and lactose is present. The absence of a repressor protein or decreased ability of the repressor to bind the operator results in constitutive (essentially constant) expression of the *lac* operon.

33.2. Which of the following is best described as cis-acting?

- A. Cyclic adenosine monophosphate response element-binding protein
- B. Operator
- C. Repressor protein
- D. Thyroid hormone nuclear receptor
- E. Histone modification

Correct answer = B. The operator is part of the DNA itself, and so is cis-acting. The cyclic adenosine monophosphate response element-binding protein, repressor protein, and thyroid hormone nuclear receptor protein are molecules that diffuse (transit) to the DNA, bind, and affect the expression of that DNA and so are trans-acting.

33.3. Which of the following is the basis for the intestine-specific expression of apolipoprotein B-48?

- A. DNA rearrangement and loss
- B. DNA transposition
- C. RNA alternative splicing
- D. RNA editing
- E. RNA interference

Correct answer = D. The production of apolipoprotein (apo) B-48 in the intestine and apo B-100 in liver is the result of RNA editing in the intestine, where a sense codon is changed to a nonsense codon by posttranscriptional deamination of cytosine to uracil. DNA rearrangement and transposition, as well as RNA interference and alternative splicing, do alter gene expression but are not the basis of apo B-48 tissue-specific production.

33.4. Which of the following is a likely consequence of the increased iron accumulation seen in patients with the disease hemochromatosis?

- A. The messenger RNA for the transferrin receptor is stabilized by the binding of iron regulatory proteins to its 3'-iron-responsive elements.
- B. The messenger RNA for the transferrin receptor is not bound by iron regulatory proteins and is degraded.
- C. The messenger RNA for ferritin is not bound by iron regulatory proteins at its 5'-iron-responsive element and is translated.
- D. The messenger RNA for ferritin is bound by iron regulatory proteins and is not translated.
- E. Both B and C are correct.

Correct answer = E. When iron levels in the body are high, as is seen with hemochromatosis, there is increased synthesis of the iron-storage molecule, ferritin, and decreased synthesis of the transferrin receptor (TfR) that mediates iron uptake by cells. These effects are the result of cis-acting iron-responsive elements not being bound by trans-acting iron regulatory proteins, resulting in degradation of the messenger RNA (mRNA) for TfR and increased translation of the mRNA for ferritin.

33.5. Patients with estrogen receptor–positive (hormone responsive) breast cancer may be treated with the drug tamoxifen, which binds the estrogen nuclear receptor without activating it. Which of the following is the most logical outcome of tamoxifen use?

- A. Increased acetylation of estrogen-responsive genes
- B. Increased growth of estrogen receptor–positive breast cancer cells
- C. Increased production of cyclic adenosine monophosphate
- D. Inhibition of the estrogen operon
- E. Inhibition of transcription of estrogen-responsive genes

Correct answer = E. Tamoxifen competes with estrogen for binding to the estrogen nuclear receptor. Tamoxifen fails to activate the receptor, preventing its binding to DNA sequences that upregulate expression of estrogen-responsive genes. Tamoxifen, then, blocks the growth-promoting effects of these genes and results in growth inhibition of estrogen-dependent breast cancer cells. Acetylation increases transcription by relaxing the nucleosome. Cyclic adenosine monophosphate is a regulatory signal mediated by cell-surface rather than nuclear receptors. Mammalian cells do not have operons.

33.6. The ZYA region of the *lac* operon will be maximally expressed if:

- A. cyclic adenosine monophosphate levels are low.
- B. glucose and lactose are both available.
- C. the attenuation stemloop is able to form.
- D. the CAP site is occupied.
- E. the Shine–Dalgarno sequence is not accessible.

Correct answer = D. It is only when glucose is gone, cyclic adenosine monophosphate (cAMP) levels are increased, the cAMP–catabolite activator protein (CAP) complex is bound to the CAP site, and lactose is available that the operon is maximally expressed (induced). If glucose is present, the operon is off as a result of catabolite repression. The *lac* operon is not regulated by attenuation, a mechanism for stopping transcription in some operons such as the *trp* operon.

33.7. X chromosome inactivation is a process by which one of two X chromosomes in mammalian females is condensed and inactivated to prevent overexpression of X-linked genes. What would most likely be true about the degree of DNA methylation and histone acetylation on the inactivated X chromosome?

Cytosines in CpG islands would be hypermethylated, and histone proteins would be deacetylated. Both conditions are associated with decreased gene expression, and both are important in maintaining X inactivation.

