Chapter 19—
Regulation of Gene Expression

John E. Donelson

19.1— Overview

To survive, a living cell must be able to respond to changes in its environment. One of many ways in which cells adjust to changes is to alter expression of specific genes, which, in turn, affects the number of corresponding protein molecules in the cell. This chapter focuses on some of the molecular mechanisms that determine when a given gene will be expressed and to what extent. The attempt to understand how expression of genes is regulated is one of the most active areas of biochemical research today.

It makes sense for a cell to vary the amount of a given gene product available under different conditions. For example, the bacterium *Escherichia coli* (*E. coli*) contains genes for about 3000 different proteins, but it does not need to synthesize all of these proteins at the same time. Therefore it regulates the number of molecules of these proteins that are made. The classic illustration of this phenomenon is the regulation of the number of **β-galactosidase** molecules in the cell. This enzyme converts the disaccharide lactose into the monosaccharides, glucose and galactose. When *E. coli* is growing in a medium containing glucose as the carbon source, β-galactosidase is not required and only about five molecules of the enzyme are present in the cell. When lactose is the sole carbon source, however, 5000 or more molecules of β-galactosidase occur in the cell. Clearly, the bacteria respond to the need to metabolize lactose by increasing the synthesis of β-galactosidase molecules. If lactose is removed from the medium, the synthesis of this enzyme stops as rapidly as it began.

The complexity of eukaryotic cells means that they have even more extensive mechanisms of gene regulation than do prokaryotic cells. The differentiated cells of higher organisms have a much more complicated physical structure and often a more specialized biological function that is determined, again, by the expression of their genes. For example, **insulin** is synthesized in β cells of the pancreas and not in kidney cells even though the nuclei of all cells of the body contain the insulin genes. Molecular regulatory mechanisms facilitate the expression of insulin in pancreas and prevent its synthesis in kidney and other cells. In addition, during development of the organism appearance or disappearance of proteins in specific cell types is tightly controlled with respect to timing and sequence of developmental events.

As expected from the differences in complexities, far more is understood about the regulation of genes in prokaryotes than in eukaryotes. However, studies on the control of gene expression in prokaryotes often provide exciting new ideas that can be tested in eukaryotic systems. Sometimes, discoveries about eukaryotic gene structure and regulation alter the interpretation of data on the control of prokaryotic genes.

Several of the best studied examples of gene regulation in bacteria will be discussed, followed by some illustrations of the organization and regulation of related genes in the human genome. Finally, the use of recombinant DNA techniques to express some human genes of clinical interest will be presented.

19.2— Unit of Transcription in Bacteria: The Operon

The single *E. coli* chromosome is a circular double-stranded DNA molecule of about four million base pairs. Most of the approximately 3000 *E. coli* genes are not distributed randomly throughout this DNA; instead, the genes that code for the enzymes of a specific metabolic pathway are clustered in one region of the DNA. In addition, genes for associated structural proteins, such as the 70 or so proteins that comprise the ribosome, are frequently adjacent to one another. Members of a set of clustered genes are usually coordinately regulated; they are transcribed together to form a "polycistronic" mRNA species that contains the coding sequences for several proteins. The term **operon** describes the

complete regulatory unit of a set of clustered genes. An operon includes the adjacent **structural genes** that code for the related enzymes or associated proteins, a **regulatory gene** or genes that code for regulator protein(s), and **control elements** that are sites on the DNA near the structural genes at which regulator proteins act. Figure 19.1 shows a partial genetic map of the *E. coli* chromosome that gives locations of structural genes of some of the different operons.

When transcription of the structural genes of an operon increases in response to the presence of a specific substrate in the medium, the effect is known as **induction.** The increase in transcription of the β-galactosidase gene when lactose is the sole carbon source is an example of induction. Bacteria also respond to nutritional changes by quickly turning off the synthesis of enzymes that are no longer needed. As will be described below, *E. coli* synthesizes the amino acid tryptophan as the end product of a specific biosynthetic pathway. However, if tryptophan is supplied in the medium, the bacteria do not need to make it themselves, and synthesis of enzymes for this metabolic pathway is stopped. This process is called **repression.** It permits the bacteria to avoid using their energy for making unnecessary and even harmful proteins.

Induction and repression are manifestations of the same phenomenon. In one case the bacterium changes its enzyme composition so that it can utilize a specific substrate in the medium; in the other it reduces the number of enzyme

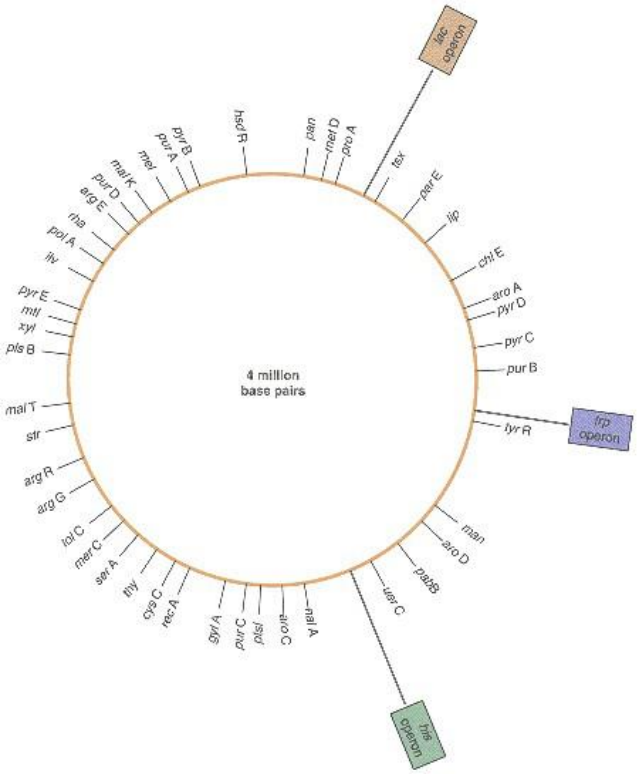


Figure 19.1
Partial genetic map of *E. coli*.
The locations of only a few of the genes identified and mapped in *E. coli* are shown here. Three operons discussed in this chapter are indicated.
Reproduced with permission from Stent, G. S., and Calendar, R. Molecular Genetics, An Introductory Narrative. San Francisco: Freeman, 1978, p. 289; modified from Bachmann, B. J., Low, K. B., and Taylor, A. L. *Bacteriol Rev*. 40:116, 1976.

molecules so that it does not overproduce a specific metabolic product. The signal for each type of regulation is the small molecule that is a substrate for the metabolic pathway or a product of the pathway, respectively. These small molecules are called **inducers** when they stimulate induction and **corepressors** when they cause repression to occur.

Section 19.3 will describe in detail the lactose operon, the best studied example of a set of inducible genes. Section 19.4 will present the tryptophan operon, an example of a repressible operon. Sections 19.5–19.7 will briefly describe some other operons as well as some gene systems in which physical movement of the genes themselves within the DNA (i.e., gene rearrangements) plays a role in their regulation.

19.3— Lactose Operon of *E. Coli*

The lactose operon contains three adjacent structural genes as shown in Figure 19.2. *Lac*Z codes for the enzyme β-galactosidase, which is composed of four identical subunits of 1021 amino acids. *Lac*Y codes for a permease, which is a 275-amino acid protein that occurs in the cell membrane and participates in the transport of sugars, including lactose, across the membrane. The third gene, *lac*A, codes for β-galactoside transacetylase, a 275-amino acid enzyme that transfers an acetyl group from acetyl CoA to β-galactoside. Of these three proteins, only β-galactosidase actually participates in a known metabolic pathway. However, the permease is clearly important in the utilization of lactose since it is involved in transporting lactose into the cell. The acetylation reaction may be associated with detoxification and excretion reactions of nonmetabolized analogs of β-galactosides.

Mutations in *lac*Z or *lac*Y that destroy the function of β-galactosidase or permease prevent cells from cleaving lactose or acquiring it from the medium, respectively. Mutations in *lac*A that destroy transacetylase activity do not seem to have an identifiable effect on cell growth and division. Perhaps there are other related enzymes in the cell that serve as backups for this enzyme, or perhaps it has an unknown function that is required only under certain conditions.

A single mRNA species containing the coding sequences of all three structural genes is transcribed from a promoter that occurs just upstream from the *lac*Z gene. Induction of these three genes occurs during initiation of their

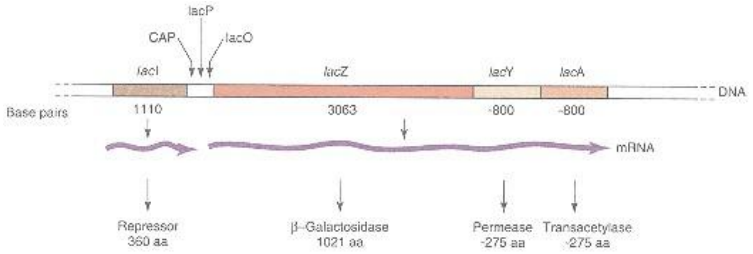


Figure 19.2
Lactose operon of E. coli.
The lactose operon is composed of the *lacI* gene, which codes for a repressor, the control elements of *CAP*, *lacP*, and *lacO*, and three structural genes, *lacZ*, *lacY*, and *lacA*, which code for β-galactosidase, a permease, and a transacetylase, respectively. The *lacI* gene is transcribed from its own promotor. Three structural genes are transcribed from the promoter, *lacP*, to form a polycistronic mRNA from which the three proteins are translated.

transcription. Without the inducer, transcription of the gene cluster occurs only at a very low level. In the presence of the inducer, transcription begins at the **promoter,** called *lac*P, and goes through all three genes to a transcription terminator located slightly beyond the end of *lac*A. Therefore the genes are **coordinately expressed**; either all three are transcribed in unison or none is transcribed.

The presence of three coding sequences on the same mRNA molecule suggests that the relative amounts of the three proteins are always the same under varying conditions of induction. An inducer that causes a high rate of transcription will result in a high level of all three proteins; an inducer that stimulates only a little transcription of the operon will result in a low level of the proteins. The inducer can be thought of as a molecular switch that influences synthesis of the single mRNA species for all three genes. The number of molecules of each protein in the cell may be different, but this does not reflect differences in transcription; it reflects differences in translation rates of the coding sequences or in degradation of the proteins themselves.

The mRNA induced by lactose is very unstable; it is degraded with a half-life of about 3 min. Therefore expression of the operon can be altered very quickly. Transcription ceases as soon as inducer is no longer present, existing mRNA molecules disappear within a few minutes, and cells stop making the proteins.

Repressor of the Lactose Operon Is a Diffusible Protein

The regulatory gene of the lactose operon, *lac*I, codes for a protein whose only function is to control the transcription initiation of the three *lac* structural genes. This regulator protein is called the *lac* **repressor**. The *lac*I gene is located just in front of the controlling elements for the *lac*ZYA gene cluster. However, it is not obligatory that a regulatory gene be physically close to the gene cluster it regulates. In some of the other operons it is not. Transcription of *lac*I is not regulated; instead, this single gene is always transcribed from its own promoter at a low rate that is relatively independent of the cell's status. Therefore affinity of the *lac*I promoter for **RNA polymerase** seems to be the only factor involved in its transcription initiation.

The lac repressor is initially synthesized as a monomer of 360 amino acids and four monomers associate to form a tetramer, the active form of the repressor. Usually there are about 10 tetramers per cell. The repressor has a strong affinity for a specific DNA sequence that lies between *lac*P and the start of *lac*Z. This sequence is called the **operator** and is designated *lac*O. The operator overlaps the promoter somewhat so that presence of repressor bound to the operator physically prevents RNA polymerase from binding to the promoter and initiating transcription.

In addition to recognizing and binding to the lac operator DNA sequence, the repressor also has a strong affinity for the inducer molecules of the *lac* operon. Each monomer has a binding site for an inducer molecule. Binding of inducer to the monomers causes an **allosteric change** in the repressor that greatly lowers its affinity for the operator sequence (Figure 19.3). In other words, when inducer molecules are bound to their sites on the repressor, a conformational change in the repressor occurs that alters the binding site for the operator. The result is that repressor no longer binds to the operator so that RNA polymerase, in turn, can begin transcription from the promoter. A repressor molecule that is already bound to the operator when the inducer becomes available can still bind to inducer so that the repressor–inducer complex immediately disassociates from the operator.

A study of the lactose operon has been greatly facilitated by the discovery that some small molecules fortuitously serve as inducers but are not metabolized by β-galactosidase. Isopropylthiogalactoside **(IPTG)** is one of several thiogalac-

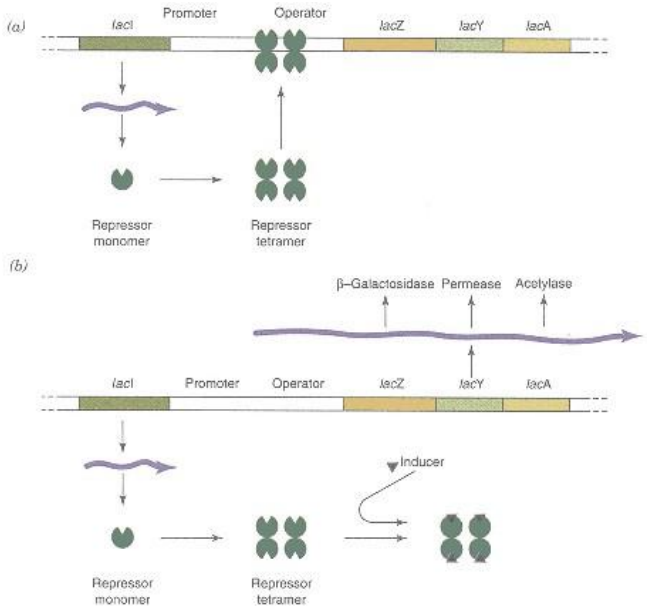


Figure 19.3
Control of *lac* operon.
(a) Repressor tetramer binds to operator and prevents transcription of structural genes.
(b) Inducer binds to repressor tetramer, which prevents repressor from binding to operator. Transcription of three structural genes can occur from the promoter.

tosides with this property. They are called **gratuitous inducers.** They bind to inducer sites on the repressor molecule causing the conformational change but are not cleaved by the induced β-galactosidase. Therefore they affect the system without themselves being altered (metabolized) by it. If it were not possible to manipulate experimentally the system with these gratuitous inducers, it would have been much more difficult to reach our current understanding of the lactose operon in particular and bacterial gene regulation in general.

The product of the *lac*I gene, the repressor protein, acts in trans; that is, it is a diffusible product that moves through the cell to its site of action. Therefore mutations in the *lac*I gene can exert an effect on the expression of other genes located far away or even on genes located on different DNA molecules. *Lac*I mutations can be of several types. One class of mutations changes or deletes amino acids of the repressor that are located in the binding site for the inducer. These changes interfere with interaction between the inducer and the repressor but do not affect the affinity of repressor for the operator. Therefore the repressor is always bound to the operator, even in the presence of inducer, and the *lac*ZYA genes are never transcribed above a very low basal level. Another class of *lac*I mutations changes the amino acids in the operator-binding site of the repressor. Most of these mutations lessen the affinity of the repressor for the operator. This means that repressor does not bind to the operator and *lac*ZYA genes are always being transcribed. These mutations are called **repressor-constitutive mutations** because *lac* genes are permanently turned on. Interestingly, a few rare *lac*I mutants actually increase the affinity of repressor for the operator over that of wild-type repressor. In these cases inducer molecules can still bind to repressor, but they are less effective in releasing repressor from the operator.

Repressor-constitutive mutants illustrate the features of a negative control system. An active repressor, in the absence of an inducer, shuts off the expres-

sion of the *lac* structural genes. An inactive repressor results in the constitutive, unregulated, expression of these genes. It is possible, using the recombinant DNA techniques described in Chapter 18, to introduce into constitutive *lac*I mutant cells a recombinant plasmid containing the wild-type *lac*I gene (but not the rest of the *lac* operon). Therefore these cells have one wild-type and one mutant *lac*I gene and will synthesize both active and inactive repressor molecules. Under these conditions, normal wild-type regulation of the lactose operon occurs. In genetic terms, the wild-type induction is dominant over the mutant constitutivity. This property is the main feature of a negative control system.

Operator Sequence of the Lactose Operon Is Contiguous on DNA with a Promoter and Three Structural Genes

The known control elements in front of the structural genes of the lactose operon are the operator and promoter. The operator was originally identified, like the *lac*I gene, by mutations that affected the transcription of the *lac*ZYA region. Some of these mutations also result in the constitutive synthesis of lac mRNA; that is, they are **operator-constitutive mutations.** In these cases the operator DNA sequence has undergone one or more base pair changes so that the repressor no longer binds as tightly to the sequence. Thus the repressor is less effective in preventing RNA polymerase from initiating transcription.

In contrast to mutations in the *lac*I gene that affect the diffusible repressor, mutations in the operator do not affect a diffusible product. They exert their influence on the transcription of only the three *lac* genes that lie immediately downstream of the operator on the same DNA molecule. This means that if a second lac operon is introduced into a bacterium on a recombinant plasmid, the operator of one operon does not influence action on the other operon. Therefore an operon with a wild-type operator will be repressed under the usual conditions, whereas in the same bacterium a second operon that has an operator-constitutive mutation will be transcribed continuously.

Operator mutations are frequently referred to as **cis-dominant** to emphasize that these mutations affect only adjacent genes on the same DNA molecule and that they are not influenced by the presence in the cell of other copies of the unmutated sequence. Cis-dominant mutations occur in DNA sequences that are recognized by proteins rather than in DNA sequences that code for the diffusible proteins. **Trans-dominant** mutations occur in genes that specify the diffusible products. Therefore cis-dominant mutations also occur in promoter and transcription termination sequences, whereas trans-dominant mutations also occur in the genes for the subunit proteins of RNA polymerase, the ribosomes, and so on.

Figure 19.4 shows the sequence of both the *lac* operator and promoter. The operator sequence has an axis of dyad symmetry. The sequence of the upper strand on the left side of the operator is nearly identical to the lower strand on the right side; only three differences occur between these inverted

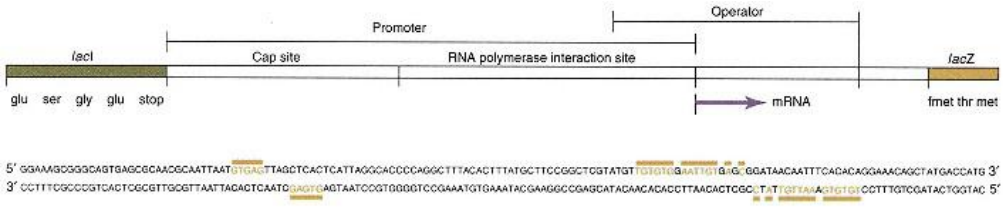


Figure 19.4
Nucleotide sequence of control elements of lactose operon.
The end of the *lacI* gene (coding for the lactose repressor) and the beginning of the *lacZ* gene (coding for *β*-galactosidase) are also shown. Lines above and below the sequence indicate symmetrical sequences within the CAP site and operator.

DNA repeats. This symmetry in the DNA recognition sequence reflects symmetry in the tetrameric repressor. It probably facilitates the tight binding of the subunits of the repressor to the operator, although this has not been definitively demonstrated. A common feature of many protein-binding or recognition sites on double-stranded DNA, including most recognition sites for restriction enzymes, is a **dyad symmetry** in the nucleotide sequence.

The 30 bp that constitute the *lac* operator are an extremely small fraction of the total *E. coli* genome of 4×10^6 bp and occupy an even smaller fraction of the total volume of the cell. Therefore it would seem that the approximately 10 tetrameric repressors in a cell might have trouble finding the *lac* operator if they just randomly diffuse about the cell. Although this remains a puzzling consideration, there are factors that confine the repressor to a much smaller space than the entire volume of the cell. First, it probably helps that the repressor gene is very close to the *lac* operator. This means that the repressor does not have far to diffuse if its translation begins before its mRNA is fully synthesized. Second, and more importantly, the repressor possesses a low general affinity for all DNA sequences. When the inducer binds to the repressor, its affinity for the operator is reduced about a 1000-fold, but its low affinity for random DNA sequences is unaltered. Therefore all of the *lac* repressors of the cell probably spend the majority of the time in loose association with the DNA. As the binding of the inducer releases a represser molecule from the operator, it quickly reassociates with another nearby region of the DNA. Therefore induction redistributes the repressor on the DNA rather than generates freely diffusing repressor molecules. This confines the repressor to a smaller volume within the cell.

Another question is how does lactose enter a *lac*-repressed cell in the first place if the *lac*Y gene product, the permease, is repressed yet is required for lactose transport across the cell membrane? The answer is that even in the fully repressed state, there is a very low basal level of transcription of the *lac* operon that provides five or six molecules of the permease per cell. Perhaps this is just enough to get a few molecules of lactose inside the cell and begin the process.

An even more curious observation is that, in fact, lactose is not the natural inducer of the lactose operon as we would expect. When the repressor is isolated from fully induced cells, the small molecule bound to each repressor monomer is **allolactose,** not lactose. Allolactose, like lactose, is composed of galactose and glucose, but the linkage between the two sugars is different. It turns out that a side reaction of β-galactosidase (which normally breaks down lactose to galactose and glucose) converts these two products to allolactose. Therefore it appears that a few molecules of lactose are taken up and converted by β-galactosidase to allolactose, which then binds to the repressor and induces the operon. Further confirmation that lactose itself is not the real inducer comes from experiments indicating that lactose binding to the purified repressor slightly increases the repressor's affinity for the operator. Therefore, in the induced state, a small amount of allolactose must be present in the cell to overcome this "anti-inducer" effect of the lactose substrate.

Promoter Sequence of Lactose Operon Contains Recognition Sites for RNA Polymerase and a Regulator Protein

Immediately in front of the *lac* operator sequence is the promoter sequence. This sequence contains the recognition sites for two different proteins, RNA polymerase and the **CAP-binding protein** (Figure 19.4). The site at which RNA polymerase interacts with the DNA to initiate transcription has been identified using several different genetic and biochemical approaches. Point mutations in this region frequently affect the affinity to which RNA polymerase will bind the DNA. Deletions (or insertions) that extend into this region also dramatically affect the binding of RNA polymerase to the DNA. The end points of the sequence to which RNA polymerase binds were identified by **DNase protection**

experiments. Purified RNA polymerase was bound to the *lac* promoter region cloned in a bacteriophage DNA or a plasmid, and this protein–DNA complex was digested with DNase I. The DNA segment protected from degradation by DNase was recovered and its sequence determined. The ends of this protected segment varied slightly with different DNA molecules but corresponded closely to the boundaries of the RNA polymerase interaction site shown in Figure 19.4.

The sequence of the RNA polymerase interaction site is not composed of symmetrical elements similar to those described for the operator sequence. This is not surprising since RNA polymerase must associate with the DNA in an asymmetrical fashion for RNA synthesis to be initiated in only one direction from the binding site. However, that portion of the promoter sequence recognized by the CAP-binding protein does contain some symmetry. A **DNA–protein interaction** at this region enhances transcription of the lac operon as described in the next section.

Catabolite Activator Protein Binds at a Site on the Lactose Promoter

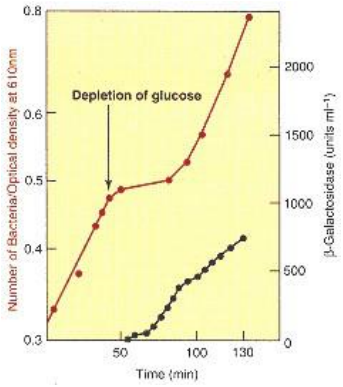


Figure 19.5
Lack of synthesis of β-galactosidase in *E. coli* when glucose is present.
The bacteria are growing in a medium containing initially 0.4 mg mL^{-1} of glucose and 2 mg mL^{-1} lactose. The left-hand ordinate indicates the optical density of the growing culture, an indicator of the number of bacterial cells. The right-hand ordinate indicates the units of β-galactosidase per milliliter. Note that the appearance of β-galactosidase is delayed until the glucose is depleted. Redrawn from Epstein, W., Naono, S., and Gros, F. *Biochem. Biophys. Res. Commun.* 24:588, 1966.

Escherichia coli prefers to use glucose instead of other sugars as a carbon source. For example, if the concentrations of glucose and lactose in the medium are the same, the bacteria will selectively metabolize the glucose and not utilize the lactose. This phenomenon is illustrated in Figure 19.5, which shows that the appearance of β-galactosidase, the *lac*Z product, is delayed until all of the glucose in the medium is depleted. Only then can lactose be used as the carbon source. This delay indicates that glucose interferes with the induction of the lactose operon. This effect is called **catabolite repression** because it occurs during the catabolism of glucose and may be due to a catabolite of glucose rather than glucose itself. An identical effect is exerted on a number of other inducible operons, including the arabinose and galactose operons, which code for enzymes involved in the utilization of various substances as energy sources. It probably is a general coordinating system for turning off synthesis of unwanted enzymes whenever the preferred substrate, glucose, is present.

Catabolite repression begins in the cell when glucose lowers the concentration of intracellular **cyclic AMP** (cAMP). The exact mechanism by which this reduction in the cAMP level is accomplished is not known. Perhaps glucose influences either the rate of synthesis or degradation of cAMP. At any rate, cAMP can bind to another regulatory protein, which has not been discussed yet, called CAP (for **catabolite activator protein**) or CRP (for cAMP receptor protein). CAP is an **allosteric protein,** and when it is combined with cAMP, it is capable of binding to the CAP regulatory site that is at the promoter of the *lac* (and other) operons. The CAP–cAMP complex exerts positive control on the transcription of these operons. Its binding to the CAP site on the DNA facilitates the binding of RNA polymerase to the promoter (Figure 19.6). Alternatively, if the CAP site is not occupied, RNA polymerase has more difficulty binding to the promoter, and transcription of the operon occurs much less efficiently. Therefore, when glucose is present, the cAMP level drops, the CAP–cAMP complex does not form, and the positive influence on RNA polymerase does not occur. Conversely, if glucose is absent, the cAMP level is high, a CAP–cAMP complex binds to the CAP site, and transcription is enhanced.

19.4— Tryptophan Operon of *E. Coli*

Tryptophan is essential for bacterial growth; it is needed for the synthesis of all proteins that contain tryptophan. Therefore, if tryptophan is not present in sufficient amount by the medium, the cell must make it. In contrast, lactose is not absolutely required for the cell's growth; many other sugars can substitute for it, and, in fact, as we saw in the previous section, the bacterium prefers to

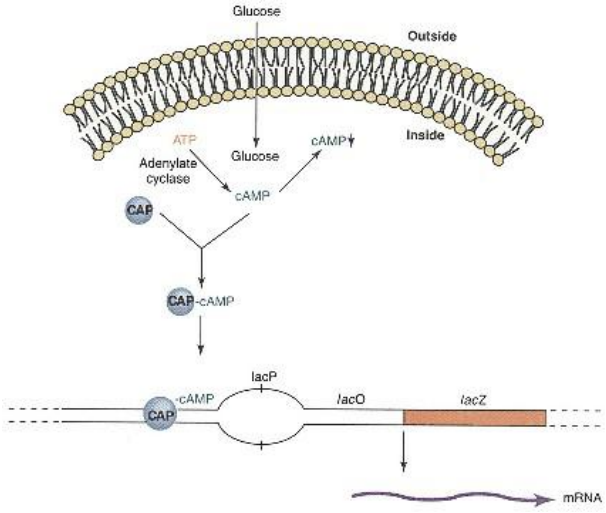


Figure 19.6
Control of *lacP* by cAMP.
A CAP–cAMP complex binds to the CAP site and enhances transcription at *lacP*. Catabolite repression occurs when glucose lowers the intracellular concentration of cAMP. This reduces the amount of the CAP–cAMP complex and decreases transcription from *lacP* and from the promoters of several other operons.

use some of these other sugars for the carbon source. As a result, synthesis of the tryptophan biosynthetic enzymes is regulated differently than synthesis of the proteins encoded by the lactose operon.

Tryptophan Operon Is Controlled by a Repressor Protein

In *E. coli* tryptophan is synthesized from chorismic acid in a five-step pathway that is catalyzed by three different enzymes as shown in Figure 19.7. The **tryptophan operon** contains the five structural genes that code for these three enzymes (two of which have two different subunits). Upstream from this gene cluster is a promoter where transcription begins and an operator to which binds a repressor protein encoded by the unlinked *trp*R gene. Transcription of the lactose operon is generally "turned off" unless it is induced by the small molecule inducer. The tryptophan operon, on the other hand, is always "turned on" unless it is repressed by the presence of a small molecule **corepressor** (a term used to distinguish it from the repressor protein). Hence the lac operon is inducible, whereas the *trp* operon is repressible. When the *trp* operon is being actively transcribed, it is said to be **derepressed;** that is, the trp repressor is not preventing RNA polymerase from binding. This is mechanistically the same as an induced lactose operon in which the *lac* repressor is not interfering with RNA polymerase.

The biosynthetic pathway for tryptophan synthesis is regulated by mechanisms that affect both the synthesis and activity of the enzymes that catalyze the pathway. For example, anthranilate synthetase, which catalyzes the first step of the pathway, is encoded by the *trp*E and *trp*D genes of the trp operon. The number of molecules of this enzyme that is present in the cell is determined by the transcriptional regulation of the *trp*operon. However, the catalytic activity of the existing molecules of the enzyme is regulated by **feedback inhibition.** This is a common short-term means of regulating the first committed step in a metabolic pathway. In this case, tryptophan, the end product of the pathway, can bind to an allosteric site on the anthranilate synthetase and interfere with its catalytic activity at another site. Therefore, as the concentration of tryptophan

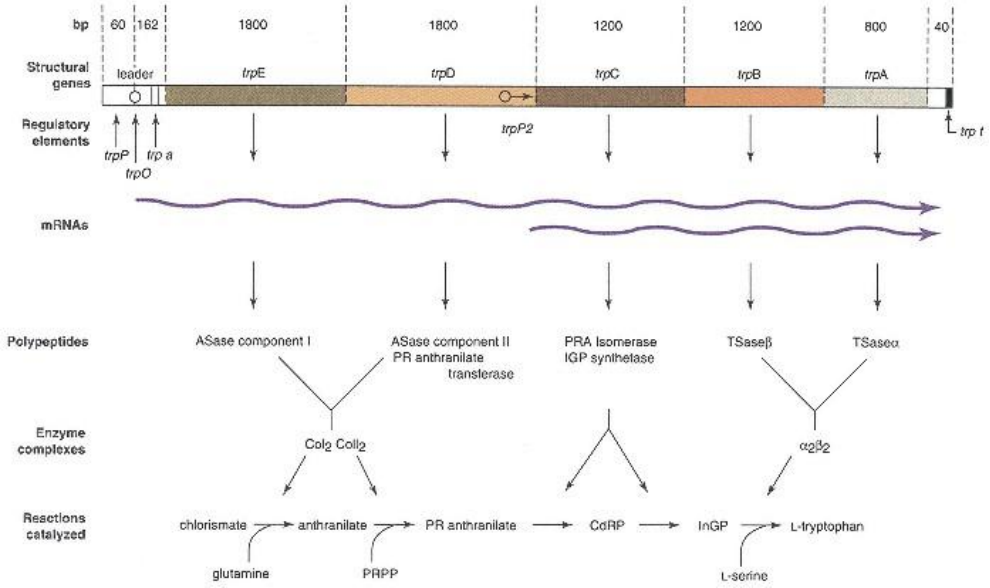


Figure 19.7
Genes of tryptophan operon of *E. coli*.
Regulatory elements are the primary promoter (*trpP*), operator (*trpO*), attenuator (*trp* a), secondary internal promoter (*trpP2*), and terminator (*trp t*) Direction of mRNA synthesis is indicated on the wavy lines representing mRNAs. Col_2 and $CoII_2$ signify components I and II, respectively, of the anthranilate synthetase (ASase) complex; PR-anthranilate is *N*-5 -phosphoribosyl-anthranilate; CdRP is 1-(o-carboxy-phenylamino)-1-deoxyribulose-5-phosphate; InGP is indole-3-glycerol phosphate; PRPP is 5-phosphoribosyl-1-pyrophosphate; and TSase is tryptophan synthetase.
Redrawn from Platt, T. The tryptophan operon. In: J. H. Miller and W. Reznikoff (Eds.), *The Operon.* Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press, 1978, p. 263.

builds up in the cell, it begins to bind to anthranilate synthetase and immediately decreases its activity on the substrate, chorismic acid. In addition, tryptophan also acts as a **corepressor** to shut down the synthesis of new enzyme molecules from the *trp* operon. Thus feedback inhibition is a short-term control that has an immediate effect on the pathway, whereas repression takes a little longer but has the more permanent effect of reducing the number of enzyme molecules.

The *trp* repressor is a tetramer of four identical subunits of about 100 amino acids each. Under normal conditions about 20 molecules of the repressor tetramer are present in the cell. The repressor by itself does not bind to the *trp* operator. It must be complexed with tryptophan in order to bind to the operator and therefore acts *in vivo* only in the presence of tryptophan. This is exactly the opposite of the *lac* repressor, which binds to its operator only in the absence of its small molecule inducer. Interestingly, *trp* repressor also regulates transcription of *trp*R, its own gene. As *trp* repressor accumulates in cells, the repressor-tryptophan complex binds to a region upstream of this gene, turning off its transcription and maintaining the equilibrium of 20 repressors per cell. Another difference from the *lac* operon is that the *trp* operator occurs entirely within the *trp* promoter rather than adjacent to it, as shown in Figure 19.8. The operator sequence is a region of dyad symmetry, and the mechanism

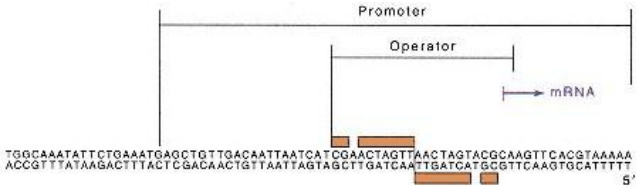


Figure 19.8
Nucleotide sequence of control elements of the tryptophan operon.
Lines above and below sequence indicate symmetrical sequences within operator.

of preventing transcription is the same as in the *lac* operon. Binding of the repressor–corepressor complex to the operator physically blocks the binding of RNA polymerase to the promoter.

Repression results in about a 70-fold decrease in the rate of transcription initiation at the *trp* promoter. (In contrast, the basal level of *lac* gene products is about 1000-fold lower than the induced level.) However, the *trp* operon contains additional regulatory elements that impose further control on the extent of its transcription. One of these additional control sites is a secondary promoter, designated *trp*P2, which is located within the coding sequence of the *trp*D gene (shown in Figure 19.7). This promoter is not regulated by the *trp* repressor. Transcription from it occurs constitutively at a relatively low rate and is terminated at the same location as transcription from the regulated promoter for the whole operon, *trp*P. The resulting transcription product from *trp*P2 is an mRNA that contains the coding sequences for *trp*CBA, the last three genes of the operon. Therefore two polycistronic mRNAs are derived from the *trp* operon, one containing all five structural genes and one possessing only the last three genes. Under conditions of maximum repression the basal level of mRNA coding sequence for the last three genes is about five times higher than the basal mRNA level for the first two genes.

The reason for a second internal promoter is unclear. Perhaps the best alternative comes from the observation that three of the five proteins do not contain tryptophan; only the *trp*B and *trp*C genes contain the single codon that specifies tryptophan. Therefore, under extreme tryptophan starvation, these two proteins would not be synthesized, which would prevent the pathway from being activated. However, since both of these genes lie downstream of the unregulated second promoter, their protein products will always be present at the basal level necessary to maintain the pathway.

Tryptophan Operon Has a Second Control Site: The Attenuator Site

Another important control element of the *trp* operon not present in the *lac* operon is the **attenuator** site (Figure 19.9). It lies within 162 nucleotides between the start of transcription from *trp*P and the initiator codon of the *trp*E gene. Its existence was first deduced by the identification of mutations that mapped in this region and increased transcription of all five structural genes. Within the 162 nucleotides, called the **leader sequence,** are 14 adjacent codons that begin with a methionine codon and end with an in-phase termination codon. These codons are preceded by a canonical ribosome-binding site and

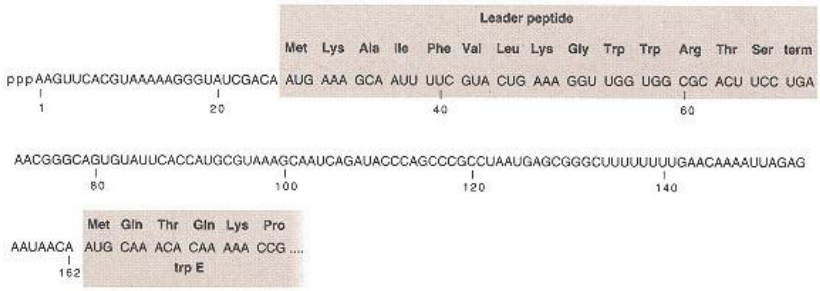


Figure 19.9
Nucleotide sequence of leader RNA from *trp* operon.
The 14 amino acids of the putative leader peptide are indicated over their codons.
Redrawn with permission from Oxender, D. L., Zurawski, G., and Yanofsky, C.
Proc. Natl. Acad. Sci. USA 76:5524, 1979.

could potentially specify a 14-residue leader peptide. This peptide has never been detected in bacterial cells, perhaps because it is degraded very rapidly. The ribosome-binding site does function properly when its corresponding DNA sequence is ligated upstream of a structural gene using recombinant DNA techniques.

The attenuator region provides RNA polymerase with a second chance to stop transcription if the trp enzymes are not needed by the cell. In the presence of tryptophan, it acts like a rho-independent transcription termination site to produce a short 140-nucleotide transcript. In the absence of tryptophan, it has no effect on transcription, and the entire polycistronic mRNA of the five structural genes is synthesized. Therefore, at both the operator and attenuator, tryptophan exerts the same general influence. At the operator it participates in repressing transcription, and at the attenuator it participates in stopping transcription by those RNA polymerases that have escaped repression. It has been estimated that attenuation has about a 10-fold effect on transcription of the trp structural genes. When multiplied by the 70-fold effect of derepression at the operator, about a 700-fold range exists in the level at which the *trp* operon can be transcribed.

The molecular mechanism by which transcription is terminated at the attenuator site is a marvelous example of cooperative interaction between bacterial transcription and translation to achieve desired levels of a given mRNA. The first hints that ribosomes were involved in the mechanism of attenuation came from the observation that mutations in the gene for **tRNA^{Trp} synthetase** (the enzyme that charges the tRNA with tryptophan) or the gene for an enzyme that modifies some bases in the tRNA prevent attenuation. Therefore a functional tRNA^{Trp} must participate in the process.

The leader peptide (Figure 19.9) of 14 residues contains two adjacent tryptophans in positions 10 and 11. This is unusual because tryptophan is a relatively rare amino acid in *E. coli*. It also provides a clue about the involvement of tRNA^{Trp} in attenuation. If the tryptophan in the cell is low, the amount of charged tRNA^{Trp} will also be low and the ribosomes may be unable to translate through the two *trp* codons of the leader peptide region. Therefore they will stall at this place in the leader RNA sequence.

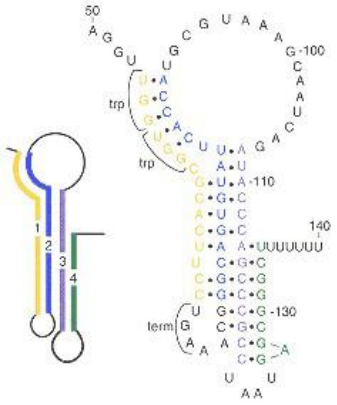


Figure 19.10
Schematic diagram showing the proposed secondary structures in trp leader RNA from *E. coli*.
Four regions can base pair to form three stem-and-loop structures. These are shown as 1–2, 2–3, and 3–4.
Reproduced with permission from Oxender, D. L., Zurawski, G., and Yanofsky, C. *Proc. Natl. Acad. Sci. USA* 76:5524, 1979.

It turns out that the RNA sequence of the attenuator region can adopt several possible secondary structures (Figure 19.10). The position of the ribosome within the leader peptide-coding sequence determines the secondary structure that will form. This secondary structure, in turn, is recognized (or sensed) by the RNA polymerase that has just transcribed through the attenuator coding region and is now located a small distance downstream. The RNA secondary structure that forms when a ribosome is not stalled at the *trp* codons is a termination signal for the RNA polymerase. Under these conditions the cell does not need to make tryptophan, and transcription stops after the synthesis of a 140-nucleotide transcript, which is quickly degraded. On the other hand, the secondary structure that results when the ribosomes are stalled at the *trp* codons is not recognized as a termination signal, and the RNA polymerase continues on into the *trpE* gene. Figure 19.11 shows these different secondary structures in detail.

The structure in Figure 19.11*a* shows the situation when a ribosome does not stall at the two tandem *trp* codons, UGG-UGG, near the beginning of region 1, but instead moves on to region 2. When the ribosome is in region 2, regions 1 and 2 cannot base pair but regions 3 and 4 can form base pairs, resulting in a hairpin loop followed by eight U residues, a structure common to sequences that signal transcription termination. Thus when the leader RNA sequence is being synthesized in the presence of sufficient tryptophan (and charged tryptophanyl-tRNA^{Trp}), it is likely that a loop between regions 3 and 4 will occur and be recognized as a signal for termination by the RNA polymerase.

A different structure occurs if the ribosome is stalled at the *trp* codons and

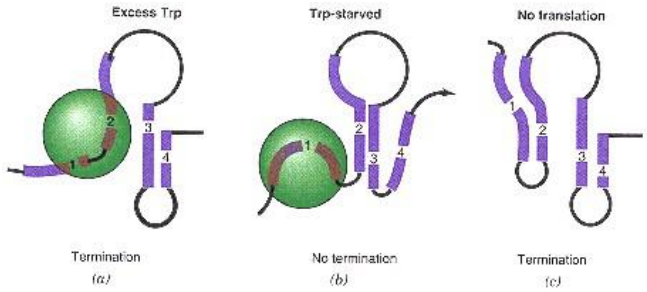


Figure 19.11
Schematic diagram showing the model for attenuation in the *trp* operon of *E. coli*.
(a) Under conditions of excess tryptophan, the ribosome (green sphere) translating the newly transcribed leader RNA will synthesize the complete leader peptide. During this synthesis the ribosome will bind to regions 1 and 2 of the RNA and prevent formation of stem and loop 1–2 or 2–3. Stem and loop 3–4 will be free to form and signal the RNA polymerase molecule (not shown) to terminate transcription.
(b) Under conditions of tryptophan starvation, tryptophanyl-tRNA^{Trp} will be limiting, and the ribosome will stall at the adjacent *trp* codons at the beginning of region 1 in the leader peptide-coding region. Because region 1 is bound to the ribosomes, stem and loop 2–3 will form, excluding formation of stem and loop 3–4, which is required as the signal for transcription termination. Therefore RNA polymerase will continue transcription into the structural genes.
(c) Under conditions in which the leader peptide is not translated, stem and loop 1–2 will form, preventing formation of stem and loop 2–3, and thereby permit formation of stem and loop 3–4. This will signal transcription termination.
Reproduced with permission from Oxender, D. L., Zurawski, G., and Yanofsky, C. *Proc. Natl. Acad. Sci. USA* 76:5524, 1979.

region 1 is prevented from base pairing with region 2 (Figure 19.11*b*). Under these circumstances, region 2 now can base pair with region 3. This region 2 and 3 hairpin ties up the sequence complementary to region 4, so that region 4 remains single stranded. Therefore the region 3 and 4 hairpin loop that serves as the termination signal does not form, and the RNA polymerase continues on with its transcription. Thus for transcription to proceed past the attenuator, region 1 must be prevented from pairing with region 2. This is accomplished if the ribosome stalls in region 1 due to an insufficient amount of charged tryptophan-tRNA for translation of the leader peptide to continue beyond two *trp* codons. When this happens, region 1 is bound within the ribosome and cannot pair with region 2. Since regions 2 and 3 are synthesized before region 4, they, in turn, will base pair before region 4 appears in the newly transcribed RNA. Therefore region 4 remains single stranded, the termination hairpin does not form, and RNA polymerase continues transcription into the structural genes.

Since the two *trp* codons occur in region 1, if the ribosome happens to stall at an earlier codon in the leader sequence, it will have little effect on attenuation. For example, starvation for lysine, valine, or glycine would be expected to reduce the amount of the corresponding charged tRNA and stall the ribosome at that codon, but a deficiency in these amino acids has no effect on transcription of the *trp* operon. An exception is arginine whose codon occurs immediately after the *trp* codons. Starving for arginine does attenuate transcription termination somewhat, probably because of ribosome stalling at this codon, but to less of an extent than a deficiency in tryptophan.

Cis-acting mutations in the attenuator region support this alternate hairpin model. Most of these mutations result in increased transcription because they disrupt base pairing in the double-stranded portion of the termination hairpin

Operon	Leader peptide sequence	Regulatory amino acids
his	Met-Thr-Arg-Val-Gln-Phe-Lys-His-His-His-His-His-His-His-Pro-Asp	His
pheA	Met-Lys-His-Ile-Pro-Phe-Phe-Phe-Ala-Phe-Phe-Phe-Thr-Phe-Pro	Phe
thr	Met-Lys-Arg-Ile-Ser-Thr-Thr-Ile-Thr-Thr-Thr-Ile-Thr-Ile-Thr-Thr-Gly-Asn-Gly-Ala-Gly	Thr Ile
leu	Met-Ser-His-Ile-Val-Arg-Phe-Thr-Gly-Leu-Leu-Leu-Leu-Asn-Ala-Phe-Ile-Val-Arg-Gly-Arg-Pro-Val-Gly-Gly-Ile-Gln-His	Leu
ilv	Met-Thr-Ala-Leu-Leu-Arg-Val-Ile-Ser-Leu-Val-Val-Ile-Ser-Val-Val-Val-Ile-Ile-Ile-Pro-Pro-Cys-Gly-Ala-Ala-Leu-Gly-Arg-Gly-Lys-Ala	Leu, Val, Ile

Figure 19.12
Leader peptide sequences specified by biosynthetic operons of *E. coli*.
All contain multiple copies of amino acid(s) synthesized by enzymes coded for by operon.

and render it less stable. Some mutations, however, increase termination at the attenuator. One of these interferes with base pairing between regions 2 and 3, allowing region 3 to be available for pairing with region 4 even when region 1 is bound to a stalled ribosome. Another mutation occurs in the AUG initiator codon for the leader peptide so that the ribosome cannot begin its synthesis.

Transcription Attenuation Is a Mechanism of Control in Operons for Amino Acid Biosynthesis

Attenuation is a common phenomenon in bacterial gene expression; it occurs in at least six other operons that code for enzymes catalyzing amino acid biosynthetic pathways. Figure 19.12 shows the corresponding **leader peptide** sequences specified by each of these operons. In each case, the leader peptide contains several codons for the amino acid product of the biosynthetic pathway. The most extreme case is the 16-residue leader peptide of the histidine operon that contains seven contiguous histidines. Starvation for histidine results in a decrease in the amount of histidinyl-tRNA^{His} and a dramatic increase in transcription of the *his* operon. As with the *trp* operon, this effect is diminished by mutations that interfere with the level of charged histidinyl-tRNA^{His}. Furthermore, the nucleotide sequence of the attenuator region suggests that ribosome stalling at the histidine codons also influences the formation of alternate hairpin loops, one of which resembles a termination hairpin followed by several U residues. In contrast to the *trp* operon, transcription of the *his* operon is regulated entirely by attenuation; it does not possess an operator that is recognized by a repressor protein. Instead, the **ribosome** acts rather like a positive regulator protein, similar to the cAMP–CAP complex discussed with the *lac* operon. If the ribosome is bound to (i.e., stalled at) the attenuator site, then transcription of the downstream structural genes is enhanced. If the ribosome is not bound, then transcription of these genes is greatly reduced.

Transcription of the other operons shown in Figure 19.12 can be attenuated by more than one amino acid. For example, the *thr* operon is attenuated by either threonine or isoleucine; the *ilv* operon is attenuated by leucine, valine, or isoleucine. This effect can be explained in each case by stalling of the ribosome at the corresponding codon, which, in turn, interferes with the formation of a termination hairpin. Although not proved, it is possible that in the cases of the longer leader peptides, stalling at more than one codon is necessary to achieve maximal transcription through the attenuation region.

19.5— Other Bacterial Operons

Synthesis of Ribosomal Proteins Is Regulated in a Coordinated Manner

Many other bacterial operons have been studied and found to possess the same general regulatory mechanisms as the *lac*, *trp*, and *his* operons, as discussed

Operon	Regulator protein	Proteins specified by the operon
Spc	S8	L14-L24-L5-S14-S8-L6-L18-S5-L15-L30
S10	L4	S10-L3-L2-L4-L23-S19-L22-S3-S17-L16-L29
str	S7	S12-S7-EF•G-EF•Tu
α	S4	S13-S11-S4-α-L17
L11	L1	L11-L1
rif	L10	L10-L7-β-

Figure 19.13
Operons containing genes for ribosomal proteins *E. coli*.
Genes for the protein components of the small (S) and large (L) ribosomal subunits of *E. coli* are clustered on several operons. Some of these operons also contain genes for RNA polymerase subunits α, β, and , and protein synthesis factors EF·G and EF·Tu. At least one of the protein products of each operon usually regulates expression of that operon (see text).

in Section 19.4. However, each operon has evolved its own distinctive quirks. For example, one interesting group of operons are those containing the structural genes for the 70 or more proteins that comprise the ribosome (Figure 19.13). Each ribosome contains one copy of each **ribosomal protein** (except for protein L7-L12, which is probably present in four copies). Therefore all 70 proteins are required in equimolar amounts, and it makes sense that their synthesis is regulated in a coordinated fashion. Characterization of this set of operons is not yet complete, but six operons, containing about one-half of the ribosomal protein genes, occur in two major gene clusters. One cluster contains four adjacent operons (*str, Spc, S10,* and *a*), and the other two operons are near each other elsewhere in the *E. coli* chromosome. There is no obvious pattern to distribution of these genes among different operons. Some operons contain genes for proteins of just one ribosomal subunit; others code for proteins of both subunits. In addition to structural genes for ribosomal proteins, these operons also contain genes for other (related) proteins. For example, *str* operon contains genes for the two soluble **translation elongation factors,** EF·Tu and EF·G, as well as genes for some proteins in the 30S ribosomal subunit. The α operon has genes for proteins of both 30S and 50S ribosomal subunits plus a gene for one of the subunits of RNA polymerase. The *rif* operon has genes for two other protein subunits of RNA polymerase and genes for ribosomal proteins.

A common theme among the six ribosomal operons is that their expression is regulated by one of their own structural gene products; that is, they are **self-regulated.** The precise mechanism of this self-regulation varies considerably with each operon and is not yet understood in detail. However, in some cases the regulation occurs at the level of translation, not transcription as discussed for the *lac* and *trp* operons. After the polycistronic mRNA is made, the "regulatory" ribosomal protein binds to this mRNA and determines which regions, if any, are translated. In general, the ribosomal protein that regulates expression of its own operon, or part of its own operon, is a protein that is associated with one of the ribosomal RNAs (rRNAs) in the intact ribosome. This ribosomal protein has a high affinity for the rRNA and a lower affinity for one or more regions of its own mRNA. Therefore a competition between the rRNA and the operon's mRNA for binding with the ribosomal protein occurs. As the ribosomal protein accumulates to a higher level than the free rRNA, it binds to its own mRNA and prevents the initiation of protein synthesis at one or more of the coding sequences on this mRNA (Figure 19.14). As more ribosomes are formed, the excess of this particular ribosomal protein is used up and translation of its coding sequence on the mRNA can begin again.

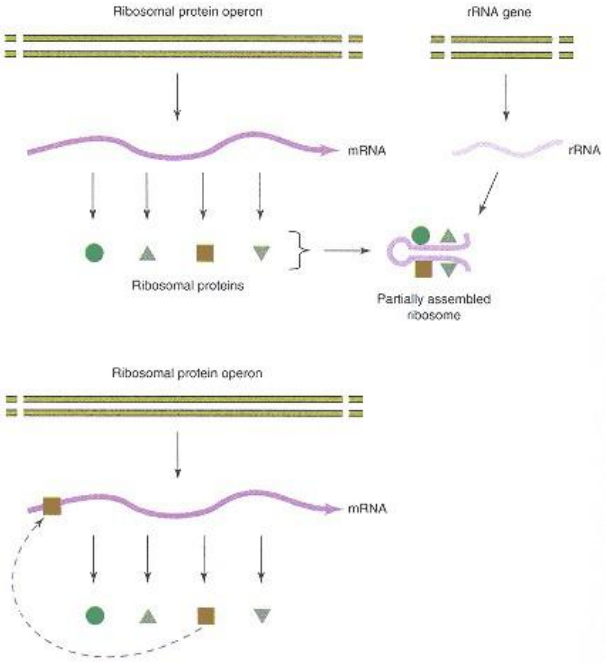


Figure 19.14
Self-regulation of ribosomal protein synthesis.
If free rRNA is not available for assembly of new ribosomal subunits, individual ribosomal proteins bind to polycistronic mRNA from their own operon, blocking further translation.

Stringent Response Controls Synthesis of rRNAs and tRNAs

Bacteria have several ways in which to respond molecularly to emergency situations; that is, times of **extreme general stress.** One of these situations is when the bacterium does not have a sufficient pool of amino acids to maintain protein synthesis. Under these conditions the cell invokes what is called the **stringent response,** a mechanism that reduces the synthesis of the rRNAs and tRNAs about 20-fold. This places many of the activities within the cell on hold until conditions improve. The mRNAs are less affected, but there is also about a three-fold decrease in their synthesis.

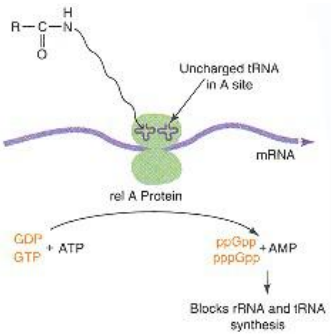


Figure 19.15
Stringent control of protein synthesis in *E. coli*.
During extreme amino acid starvation, an uncharged tRNA in the A site of the ribosome activates the relA protein to synthesize ppGpp and pppGpp, which, in turn, are involved in decreasing transcription of the genes coding for rRNAs and tRNAs.

The stringent response is triggered by the presence of an uncharged tRNA in the A site of the ribosome. This occurs when the concentration of the corresponding charged tRNA is very low. The first result, of course, is that further peptide elongation by the ribosome stops. This event causes a protein called the **stringent factor,** the product of the *relA* gene, to synthesize **guanosine tetraphosphate** (ppGpp) and **guanosine pentaphosphate** (pppGpp), from ATP and GTP or GDP as shown in Figure 19.15. Stringent factor is loosely associated with a few, but not all, ribosomes of the cell. Perhaps a conformational change in the ribosome is induced by occupation of the A site by an uncharged tRNA, which, in turn, activates the associated stringent factor. The exact functions of ppGpp and pppGpp are unknown. However, they seem to inhibit transcription initiation of the rRNA and tRNA genes. In addition they affect transcription of some operons more than others.

CLINICAL CORRELATION 19.1

Transmissible Multiple Drug Resistances

Pathogenic bacteria are becoming increasingly resistant to a large number of antibiotics, which is viewed with alarm by many physicians. Many cases have been documented in which a bacterial strain in a patient being treated with one antibiotic suddenly became resistant to that antibiotic and, simultaneously, to several other antibiotics even though the bacterial strain had never been previously exposed to these other antibiotics. This occurs when the bacteria suddenly acquire from another bacterial strain a plasmid that contains several different transposons, each containing one or more antibiotic-resistance genes. Examples include the genes encoding β-lactamase, which inactivates penicillins and cephalosporins, chloramphenicol acetyltransferase, which inactivates chloramphenicol, and phosphotransferases, which modify aminoglycosides such as neomycin and gentamycin.

Neu, H. C. The crisis in antibiotic resistance. *Science* 257: 1064, 1992.

19.6— Bacterial Transposons

Transposons Are Mobile Segments of DNA

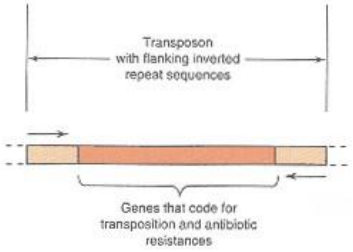


Figure 19.16
General structure of transposons.
Transposons are relatively rare mobile segments of DNA that contain genes coding for their own rearangement and (usually) genes that specify resistance to various antibiotics.

So far we have only discussed the regulation of bacterial genes whose locations are fixed in the chromosome. Their positions relative to the neighboring genes do not change. The vast majority of bacterial genes are of this type. In fact, genetic maps of *E. coli* and *Salmonella typhimurium* are quite similar, indicating the lack of much evolutionary movement of most genes within the bacterial chromosome. There is a class of bacterial genes, however, in which newly duplicated gene copies "jump" to another genomic site with a frequency of about 10^{-7} per generation, the same rate as spontaneous point mutations occur. The mobile segments of DNA containing these genes are called **transposable elements** or **transposons** (Figure 19.16). Transposons were first detected as rare insertions of foreign DNA into structural genes of bacterial operons. Usually, these insertions interfere with the expression of the structural gene into which they have inserted and all downstream genes of the operon. This is not surprising since they can potentially destroy the translation reading frame, introduce transcription termination signals, affect the mRNA stability, and so on. Many transposons and the sites into which they insert have been isolated using recombinant DNA techniques and have been extensively characterized. These studies have revealed many interesting features about the mechanisms of transposition and the nature of genes located within transposons.

Transposons vary tremendously in length. Some are a few thousand base pairs and contain only two or three genes; others are many thousands of base pairs long, containing several genes. Several small transposons can occur within a large transposon. All active transposons contain at least one gene that codes for a **transposase,** an enzyme required for the transposition event. Often they contain genes that code for resistance to antibiotics or heavy metals. Most transpositions involve generation of an addition copy of the transposon and insertion of this copy into another location. The original transposon copy is the same after the duplication as before; that is, the donor copy is unaffected by insertion of its duplicate into the recipient site. Transposons contain short inverted **terminal repeat sequences** that are essential for the insertion mechanism, and in fact these inverted repeats are often used to define the two boundaries of a transposon. The multiple target sites into which most transposons can insert seem to be fairly random in sequence; other transposons have a propensity for insertion at specific "hot spots." The duplicated transposon can be located in a different DNA molecule than its donor. Frequently, transposons are found on plasmids that pass from one bacterial strain to another and are the source of a suddenly acquired resistance to one or more antibiotics by a bacterium (Clin. Corr. 19.1).

As with bacterial operons, each transposon or set of transposons has its own distinctive characteristics. The well-characterized transposon *Tn3* will be discussed as an example of their general properties.

Tn3 *Transposon Contains Three Structural Genes*

The **transposon Tn3** has been cloned using recombinant DNA techniques and its complete sequence determined. It contains 4957 base pairs including 38 base pairs at one end that occur as an inverted repeat at the other end (Figure 19.17). Three genes are present in *Tn3*. One gene codes for the enzyme β-lactamase, which hydrolyzes ampicillin and renders the cell resistant to this antibiotic. The other two genes, *tnp*A and *tnp*R, code for a transposase and a repressor protein, respectively. The transposase has 1021 amino acids and binds to single-stranded DNA. Little else is known about its action, but it is thought

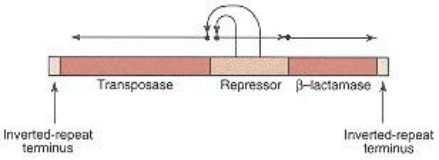


Figure 19.17
Functional components of the transposon *Tn3*.
Genetic analysis shows there are at least four kinds of regions: the inverted repeat termini; a gene for the enzyme *β*-lactamase, which confers resistance to ampicillin and related antibiotics; a gene encoding an enzyme required for transposition (transposase); and a gene for a repressor protein that controls transcription of genes for transposase and for repressor itself. The horizontal arrows indicate direction in which DNA of various regions is transcribed.
Redrawn from Cohen, S. N., and Shapiro, J. A. *Sci. Am.* 242:40, 1980. W. H. Freeman and Company, Copyright © 1980.

to recognize the repetitive ends of the transposon and to participate in the cleavage of the recipient site into which the new transposon copy inserts. The *tnp*R gene product is a protein of 185 amino acids. In its role as a repressor it controls transcription of both the transposase gene and its own gene. The *tnp*A and *tnp*R genes are transcribed divergently from a 163 base pair control region located between the two genes that is recognized by the repressor. The *tnp*R product also participates in the recombination process that results in the insertion of the new transposon. Transcription of the ampicillin-resistance gene is not affected by the *tnp*R gene product.

Mutations in the transposase gene generally decrease the frequency of Tn3 transposition, demonstrating its direct role in the transposition process. Mutations that destroy the repressor function of the *tnp*R product cause an increased frequency of transposition. These mutations derepress the *tnp*A gene, resulting in more molecules of the transposase, which increases the formation of more transposons. They also derepress the *tnp*R gene but, since the repressor is inactive, this has no effect on the system.

When a transposon, containing its terminal inverted repeats, inserts into a new site, it generates short (5–10 bp) direct repeats of the sequences at the recipient site that flank the new transposon. This is due to the mechanism of recombination that occurs during the insertion process (Figure 19.18). The first step is the generation of staggered nicks at the recipient sequence. These staggered single-strand, protruding 5 ends then join covalently to the inverted repeat ends of the transposon. The resulting intermediate resembles two replicating forks pointing toward each other and separated by the length of the transposon. The replication machinery of the cell fills in the gaps and continues the divergent elongation of the two primers through the transposon region. This ultimately results in two copies of the transposon sequence. Reciprocal recombination within the two copies regenerates the original transposon copy at its original position and completes the process of forming a new copy at the recipient site that is flanked by direct repeats of the recipient sequence.

The practical importance of transposons located on plasmids has taken on increased significance for the use of antibiotics in treatment of bacterial infections. Plasmids that have not been altered for experimental use in the laboratory usually contain genes that facilitate their transfer from one bacterium to another. As the plasmids transfer (e.g., between different infecting bacterial strains), their transposons containing **antibiotic-resistance genes** are moved into new bacterial strains. Once inside a new bacterium, the transposon can be duplicated onto the chromosome and become permanently established in that cell's lineage.

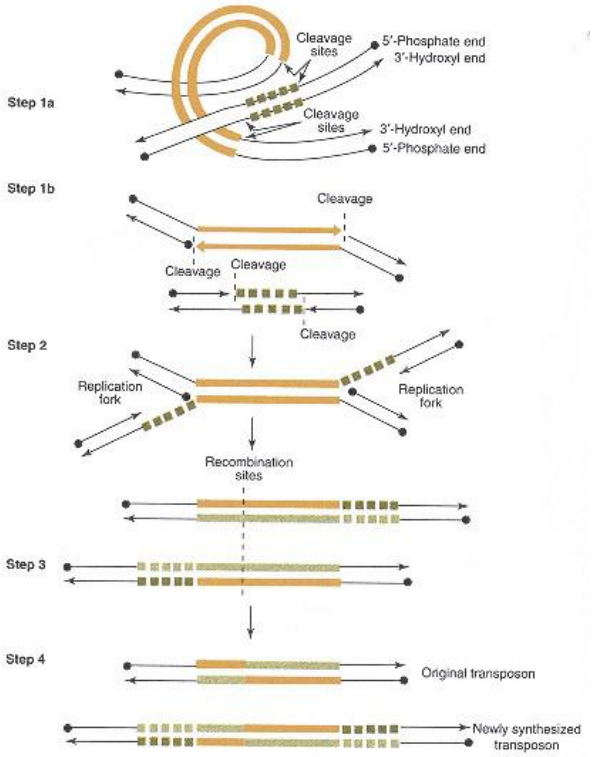


Figure 19.18
Proposed molecular pathway for transposition and chromosome rearrangements.
Donor DNA, including the transposon, shown in red, recipient DNA contains small light green. The pathway has four steps, beginning with staggered, single-strand cleavages (Step 1a) at each end of the transposable element and at each end of the "target" nucleotide sequence to be duplicated. The cleavages expose (Step 1b) the DNA strand ends involved in the next step: the joining of DNA strands from donor and recipient molecules in such a way that the double-stranded transposable element has a DNA replication fork at each end (Step 2). DNA synthesis (Step 3) replicates transposon (red bars) and target sequence (light green squares), accounting for the observed duplication. This step forms two new complete double-stranded molecules; each copy of the transposable element joins a segment of the donor molecule and a segment of the recipient molecules. (Copies of the element serve as linkers for the recombination of two unrelated DNA molecules.) In the final Step 4, reciprocal recombination between copies of the transposable element inserts the element at a new genetic site and regenerates the donor molecule.
Redrawn from Cohen, S. N., and Shapiro, J. A. *Sci. Am.* 252:40, 1980.

The result is that more and more pathogenic bacterial strains become resistant to an increasing number of antibiotics.

19.7— Inversion of Genes in *Salmonella*

A different mechanism of differential gene regulation has been discovered for one set of genes in *Salmonella*. Similar control mechanisms exist for the expression of other genes in other prokaryotes (e.g., a bacteriophage called).

Bacteria move by waving their **flagella** that are composed predominantly of subunits of a protein called flagellin. Many *Salmonella* species possess two different flagellin genes and express only one of these genes at a time. Bacteria are said to be in phase 1 if they are expressing the H1 flagellin gene and in phase 2 if they are expressing the H2 flagellin gene. A bacterial clone in one phase switches to the other phase about once every 1000 divisions. This switch is called **phase variation,** and its occurrence is controlled at the level of transcription of *H1* and *H2* genes.

Organization of the flagellin genes and their regulatory elements are shown in Figure 19.19. A 995-bp segment of DNA flanked by 14-bp repeats is adjacent to the *H2* gene and a *rhl* gene that codes for a repressor of *H1*. The *H2* and

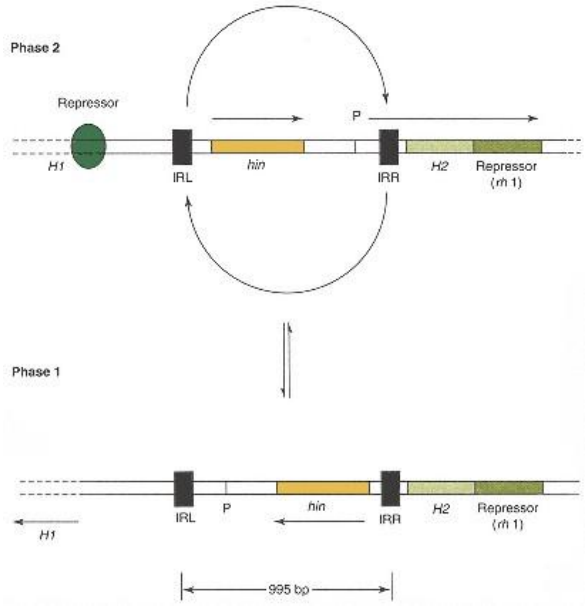


Figure 19.19
Organization of the flagellin genes of *Salmonella*.
Orientation of a 995-bp DNA segment flanked by 14-bp inverted repeats (IRL and IRR) controls the expression of *H1* and *H2* flagellin genes. In phase 2, transcription initiates at promoter P within the invertible segment and continues through *H2* and *rh1* genes. In phase 1, the orientation is reversed so that transcription of *H2* and *rh1* genes does not occur.

rhl genes are coordinately transcribed. Therefore, when *H2* is expressed, the repressor is also made and turns off *H1* expression. When H2 protein and the repressor are not made, the H1 gene is derepressed and *H1* synthesis occurs.

The promoter for the operon containing *H2* and *rhl* lies near one end of the 995-bp segment, just inside one copy of the 14-bp repeats. This segment can undergo inversions between the 14-bp repeats. In one orientation of the segment, the promoter is upstream of the *H2–rhl* transcription unit; in the other orientation it points toward the opposite direction so that *H2* and *rhl* are not transcribed. In addition to containing this promoter, the invertible segment of DNA possesses the *hin* gene whose product is an enzyme that catalyzes the inversion event itself. The *hin* gene seems to be transcribed constitutively at a low rate. Mutations in *hin* reduce the rate of inversion by 10,000-fold. Therefore phase variation is controlled by physical inversion of the segment of DNA that removes a promoter from its position in front of the *H2–rhl* operon. When the promoter is in the opposite direction, it presumably still initiates transcription, but the fate of that RNA is unknown. It does not initiate transcription of the *H1* that maps in this direction. That gene apparently has its own promoter controlled directly by the *rhl* repressor.

Inversion of the *hin* segment probably occurs via recombination between the 14-bp inverted repeats that is similar to recombination events involved in the transposition of a transposon. In fact, transposons do invert relative to their flanking sequences in a fashion exactly analogous to the *hin* inversion. Furthermore, the amino acid sequence of the *hin* product shows considerable similarity to that of the *tnp*R product of the *Tn3* transposon, which participates in the integration of the transposon into a new site. Thus it is possible, and even likely, that the two processes are evolutionarily related.

19.8— Organization of Genes in Mammalian DNA

The past 20 years have seen a virtual explosion of new information about the organization, structure, and regulation of genes in eukaryotic organisms. The reason for this enormous increase in our knowledge about eukaryotic genes has been the concurrent development of recombinant DNA techniques and DNA sequencing techniques (Chapter 18). Experiments undreamed of a few years ago are now routine accomplishments.

The human haploid genome contains 3×10^9 bp of DNA, about 1000 times more DNA than the *E. coli* chromosome. All available evidence suggests that each of the 23 haploid chromosomes in the human genome has a single unique DNA molecule. Since the distance between two adjacent base pairs is 3.4×10^{-10} meters (3.4 Å), if these 23 human chromosomal DNA molecules were stretched out end-to-end, they would extend about 1 meter. Each mammalian cell contains virtually a complete copy of this genome, and all except the haploid germline cells contain two copies.

Different types of mammalian cells express widely different proteins even though each contains the same complement of genes. In addition, widely different patterns of protein synthesis occur at different developmental stages of the same type of cells. Therefore extremely intricate and complicated mechanisms of regulation for these genes must exist, and, in fact, these mechanisms are not understood for even one mammalian gene to the extent that they are understood for many bacterial operons. Despite the great advances of the past 20 years, our understanding of gene regulation in mammals, and indeed all eukaryotes, remains fragmentary at best and probably is still very naive.

Only a Small Fraction of Eukaryotic DNA Codes for Proteins

It was appreciated even before the advent of recombinant DNA methodology that eukaryotic cells, including mammalian cells, contain far more DNA than seems necessary to code for all of the required proteins. Furthermore, organisms that appear rather similar in complexity can have a several-fold difference in cellular DNA content. A housefly, for example, has about six times the cellular DNA content of a fruitfly. Some plant cells have almost ten times more DNA than human cells. Therefore DNA content does not always correlate with the complexity and diversity of functions of the organism.

It is difficult to obtain an accurate estimate of the number of different proteins, and therefore genes, in a mammalian cell or in the entire mammalian organism. However, nucleic acid hybridization procedures indicate that a maximum of 5000–10,000 different mRNA species may be present in a mammalian cell at a given time. Most of these mRNAs code for proteins that are common to many cell types. Therefore a generous estimate is that there are approximately 100,000 genes for the entire mammalian genome. If the average coding sequence is 1500 nucleotides (specifying a 500 amino acid protein), this accounts for 5% of the mammalian genome. DNA regulatory elements, repetitive genes for rRNAs, and so on may account for another 5–10%. However, as much as 85–90% of the mammalian genome may not have a direct genetic function. This remarkable conclusion is in contrast to the bacterial genome in which virtually all of the DNA is consumed by genes and their regulatory elements.

Eukaryotic Genes Usually Contain Intervening Sequences (Introns)

As discussed in Chapter 16, coding sequences **(exons)** of eukaryotic genes are frequently interrupted by intervening sequences or **introns** that do not code for a product. These introns are transcribed into a **precursor RNA** species found in the nucleus and are removed by **RNA splicing** events during the processing of the nuclear precursor RNA to the mature mRNA in the cytoplasm.

The number and length of the introns in a gene can vary tremendously. Histone genes and interferon genes lack introns; they contain a continuous coding sequence for the protein as do bacterial genes. The mammalian collagen gene, on the other hand, has more than 50 different introns that collectively consume 90% of the gene. The largest human gene discovered to date is 2400 kb, or more than half the size of the entire *E. coli* genome of 4000 kb. This gene contains 79 introns of about 30-kb average size and encodes a 427-kDa muscle protein called **dystrophin** (Figure 19.20). Despite the fact that dystrophin is a very large protein, the dystrophin gene's introns consume more than 99% of the gene's length. Mutations in this huge dystrophin gene are responsible for **Duchenne/Becker muscular dystrophy** (see Clin. Corr. 19.2). On the basis of the many mammalian genes analyzed to date, it appears that most have three or four introns and that the presence of 50 or more introns in a single gene represents an extreme case. Nevertheless, introns of genes clearly account for some of the "excess" DNA present in eukaryotic genomes.

The significance of introns and their potential biological functions, if any, are the subject of much speculation and experimentation. In a few genes, including those for the **α- and β-globin subunits** of hemoglobin (see below), introns separate the coding regions for functional domains of the protein. In many other genes, however, no obvious correlation exists between the intron positions of a gene and the three-dimensional domains of its encoded protein. In fact, the number of introns in a given gene sometimes is not the same in different mammalian species, or even within a single species. For example, the rat haploid genome has two insulin genes, one with two introns and one with a single intron. The haploid genomes of other rodents have a single insulin gene with two introns.

One widely quoted hypothesis for the possible function of introns is that they may have served to facilitate the mixing and matching of exons during the course of evolution so that occasionally new protein-encoding genes are created, which provide a selective advantage for the organism. Some circumstantial evidence exists to support this possibility. For example, **chicken collagen** has a larger number of repeating Gly-X-Y triplets and most of the exons in its genes are multiples of 9 bp (i.e., 45, 54, 99, 108, or 162 bp per exon) beginning with a glycine codon and ending with a Y codon. Thus the collagen gene may have evolved via multiple duplications of an exon–intron unit. Genes of unicellular lower eukaryotes, such as **yeast,** have either no introns or a small number of introns that tend to be short compared to introns of higher eukaryotes. Perhaps these lower eukaryotes, which reproduce much faster than do higher organisms, have to be more efficient in their DNA and RNA metabolism and

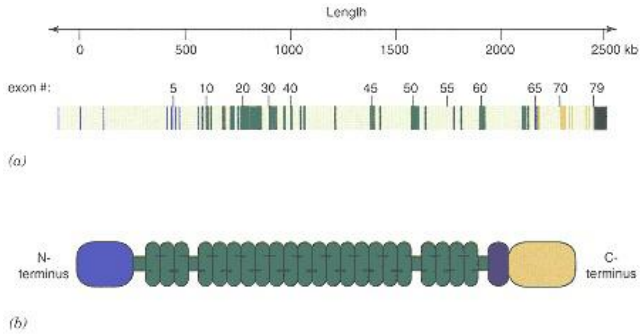


Figure 19.20
Human dystrophin gene and its protein.
(a) The 79 exons (dark thin vertical lines) of human dystrophin gene span 2.4×10^6 bp (2400 kb), more than one-half the length of the *E. coli* genome. The average dystrophin exon is 140 bp and the average dystrophin intron (light gray background regions) is more than 30,000 bp.
(b) Dystrophin (427 kDa) has 3685 amino acids. It contains an actin-binding domain blue, 24 tandem repeats of about 109 amino acids that likely form a rod-like domain (green), a cysteine-rich domain (purple), and a C terminus that may associate with the membrane (red).
Redrawn from Ahn, A. H., and Kunkel, L. M. *Nature Genetics* 3:283, 1993.

CLINICAL CORRELATION 19.2

Duchenne/Becker Muscular Dystrophy and the Dystrophin Gene

Both Duchenne muscular dystrophy (DMD) and the milder Becker muscular dystrophy (BMD) are inherited as X-linked recessive diseases. They result in degenerative disorders of skeletal muscle and are the most common of all lethal neuromuscular genetic diseases, affecting 1 in 3500 males. They are associated with abnormally high levels of serum creatine kinase levels from birth. Although most afflicted males inherit the defect from their unaffected, heterozygous mother, 30% of the cases exhibit no previous family history and appear to be "spontaneous" new mutations in the germline of either the mother or her parents. Both forms of muscular dystrophy are caused by defects in the dystrophin gene on the X chromosome. This gene is huge and complicated. It has 79 exons and spans 2.4 $\times 10^6$ bp and encodes a membrane-associated cytoskeleton protein. Its expression is regulated in a cell-specific and developmentally controlled manner from at least five different promoters. Many mutations responsible for DMD and BMD are large deletions that remove one or more of the 79 exons, but the size of the deletion does not necessarily correlate with the severity of the disorder. In DMD patients, dystrophin is undetectable or absent, whereas in BMD patients, it is reduced or altered. Genetic, biochemical, and anatomical studies suggest that dystrophin may serve diverse roles in many other tissues besides muscle. It is hoped that future studies of dystrophin may lead to an understanding of the cause and perhaps a rational treatment for muscular dystrophy.

Ahn, A. H., and Kunkel, L. M. The structural and functional diversity of dystrophin. *Nature Genetics* 3:283, 1993.

cannot tolerate large numbers of large introns. In many ways, however, introns remain as big an enigma as when first discovered.

19.9— Repetitive DNA Sequences in Eukaryotes

Another curiosity about mammalian DNA, and the DNA of most higher organisms, is that, in contrast to bacterial DNA, it contains repetitive sequences in addition to single copy sequences. This repetitive DNA falls into two general classes—**highly repetitive** simple sequences and **moderately repetitive** longer sequences of several hundred to several thousand base pairs.

Importance of Highly Repetitive Sequences Is Unknown

The highly repetitive sequences range from 5 to about 300 bp and occur in tandem. Their contribution to the total genomic size is extremely variable, but in most organisms they are repeated millions of times and in a few organisms they consume 50% or more of the total DNA. These highly repetitive sequences are sometimes called **satellite DNAs** because when total DNA isolated from a eukaryote is sheared slightly and centrifuged in a CsCl gradient, they can be separated as "satellites" of the bulk of the DNA on the basis of their differing buoyant densities. They are concentrated primarily at the **centromeres** and to a lesser extent at **telomeres** (i.e., ends of chromosomes). Figure 19.21 shows the three main repeat units of the highly repetitive sequences at the chromosomal centromeres of the fruitfly, *Drosophila virilis*. Repeats of these three sequences of 7 bp comprise 41% of the organism's DNA. They are obviously related evolutionarily since two of the repeats can be derived from the third by a single base pair change. Relatively little transcription occurs from the highly repetitive sequences, and their biological importance remains, for the most part, a mystery (see Clin. Corr. 19.3). Those repetitive sequences that occur near the telomeres are probably required for the replication of the ends of the linear DNA molecules. The ones at the centromeres might play a structural role since these regions attach to the **microtubules** of the **mitotic spindle** during chromosome pairing and segregation in mitosis and meiosis. Highly repetitive sequences occur in human DNA at both centromeres and telomeres but their repeat units at centromeres are longer and more variable in sequence than those of *Drosophila virilis* shown in Figure 19.21.

A Variety of Repeating Units Are Defined as Moderately Repetitive Sequences

The moderately repetitive sequences consist of a large number of different sequences repeated to such different extents that it is somewhat misleading to group them under one heading. Some are clustered in one region of the genome;

Genome (%)	Number of copies in genome	Predominant sequence
25	1×10^7	5 -ACAAACT- 3 3 -TGTTTGA- 5
8	3.6×10^6	5′ -ATAAACT- 3′ 3′ -TATTTGA- 5′
8	3.6×10^6	5′ -ACAAATT- 3′ 3′ -TGTTTAA- 5′

Figure 19.21
Main repeat units of repetitive sequences of the fruitfly *Drosophila virilis*.
Approximately 41% of genomic DNA of *Drosophila virilis* is comprised of three related repeat sequences of 7 bp. The bottom two sequences differ from the top sequence at one base pair shown in box.

CLINICAL CORRELATION 19.3

Huntington's Disease and Trinucleotide Repeat Expansions

Huntington's disease is an autosomal dominant neurodegenerative disorder characterized by increasing behavioral disturbance, involuntary movements, cognitive impairment, and dementia. It can be inherited from either parent. Disease onset often does not occur until age 40 and death results 10–15 years later from aspiration, trauma, or pneumonia. The defective gene on chromosome 4 responsible for the disease is dominant over the normal gene, suggesting the defect causes the gene's protein to gain a deleterious function. This gene encodes a large protein called "huntingtin" that contains 3144 amino acids found in many tissues but whose function is unknown. Near the beginning of the gene is a run of CAGs that encodes a polyglutamine tract in huntingtin. The length of this polyglutamine tract is 11–34 in normal individuals and 37–121 in Huntington's disease patients. The larger the number of repeats, the sooner the onset of the disease. Furthermore, the child of a parent with an abnormally large number of repeats will often have an even larger number of repeats, resulting in a "genetic anticipation" of the disease. Neither the cause of the trinucleotide repeat expansions nor the abnormal function of huntingtin with an expanded polyglutamine is known. However, at least seven other neurological disorders are caused by trinucleotide repeat expansions in other genes, including X-linked spinal and bulbar muscular atrophy, fragile X syndrome, and myotonic dystrophy. The reason for this neuronal toxicity is currently the subject of intense research. These diseases can be diagnosed molecularly by tests based on the polymerase chain reaction.

La Spada, A. R., Paulson, H. L., and Fischbeck, K. H. Trinucleotide repeat expansion in neurological disease. *Ann. Neurol.* 36:814, 1994.

many are scattered throughout the DNA. Some moderate repeats are several thousand base pairs in length; other repeats come in a unit size of only a hundred base pairs. Sometimes the sequence is highly conserved from one repeat to another; in other cases, different repeat units of the same basic sequence will have undergone considerable divergence. Two examples from the human genome will be described.

In mammalian cells the 18S, 5.8S, and 28S rRNAs are transcribed as a single precursor transcript that is subsequently processed to yield the mature rRNAs. In humans the length of this precursor is 13,400 nucleotides, about one-half of which is comprised of the three mature rRNA sequences. Several **posttranscriptional cleavage** steps remove the extra sequences from the ends and the middle of the precursor RNA, releasing the mature rRNA species. DNA that contains the rRNA genes is a moderately repetitive sequence of about 43,000 bp of which 30,000 bp are nontranscribed spacer DNA. Clusters of this entire DNA unit occur on five chromosomes. In total, there are about 280 repeats of this unit, which comprise about 0.3% of the total genome (Figure 19.22). The 5S rRNA genes are repeated about 2000 times but in different clusters. The need for so many rRNA genes is because the rRNAs are structural RNAs. Each transcript from the gene yields only one copy of each rRNA molecule. On the other hand, each mRNA molecule derived from a ribosomal protein gene can be translated repeatedly to give many protein molecules.

In contrast to tandemly repetitive rRNA genes clustered at a few chromosomal sites, most moderately repetitive sequences in the mammalian genome do not code for a stable gene product and are interspersed with nonrepetitive sequences that occur only once or a few times in the genome. The average size of these interspersed repetitive sequences is about 300 bp. Almost one-half of these sequences are members of a general family of moderately repetitive sequences called the **Alu family** because they can be cleaved by the restriction enzyme *Alu*I. There are about 300,000 Alu sequences scattered throughout the human haploid genome (on the high side of being moderately repetitive). Individual members are related in sequence but are frequently not identical. Their average homology with a consensus sequence is about 87%.

Additional repeat symmetry occurs within an Alu sequence. The sequence appears to have arisen by tandem duplication of a 130-bp sequence with a 31-bp insertion in one of the two adjacent repeats. Some members of the Alu family resemble bacterial transposons in that they are flanked by short direct repeats. This does not prove that an Alu repeat can be duplicated and transposed to another site like true transposons, but it suggests that such events may occur.

The biological function of Alu sequences is unknown. One suggestion is that they serve as multiple origins for the DNA replication during S phase, but more sequences occur than seem necessary for this function. Alu sequences appear in the introns of some genes and are transcribed as part of large precursor RNAs in which the Alu sequences are removed during RNA splicing. Other Alu sequences are transcribed into small RNA molecules whose function is unknown. All mammalian genomes appear to have a counterpart to the human interspersed Alu sequence family although the size of the repeat and its distribution can vary considerably between species.

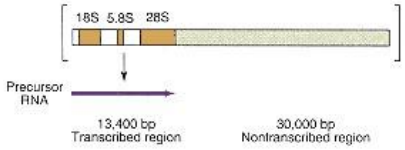


Figure 19.22
Repetitive sequence in human DNA for rRNA.
In human cells a single transcription unit of 13,400 nucleotides is processed to yield the 18S, 5.8S, and 28S rRNAs. About 280 copies of the corresponding rRNA genes are clustered on five chromosomes. Each repeat contains a nontranscribed spacer region of about 30,000 bp.

19.10— Genes for Globin Proteins

Recombinant DNA Technology Has Been Used to Clone Genes for Many Eukaryotic Proteins

Many mammalian structural genes that have been cloned by recombinant DNA techniques specify proteins that either occur in large quantity in a specific cell type, such as the **globin subunits** in the red blood cell, or after induction of a specific cell type, for example, **growth hormone** or **prolactin** in the pituitary. As a result, more is known about the regulation of these genes than of other genes whose protein products occur at lower levels in many different cell types. Increasingly, however, information is being gained about mammalian genes for "rare" proteins with low abundances in the cell. We will discuss organization, structure, and regulation of the related members of two gene families—the genes for the globin subunits and the growth hormone-like proteins.

The first step in characterizing a eukaryotic gene is usually to use recombinant DNA techniques to clone a **complementary DNA** (cDNA) copy of that gene's corresponding mRNA. In fact, this is the reason that the most extensively studied mammalian genes code for the major proteins of specific cells; a large fraction of the total mRNA isolated from these cells codes for protein of interest. **Hemoglobin** is comprised of two **α-globin** subunits (141 amino acids) and two **β-globin** subunits (146 amino acids). Almost all of the mRNA isolated from immature red cells **(reticulocytes)** codes for these two subunits of hemoglobin.

There are several experimental variations of the procedure for synthesizing double-stranded cDNA copies of isolated mRNA *in vitro*. As discussed in Chapter 18, many different plasmid and viral DNA vectors are available for cloning the (passenger) cDNA molecules. Figure 19.23 shows one protocol for constructing and cloning cDNAs prepared from mRNA of reticulocytes.

A synthetic **oligonucleotide** composed of 12–18 residues of deoxythymidine is hybridized to the 3 -polyadenylate tail of the mRNA and serves as a primer for **reverse transcriptase,** an enzyme that copies an RNA sequence into a DNA strand in the presence of the four deoxynucleoside triphosphates. The resulting RNA–DNA heteroduplex is treated with NaOH, which degrades the RNA strand and leaves the DNA strand intact. The 3 end of the remaining DNA strand can then fold back and serve as a primer for initiating synthesis of a second DNA strand at random locations by reverse transcriptase, the same enzyme used to synthesize the first strand. The hairpin loop is then nicked by **S1 nuclease,** an enzyme that cleaves single-stranded DNA but has little activity against double-stranded DNA. The ends of the resulting double-stranded cDNAs are ligated to small synthetic "linker" oligonucleotides that contain the recognition site for the restriction enzyme *Hind*III. Digestion of the resulting DNA with *Hind*III generates DNA fragments that contain *Hind*III-specific ends. These fragments can be ligated into the *Hind*III site of a plasmid, and when the resulting circular "recombinant" DNA species are incubated with *E. coli* in the presence of cations such as calcium or rubidium, a few molecules will be taken up by the bacteria. The incorporated **recombinant DNAs** will be replicated and maintained in the progeny of the original transformed bacterial cell.

The collection of cloned cDNAs synthesized from the total mRNA in a given tissue or cell type is called a cDNA library, for example, a liver cDNA library or a reticulocyte cDNA library. Since most of the mRNAs of a reticulocyte code for either α- or β-globin, it is relatively easy to identify these globin cDNAs in a reticulocyte cDNA library using procedures discussed in Chapter 16. Once identified, the nucleotide sequences of the cDNAs can be determined to confirm that they do code for the known amino acid sequences of the α- and β-globins. In cases in which the amino acid sequence of the protein is not known, other procedures (sometimes immunological) are used to confirm the identification of the desired cDNA clone.

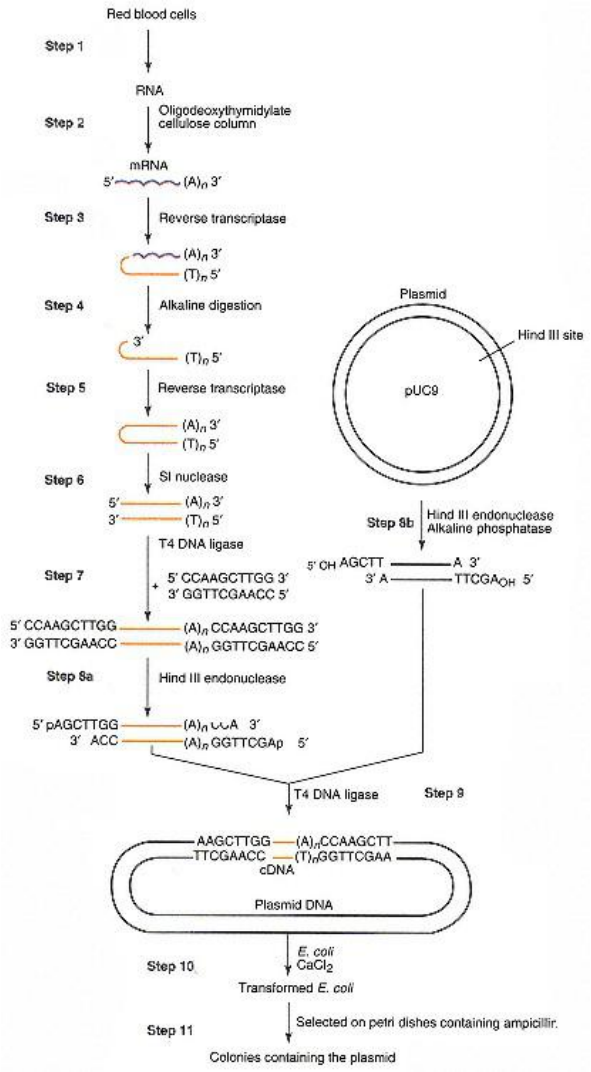


Figure 19.23
Cloning of globin cDNA.
Step 1: Total RNA is extracted from red blood cells.
Step 2: The total RNA is passed through an oligode-oxythymidylate cellulose column, which separates polyadenylated mRNA (see Chapter 15) from rRNA and tRNA. Polyadenylated mRNA of red blood cells contains predominantly hemoglobin mRNA.
Step 3: The mRNA is reverse-transcribed into first-strand cDNA using reverse transcriptase, the viral enzyme that synthesizes DNA from RNA templates (see Chapter 15).
Step 4: The mRNA is hydrolyzed with alkali whereas the DNA is unaffected.
Step 5: The single-stranded cDNA is converted into double-stranded DNA by reverse transcriptase.
Step 6: The resulting double helix contains a single-stranded hairpin loop that is removed by S1 nuclease, an enzyme that hydrolyzes single-stranded DNA.
Step 7: The cDNA is now a double helix with A-T base pairs at one end. To generate cohesive ends for the ligation of this cDNA into a plasmid, a chemically synthesized decanucleotide is attached to both ends using DNA ligase from bacteriophage T4. This decanucleotide contains the sequence recognized by *Hind*III restriction nuclease.
Step 8a: Treatment with *Hind*III produces a cDNA molecule with *Hind*III cohesive ends.
Step 8b: The plasmid pUC9, which contains an ampicillin-resistance gene, is cleaved with *Hind*III and exposed to bacterial alkaline phosphatase, an enzyme that removes the phosphates from the cleaved 5 -terminal ends of the plasmids at the *Hind*III site. This prevents the cleaved plasmid from recircularizing without the insertion of the cDNA.
Step 9: The linear plasmid and the cDNA molecules are mixed with T4 DNA ligase, and circular, dimeric, "recombinant" DNA molecules are formed.
Step 10: This ligation mixture is used to transform *E. coli*.
Step 11: Individual *E. coli* cells that take up the plasmid are selected by their ability to grow on ampicillin. The globin cDNA is confirmed by determining the nucleotide sequence of the small DNA fragment released from the plasmid DNA by *Hind*III; if the observed nucleotide sequences corresponded to those expected based on the known amino acid sequence of α- and β-globin, then the cDNA is identified.

Comparison of the α- and β-globin cDNA sequences with the corresponding globin genes, which have also been cloned using recombinant DNA techniques, reveals that all members of both sets of genes contain two introns at approximately the same positions relative to the coding sequences (Figure 19.24). The α (and α-like) genes have an intron of 95 bp between codons 31 and 32 and a second intron of 125 bp between codons 99 and 100. The β (and β-like) genes have introns of 125–150 bp and 800–900 bp located between codons 30

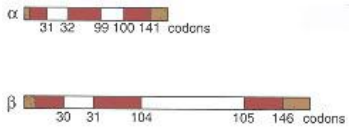


Figure 19.24
Structures of human globin genes.
Structures for the human α-like and β-like globin genes are drawn to approximate scale. Red rectangles and open rectangles represent exons and introns, respectively. Gray rectangles indicate the (5) upstream and (3) downstream nontranslated regions in the DNA. The α-like globin genes contain introns of approximately 95 and 125 bp, located between codons 31 and 32, and 99 and 100, respectively. The β-like globin genes contain introns of approximately125–150 and 800–900 bp, located between codons 30 and 31, and 104 and 105, respectively.

and 31 and codons 104 and 105, respectively. Introns separate the coding sequences of different functional domains of a few proteins, including the globins. The coding region between the two globin introns specifies the region of the protein that interacts with the heme group. The final coding region (after the second intron) encodes the region of the protein that serves as the interface with the opposite subunit, that is, the α-globin-β-globin interaction. This separation of the coding sequences for functional domains of a protein by introns is not a general phenomenon, however. The positioning of introns in other genes seems to bear little relationship to the final three-dimensional structure of the encoded protein.

Different α-like and β-like globin subunits are synthesized at different developmental stages. These developmentally distinct subunits have slightly different amino acid sequences and oxygen affinities but are closely related. In humans there are two α-like chains—that is, , which is expressed in the embryo during the first 8 weeks, and α itself, which replaces in the fetus and continues through adulthood. There are four β-like chains. Epsilon ( ) and γ are expressed in the embryo, γ in the fetus, and plus β in the adult.

Each of the different globin chains is coded by at least one gene in the haploid genome. The α-like genes are clustered on the short arm of human **chromosome 16,** and the β-like genes are clustered on the short arm of **chromosome 11.** The gene organization within these two clusters is shown in Figure 19.25. The genes within both clusters are positioned relative to one another in the order of both their transcriptional direction and their developmental expression; that is, 5 –embryonic–fetal-adult–3 .

The α-gene cluster spans about 28 kb and includes three functional genes and two **pseudogenes.** The functional genes are the embryonic gene and two α genes, α1 and α2, that code for identical α-globin proteins but have different 3 untranslated regions. The two pseudogenes, and , occur between the and α1 genes. They have sequences very similar to the functional genes, but various mutations prevent them from coding for an active globin subunit. Pseudogenes are common in eukaryotic genomes. They do not seem to be deleterious and probably arose via a duplication of a segment of DNA followed by mutations.

The β-gene cluster encompasses about 60 kb and has five active genes and one pseudogene. Of the five functional genes, two are for the γ subunit and specify proteins that differ only at position 136, which is a glycine in the G variant and an alanine in the A variant. Only a single haploid gene exists for the -, -, and β-globin subunits. Alu repetitive sequences and other moderately repetitive sequences are scattered between some genes of the α- and β-gene clusters.

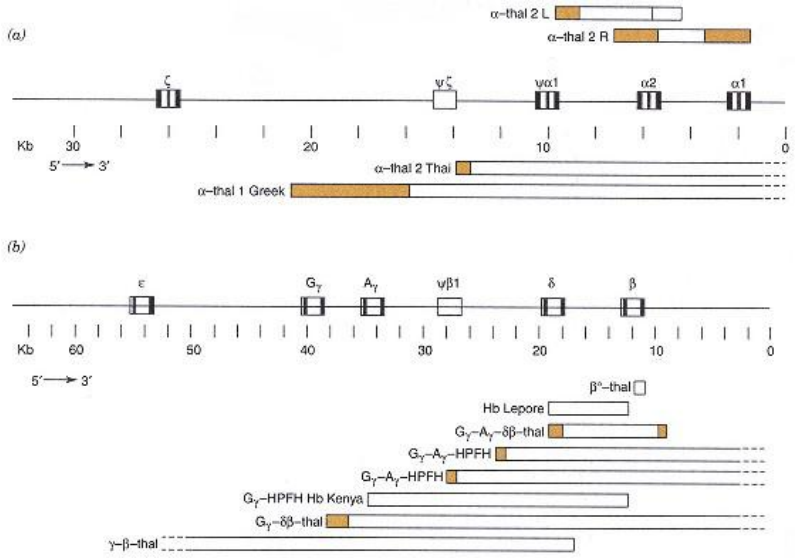


Figure 19.25
Gene organization forr α-like and β-like genes of human hemoglobin.
(a) Linkage of human α-like globin genes on chromosome 16 and locations of some known deletions within α-like gene cluster. The positions of adult (α1, α2) and embryonic ( ) α-like globin genes and two pseudogenes (, 1) are shown. Pseudogenes have mutations that prevent the formation of functional proteins from them. For each functional gene the black and white boxes represent exons and introns, respectively. Horizontal arrow indicates the direction of transcription of each gene. The locations of DNA deletions associated with the leftward and rightward types of α-thalassemia 2 are indicated above the linkage map by the rectangles labeled α-thal 2 L and α-thal 2 R. Red areas at the ends of these rectangles indicate the deletion end points have not been mapped precisely. Locations of deletions associated with two cases of α-thalassemia 1 (α-thal 1 Thai and α-thal 1 Greek) are shown below the linkage map. The light green areas and dashed lines indicate uncertainties in the left and right endpoints, respectively, of each deletion.
(b) Linkage of the human β-like globin genes on chromosome 11 and locations of deletions within the β-like gene cluster. The positions of the embryonic ( ), fetal (G_γ, A_γ), and adult (δ, β) β-like globin genes and one β-like pseudogene ($\psi\beta$1) are shown. For each functional gene the black and white boxes represent the exons and introns, respectively. The locations of various known deletions within the gene cluster are shown below the map. Open rectangles represent areas known to be deleted; Red areas and dashed lines indicate that the endpoints of the deletion have not been determined. For $\delta\beta$-thalassemia and hereditary persistence of fetal hemoglobin (HPFH), the type of fetal globin chain produced (G_γ and/or A_γ) is indicated in the name of each syndrome (e.g., in (G_γ-A_γ-$\delta\beta$-thalassemia, the G_γ- and A_γ-globin chains are produced).
Redrawn from Maniatis, T., Fritsch, E. F. Lauer, J., and Lawn, R. M. Annu. Rev. Genet. 14:145, 1980. Copyright © 1980 by Annual Reviews, Inc.; and from Karlsson, S., and Nienhuis, A. W. *Annu. Rev. Biochem.* 54:1071, 1985. Copyright © 1985 by Annual Reviews, Inc.

Other mammalian species often have a different number of globin-like genes within the two clusters. For example, rabbits have only four β-like genes, goats have seven, and mice have as many as nine. Some of these additional genes are pseudogenes.

Many patients have been identified who have abnormalities in hemoglobin structure or expression. In many cases the precise molecular defect responsible for these abnormalities is known. The two that have been the most extensively studied are **sickle cell anemia** and a family of diseases collectively called **thalassemias.**

CLINICAL CORRELATION 19.4

Prenatal Diagnosis of Sickle Cell Anemia

Sickle cell anemia can be diagnosed from fetal DNA obtained by amniocentesis. This genetic disease is caused by a single base pair change that converts a glutamate to a valine in the sixth position of β-globin. In the normal β-globin gene, the sequence that specifies amino acids 5, 6, and 7 (Pro-Glu-Glu) is CCT-GAG-GAG. In a heterozygous carrier of sickle cell anemia, this sequence is CCT-GTG-GAG. An A in the middle of the sixth codon has been changed to a T. The restriction enzyme *Mst*II recognizes and cleaves the sequence CCT-GAG-G, which is present at this position in normal DNA but not the mutated DNA. Therefore digestion of fetal DNA with *Mst*II followed by the Southern blot technique (see p. 774) using β-globin cDNA as the radioactive probe reveals whether this restriction site is present in one or both allelic copies of the gene. If it is absent in both copies, the fetus will be homozygous for the sickle trait; if it is missing in only one copy, the fetus will be heterozygous for the trait. The difference in restriction enzyme patterns observed between individuals is often called a restriction fragment length polymorphism (RFLP). Polymerase chain reaction methods can be used to amplify the desired chromosomal DNA region and greatly speed up the RFLP analysis.

Other methods are necessary if the disease mutation does not cause a change in a restriction site or is not linked to an RFLP. For example, the DNA carrying the mutation can be amplified by the polymerase chain reaction, and the alleles can be detected by hybridization with allele-specific oligonucleotides (ASOs). Two ASOs differing at usually one nucleotide are made so that one ASO matches the normal allele perfectly while the other ASO matches the abnormal allele. Hybridization conditions are used in which only the ASO matching perfectly remains bound to the DNA.

Sickle Cell Anemia Is Due to a Single Base Pair Change

A **single base pair change** within the coding region for the β-globin subunit is responsible for sickle cell anemia. This occurs in the second position of the codon for position 6 of the β chain. In the mRNA the codon, GAG, which specifies glutamate in normal β chains, is converted to GUG, which specifies valine. The resultant hemoglobin, called **hemoglobin S** (HbS), has altered surface charge properties (because the negative charge of glutamate has been replaced by valine's nonpolar group), which is responsible for clinical symptoms. This mutation occurs mainly in peoples of equatorial African descent and is the classic example of a mutation that confers an adaptive advantage as well as a **genetically inheritable disease.** Individuals heterozygous for HbS are resistant to infection by the parasites that cause **malaria** but do not acquire the symptoms of sickle cell disease exhibited by individuals homozygous for HbS. The life cycle of the malaria-causing parasites includes an obligatory stage that occurs inside erythrocytes and they do not survive in erythrocytes containing HbS. Carriers of the mutation can be detected by restriction enzyme digestion of a sample of the potential carrier's DNA followed by **Southern hybridization** technique with the β-globin cDNA as described in Clin. Corr. 19.4.

Thalassemias Are Caused by Mutations in Genes for the α or β Subunits of Globin

Thalassemias are a family of related genetic diseases that occur in people who frequently originate from the Mediterranean areas and Asia. If there is a reduced synthesis or a total lack of synthesis of α-globin mRNA, the disease is classified as **α-thalassemia;** if the β-globin mRNA level is affected, it is called **β-thalassemia.** Thalassemias can be due to the deletion of one or more globin-like genes in either of the globin gene clusters or be caused by a defect in the transcription or processing of a globin gene's mRNA.

Since each chromosome 16 contains two adjacent α-globin genes, a normal diploid individual has four copies of this gene. α-Thalassemic patients may be missing one to four α-globin genes. The condition in which one α-globin gene is missing is referred to as **α-thal 1;** when two α-globin genes are gone, the condition is **α-thal 2.** In both cases the individuals can experience mild to moderate anemia but may have no additional symptoms. When three α-globin genes are missing, many more β-globin molecules are synthesized than α-globin molecules, resulting in the formation of a globin tetramer of four β-globins, which causes **HbH disease** and accompanying anemia. When all four α-globin genes are absent, the disease **hydrops fetalis** occurs, which is fatal at or before birth. Some chromosomal deletions that have been mapped in the α-globin gene cluster are shown in Figure 19.25.

β-Thalassemias also exhibit different degrees of severity and can be caused by a variety of defects or deletions. In one case the β-globin gene is present but has undergone a mutation in the codon 17, which generates a termination codon. In another case the β-globin gene is transcribed in the nucleus but no β-globin mRNA occurs in the cytoplasm. Thus a defect has occurred in the processing and/or transport of the primary transcript of the gene.

Other β-thalassemias are caused by deletions within the β-globin gene cluster on chromosome 11 (Figure 19.25). In some cases these deletions remove the DNA between two adjacent genes, resulting in a new fusion gene. For example, in the normal person the linked -globin and β-globin genes differ in only about 7% of their positions. In **Hb Lepore** a deletion has placed the front portion of the -globin gene in register with the back portion of the β-globin gene. From this fusion gene a new β-like globin is produced in which the N-terminal sequence of -globin is joined to the C-terminal sequence of β-globin. Several variants of Hb Lepore are known, and in each case the globin

CLINICAL CORRELATION 19.5

Prenatal Diagnosis of Thalassemia

If a fetus is suspected of being thalassemic because of its genetic background, recombinant DNA techniques can be used to determine if one or more globin genes are missing from its genome. Fetal DNA can easily be obtained (in relatively small quantities) from amniotic fluid cells aspirated early during the second trimester of pregnancy. Regions of interest are amplified from the fetal DNA by polymerase chain reactions and digested with restriction enzymes that divide the globin genes among restriction fragments of several hundred to 2000 base pairs. These fragments are separated by electrophoresis through an agarose gel and hybridized with radioactive cDNA for α- and/or β-globin using the Southern blot technique (see p. 774). If one or more globin genes are missing, the corresponding restriction fragment will not be detected or its hybridization to the radioactive cDNA probe will be reduced (in the case when only one of two diploid genes is absent).

Benz, E. J. The hemoglobinopathies. In: W. N. Kelly (Ed.), *Textbook of Internal Medicine.* Philadelphia: Lippincott, 1989, pp. 1423–1432.

product is a composite of the and β sequence, but the actual fusion junction is different.

Another fusion β-like globin is **Hb Kenya.** This deletion results in a gene product that contains the N-terminal sequence of the γ-globin gene and the C-terminal sequence of the β-globin gene. Still another series of deletions has been found in which both the - and β-globin genes are removed, causing HPFH **(hereditary persistence of fetal hemoglobin).** Frequently, there are no clinical symptoms of this condition because fetal hemoglobin ($_2$ $_2$) continues to be synthesized after the time at which γ-globin gene expression is normally turned off (see Clin. Corr. 19.5.)

19.11—
Genes for Human Growth Hormone-Like Proteins

Human growth hormone (hGH, also called somatotropin) is a polypeptide of 191 amino acids. A larger precursor is synthesized in the somatotrophs of the anterior pituitary, and the mature form is secreted into the circulatory system. Growth hormone induces liver (and perhaps other) cells to produce other hormones called **somatomedins,** which are insulin-like growth factors that stimulate proliferation of mesodermal tissues such as bone, cartilage, and muscle. Infants with a deficiency in growth hormone become dwarfs, whereas those who produce too much become giants.

A closely related protein of 191 amino acids, having 85% homology with growth hormone, is human **chorionic somatomammotropin** (hCS, also called placental lactogen) synthesized in the placenta. The complete role of this hormone in normal fetal–maternal physiology is still unclear, but it participates in placental growth and contributes to mammary gland preparation for lactation during pregnancy.

The hormones hGH and hCS are examples of two very similar proteins that serve different biological functions and are synthesized in different tissues. It is to be expected that their genes also are closely related but expressed in a tissue-specific fashion. The genes for hGH and hCS are very similar and occur in the same region of chromosome 17 (Figure 19.26).

Five related genes comprise the human growth hormone gene family. They occur over a distance of about 55 kb and share a common structure of five exons and four introns, with the exon–intron boundaries always in the same locations. Alu repetitive sequences occur between some of the genes, as in the globin gene clusters. The order of the genes is 5 *hGH-N ψhCS-L, hCS-A, hGH-V, hCS-B 3′. The first gene in this cluster, hGH-N,* is expressed in the anterior

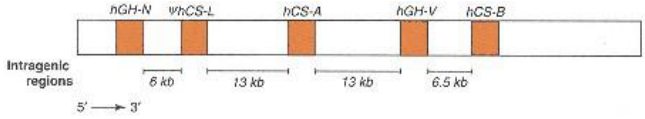


Figure 19.26
Organization of human growth hormone (*hGH*) gene family.
The five structural genes of this family occur as a linear array over about 55 kb on the long arm of chromosome 17. Two genes (*hGH-N* and *hGH-V*) code for growth hormone, two genes code for the closely related human chorionic somatomammotropin (*hCS-A* and *hCS-B*), and one gene appears to be a pseudogene (hCS-L). Only *hGH-N* is expressed in the pituitary; other genes are expressed in the placenta. The order of the genes in the array (red boxes) is 5 *hGH-N*, *hCS-L*, *hCS-A*, *hGH-V*, *hCS-B* 3 , and all are transcribed in the same direction. Each gene has the same basic structure of five exons and four introns (not shown).
Redrawn from Chen, E. Y., Liao, Y. C., Smith, D. H., Barrera-Saldana, H. A. et al. *Genomics* 2:479, 1989.

pituitary, whereas the others are expressed in the placenta. The *hGH* and *hCS* genes have different sequences about 100 bp beyond their polyadenylation sites. The *hGH-N* gene codes for normal growth hormone of 22 kDa. Alternative splicing of intron 3 of this gene occurs in about 10% of the primary transcripts, giving rise to a 20-kDa version of growth hormone whose significance is not known. (See Chapter 16.) The *hGH-V* gene codes for a variant growth hormone that can be expressed in transgenic animals (see Section 19.14), but whose function in the placenta is unknown. The *hCS-A* and *hCS-B* genes specify the same mature hormone but are expressed at different levels in the placenta. The hCS-L pseudogene has a single base substitution at an exon-intron splice site that appears to prevent normal maturation of its primary transcript into mRNA.

Expression of *hGH* and *hCS* genes is under the regulation of other hormones. **Thyroxine** and **cortisol** stimulate increased transcription of these genes. In cultured rat pituitary tumor cells these hormones act in a synergistic fashion to induce growth hormone mRNA synthesis. Pituitary cells that have only about two molecules of growth hormone mRNA per cell can be stimulated to a level of 1000 growth hormone mRNA molecules per cell, a 500-fold increase comparable in magnitude to the induction of many bacterial operons.

Only some of the details by which thyroxin and cortisol stimulate this increased transcription are known. Their regulatory effect at the molecular level is clearly more complicated than is the control of bacterial operon transcription. Two promoter sites lie just upstream of *hGH-N* and a specific transcription factor, GHF-1 (also called Pit-1), contributes to this gene's pituitary-specific expression. GHF-1 belongs to a family of **homeodomain transcription factors** found in organisms as diverse as yeast and fruitflies. The regulatory hormones are transported into the nucleus and in association either with their receptors or with a binding protein, such as GHF-1, affect transcription initiation at *hGH-N*. Alternatively, these other hormones may interact with additional factors in the cell that in turn regulate the level of transcription. The DNA regulatory site influenced by glucocorticoid hormones is known to be upstream of the site at which transcription of *hGS-N* begins. An example of the many transcription initiation protein factors that can interact with the DNA in the vicinity of eukaryotic genes is shown in Figure 16.18.

Deletions can occur within the growth hormone gene family. Deletions of *hGH-N* in both copies of chromosome 17 have been detected in some cases of severe growth hormone deficiency. These individuals are very short and do not have detectable serum growth hormone. Some such children initially respond very well to treatment with recombinant human growth hormone synthesized in the bacterium *E. coli* (see p. 834 and Figure 19.29) but they often develop antibodies against the growth hormone. Deletions also have been detected in which *hCS-A, hGH-V,* and *hCS-B* are lost from both chromosome 17 copies. Despite the fact that maternal sera of these individuals lack these hormones, fetal development usually proceeds normally, suggesting they either are unnecessary or can be compensated for by other hormones or factors.

19.12— Mitochondrial Genes

About 0.3% of the DNA of human cells occurs in the mitochondria. Human mitochondrial DNA (mtDNA) is a double-stranded circular molecule of 16,569 bp whose sequence has been completely determined. As many as 100 molecules of mtDNA can occur in a metabolically active cell. Each mtDNA codes for 2 rRNAs, 22 tRNAs, and 13 proteins, most of which are subunits of multi-subunit complexes in the mitochondrial inner membrane that catalyze oxidative phosphorylation (Figure 19.27). For example, Complex I (NADH dehydrogenase), the first of three proton-pumping complexes involved in oxidative phosphorylation, is comprised of 26 proteins. Seven of these proteins are encoded by the

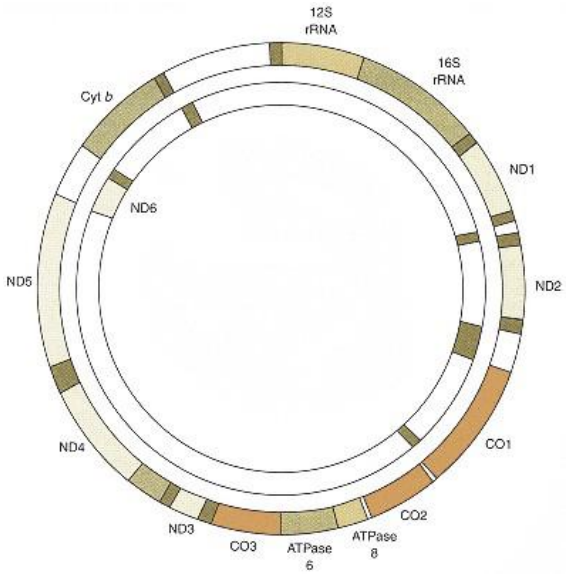


Figure 19.27
Human mitochondrial DNA.
The 16,569-bp human mtDNA molecule codes for two ribosomal RNAs (12 S and 16 S rRNA), some of the subunits for NADH dehydrogenase (ND), cytochrome oxidase (CO), ATP synthase (ATPase), and cytochrome b (cyt *b*), and 22 tRNAs (dark gray regions). Most genes occur on the outer DNA strand but genes for ND6 and a few tRNAs are on the inner strand.

CLINICAL CORRELATION 19.6

Leber's Hereditary Optic Neuropathy (LHON)

Leber's hereditary optic neuropathy, first described in 1871, is a maternally inherited genetic disease that usually strikes young adults and results in complete or partial blindness from optic nerve degeneration. Other neurological disorders such as cardiac dysrhythmia can also be associated with the disease. The cause of this defect in many patients has been traced to a single base pair mutation in the mitochondrial DNA that changes an arginine to a histidine at amino acid 340 in NADH dehydrogenase subunit 4 of Complex I in the inner mitochondrial membrane. Although it is not clear why this mutation leads to blindness, the eyes require a high level of mitochondrial activity and perhaps become sensitive over time to a small decrease in ATP synthesis by oxidative phosphorylation.

Singh, G., Lott, M. T., and Wallace, D. C. A mitochondrial DNA mutation as a cause of Leber's Hereditary Optic Neuropathy. *N. Eng. J. Med.* 320:1300, 1989.

mtDNA. Mitochondrial DNA also contains genes for three cytochrome oxidase subunits, two ATP synthase subunits, and cytochrome *b*. In contrast to the nucleus, where much of the chromosomal DNA seems to have no genetic function, virtually every base pair in mtDNA is essential. Regions between the protein-coding genes usually encode tRNAs and sometimes the last nucleotide of one gene will be the first nucleotide of the adjacent gene. Polyadenylation at the 3 ends of some of the mitochondrial mRNAs adds the last two A residues of the termination codon, UAA, to create the end of the reading frame.

Even more remarkable, the genetic code of mammalian mtDNA is not identical to the genetic code of nuclear or prokaryotic DNA. UGA codes for tryptophan instead of for termination, AUA codes for methionine rather than isoleucine, and AGA and AGG serve as stop codons instead of specifying arginine. It is not clear why mitochondria have their own altered genetic system. Perhaps mtDNA is an evolutionary vestige of an early symbiotic relationship between a bacterium and the progenitor of eukaryotic cells. What is clear is that cells makes a large investment to express the 13 mitochondria-encoded proteins. To produce those proteins a large group of nucleus-encoded ribosomal proteins and associated translation factors must be imported into the mitochondrion and assembled, as well as all of the enzymes and binding proteins required for mtDNA replication and transcription. More than 100 nucleus-encoded proteins are probably necessary to maintain the mtDNA and express its gene products.

Since mitochondria are in the cytoplasm, mtDNA molecules are maternally inherited. mtDNA sequences can be used as markers for maternal lineages. In addition, mutations in mtDNA can lead to genetic diseases that are inherited only from the mother. For example, a single base pair change in mtDNA has been found to be responsible for **Leber's hereditary optic neuropathy** (see Clin. Corr. 19.6). Similar mtDNA mutations may be the cause of two other maternally inherited genetic diseases, *myoclonic epilepsy* and **infantile bilateral striatal necrosis.**

19.13— Bacterial Expression of Foreign Genes

Recombinant DNA techniques are now frequently used to construct bacteria that are "factories" for making large quantities of specific human proteins useful in the diagnosis or treatment of disease. The two examples to be illustrated here are the construction of bacteria that synthesize human insulin and human growth hormone.

Many factors must be considered in designing recombinant plasmids that contain a eukaryotic gene to be expressed in bacteria. First, the cloned eukaryotic gene cannot have any introns since the bacteria do not have the RNA-splicing enzymes that correctly remove introns from the initial transcript. Thus the actual eukaryotic chromosomal gene is usually not used for these experiments; instead, the cDNA or a synthetic equivalent of the coding sequence, or a combination of both, is placed in the bacterial plasmid.

Another consideration is that different nucleotide sequences comprise the binding sites for RNA polymerase and ribosomes in bacteria and eukaryotes. Therefore, to achieve expression of the desired protein, it is necessary to insert the eukaryotic coding sequence directly behind a set of bacterial regulatory elements. This has the advantage that the foreign gene is now under the regulation of the bacterial control elements, but its disadvantage is that considerable recombinant DNA manipulation is required to make the appropriate plasmid. Still other factors to be considered are that the foreign gene product must not be degraded by bacterial proteases or require modification before it is active (e.g., specific glycosylation events that the bacteria cannot perform) and must not be toxic to the bacteria. Even when the bacteria do synthesize the desired product, it must be isolated from the 1000 or more endogenous bacterial proteins.

Recombinant Bacteria Can Synthesize Human Insulin

Insulin is produced by the β cells of the pancreatic islets of Langerhans. It is initially synthesized as **preproinsulin,** a precursor polypeptide that possesses an N-terminal signal peptide and an internal C peptide of 33 amino acids that are removed during the subsequent maturation and secretion of insulin (see p. 40). The **A peptide** (21 amino acids) and **B peptide** (30 amino acids) of mature insulin are both derived from this initial precursor and are held together by two disulfide bridges. Bacteria do not have the processing enzymes that convert the precursor form to mature insulin. Therefore the initial strategy for bacterial synthesis of human insulin involved the production of the A and B chains by separate bacteria followed by purification of the individual chains and subsequent formation of the proper disulfide linkages.

The first step was to use synthetic organic chemistry methods to prepare a series of single-stranded oligonucleotides (11–18 nucleotides) that were both complementary and overlapping with each other. When these oligonucleotides were mixed together in the presence of **DNA ligase** under proper conditions, they formed a double-stranded fragment of DNA with termini equivalent to those formed by specific restriction enzymes (Figure 19.28). The sequences of the oligonucleotides were carefully chosen so that one of the two strands contained a methionine codon followed by the coding sequence of the A chain of insulin and a termination codon. A second set of overlapping complementary oligonucleotides were prepared and ligated together to form another double-stranded DNA fragment that contained a methionine codon followed by 30 codons specifying the B chain of insulin and a termination codon.

These two double-stranded fragments were then individually cloned at a restriction site in the β-galactosidase gene of the lactose operon in a plasmid. These two recombinant plasmids were introduced into bacteria. The bacteria could now produce a fusion protein of β-galactosidase and the A chain (or B chain) whose expression was under control of the lactose operon. In the absence

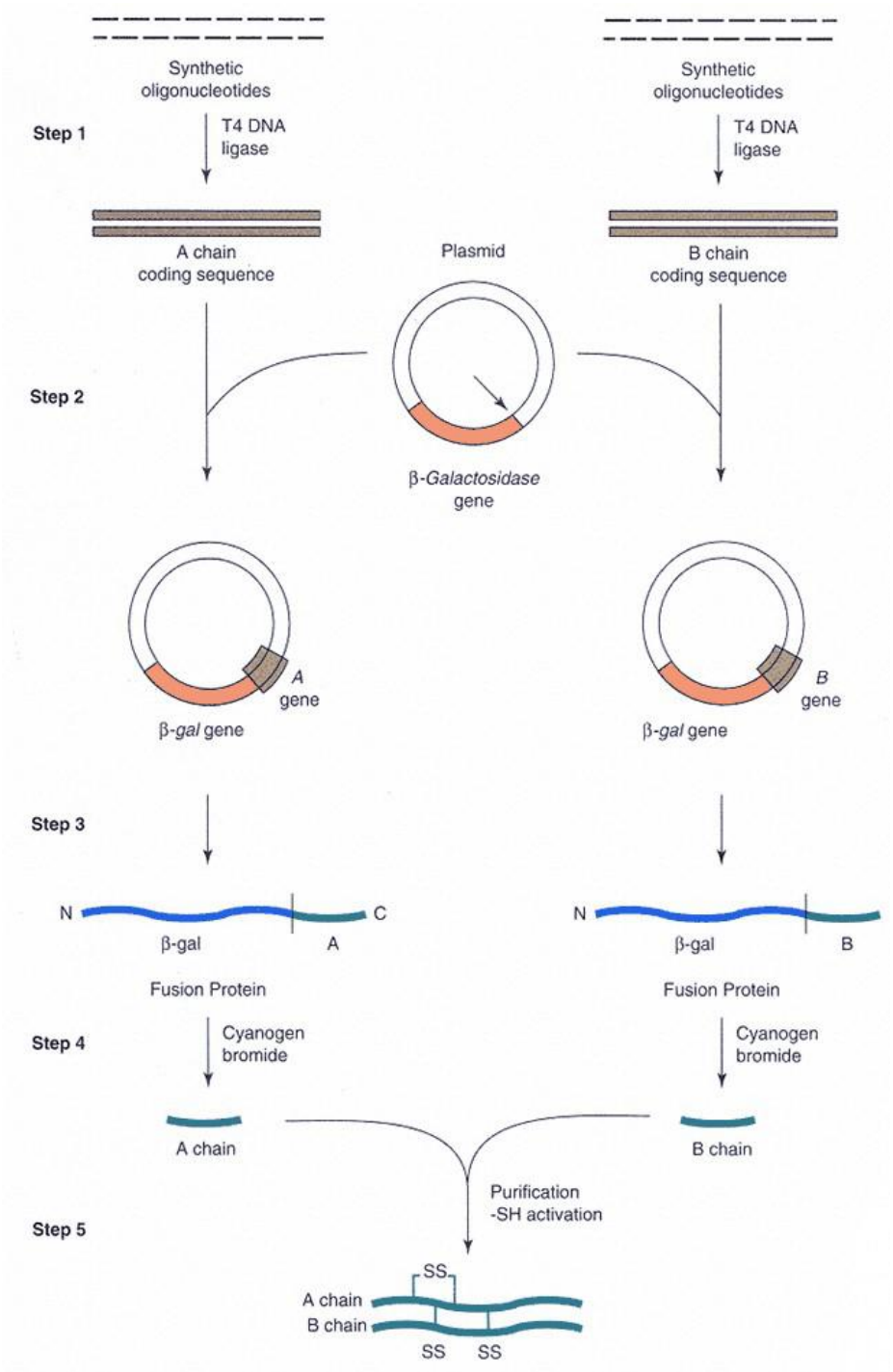


Figure 19.28
Bacterial expression of the A and B chains of human insulin.
Step 1: A series of overlapping, complementary oligonucleotides (11 for the A chain and 18 for the B chain) were synthesized and ligated together. One strand of the resulting small DNA fragments contained a methionine codon followed by coding sequence for A chain and B chain, respectively.
Step 2: The small DNA fragments were ligated into a restriction site near the end of the *β*-galactosidase gene of the lactose operon in a plasmid.
Step 3: Recombinant plasmids were introduced into E. coli and the *β*-galactosidase gene was induced with IPTG, an inducer of the lactose operon. A fusion protein was produced that contained most of the *β*-galactosidase sequence at the N terminus and the A chain (or B chain) at the C terminus.
Step 4: Bacterial cell lysates containing the fusion protein were treated with cyanogen bromide, which cleaves peptide bonds following methionine residues. Step 5: A and B chains were purified away from all other cyanogen bromide peptides using biochemical and immunological separation techniques. The –SH groups on the cysteines were activated and reacted to form intra- and interchain disulfide bridges found in mature human insulin.
Redrawn from Crea, R., Krazewski, A., Hirose, T., and Itakura, K. *Proc. Natl. Acad. Sci. USA* 75:5765, 1980.

of lactose in the bacterial medium, the lactose operon is repressed and only very small amounts of the fusion protein are synthesized. Using induction with IPTG and some additional genetic tricks, the bacteria can be forced to synthesize as much as 20% of their protein as the fusion protein. The A peptide (or B peptide) can be released from this fusion protein by treatment with cyanogen bromide, which cleaves on the carboxyl side of methionine residues. Since neither the A nor B peptide contains a methionine, they will be liberated intact and can subsequently be purified to homogeneity. The final steps involve chemically activating the free –SH groups on the cysteines and mixing the activated A and B chains together in a way that the proper disulfide linkages form to generate molecules of mature human insulin.

Recombinant Bacteria Can Synthesize Human Growth Hormone

The strategy for generating a recombinant DNA plasmid from which bacteria can synthesize human growth hormone is somewhat different than for insulin synthesis. First, human growth hormone is 191 amino acids long so the synthetic construction of the corresponding DNA coding sequence is more difficult (although certainly not impossible) than in the insulin case. On the other hand, growth hormone is a single polypeptide so it is not necessary to deal with the production of two chains and their subsequent dimerization to form a protein with biological activity. Because of these considerations, the coding sequence was initially cloned into a bacterial expression plasmid using part of a cloned growth hormone cDNA and several synthetic oligonucleotides (Figure 19.29). The overlapping oligonucleotides were prepared so that, when ligated together, they would form a small double-stranded DNA containing the codons for the first 24 amino acids of mature human growth hormone. One end of this DNA

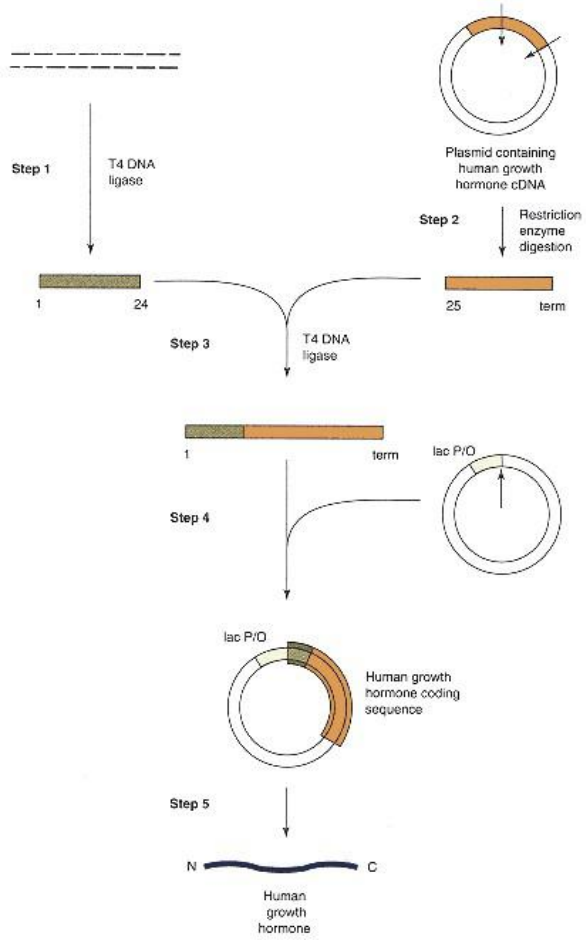


Figure 19.29
Expression of human growth hormone in *E. coli*.
Step 1: Several overlapping, complementary, oligonucleotides were synthesized and ligated together. One strand of the resulting small DNA fragment contains the coding sequence for the first 24 amino acids of mature human growth hormone (after removal of the N-terminal signal peptide).
Step 2: A recombinant plasmid with a full length human growth hormone (hGH) cDNA, which is not expressed, is cleaved with restriction enzymes that release a fragment containing the complete growth hormone coding sequence after codon 24.
Step 3: The synthetic fragment and the partial cDNA-containing fragment are ligated together to yield a new fragment containing the complete coding sequence of mature hGH.
Step 4: The new fragment is ligated into a restriction site just downstream from the lactose promoter–operator region cloned in a plasmid.
Step 5: The resulting recombinant DNA plasmid is introduced into bacteria in which synthesis of hGH can be induced with IPTG, an inducer of the lactose operon.

fragment was designed so that the fragment could be ligated in front of a restriction fragment of growth hormone cDNA that provided the rest of the coding sequence, including the termination codon. The other end of the synthetic fragment was chosen so that the composite coding sequence could easily be inserted into a site immediately downstream of the promoter–operator–ribosome binding site of the lactose operon cloned in a plasmid. After the introduction into bacteria, the bacteria were induced with IPTG to transcribe this foreign coding region and the greatly overproduced human growth hormone subsequently was purified away from the bacterial proteins.

19.14— Introduction of Rat Growth Hormone Gene into Mice

The previous section described the use of bacteria to produce large quantities of human proteins for treatment of disease. It is possible to microinject molecules of purified RNA or DNA directly into eukaryotic cells. This provides a very powerful approach for identifying conditions under which specific genes are expressed in eukaryotic cells. One of the most dramatic illustrations of this approach was the microinjection of a chromosomal DNA fragment containing the structural gene for rat growth hormone into the **pronuclei** of fertilized mouse eggs. The eggs were then reimplanted into the reproductive tracts of foster mouse mothers. Some of the mice that developed from this procedure were **transgenic;** one or more copies of the microinjected growth hormone gene integrated into a host mouse chromosome at an early stage of embryo development. These foreign genes were transmitted through the germline and became a permanent feature in the host chromosomes of the progeny (Figure 19.30).

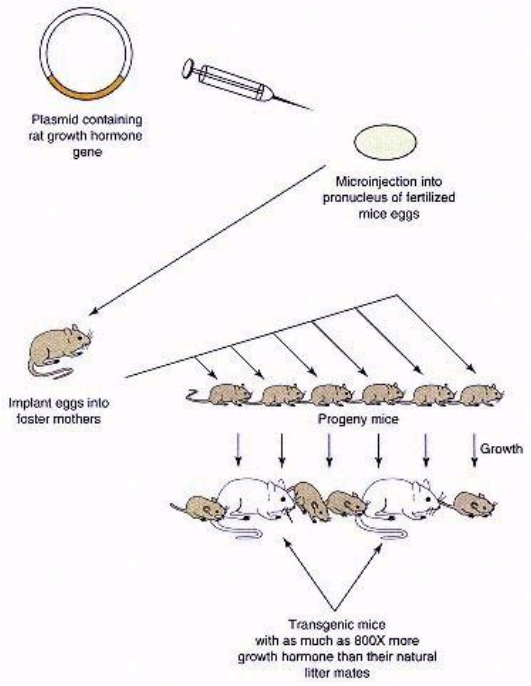


Figure 19.30
Schematic illustration of the introduction of rat growth hormone gene into mice.
Copies of a recombinant plasmid DNA containing rat growth hormone gene were microinjected into fertilized mouse eggs that were reimplanted into foster mothers. Some of the resulting progeny contained the foreign gene integrated into their own genome and greatly overexpressed growth hormone, growing much larger than their normal-sized littermates. Redrawn from Palmiter, R. D., Brinster, R. L., Hammer, R. E. et al. *Nature* 300:611, 1982.

Analysis of these transgenic mice revealed that in some cases several tandem copies of the rat growth hormone gene had integrated into a mouse chromosome; in other cases only one gene copy was present. In all cases at least some transcription occurred from the integrated gene(s), and in a few cases a dramatic overproduction of rat growth hormone resulted. In these latter cases, as much as 800 times more growth hormone was present in the transgenic mice than in normal mice, resulting in animals more than three times the size and weight of their unaffected littermates.

These results present many potential experimental possibilities for the future and raise a number of issues. One implied possibility is the use of similar growth hormone gene insertions to stimulate rapid growth of commercially valuable animals. This could result in a shorter production time and increased efficiency of food utilization. Another long-term possibility is the use of this approach to correct certain human genetic diseases or mimic the diseases in experimental animals so that they can be studied more carefully. One obvious human disease that is a candidate for this **"gene therapy"** approach is thalassemia. For example, an individual with two to three missing α-globin genes might benefit tremendously from receiving bone marrow transplants of his/her own cells that have been established in culture and microinjected with additional copies of the normal α-globin gene. This approach to gene therapy is being investigated. Insertion of normal genes into human somatic cells of a defective tissue or organ does not result in transmission of these genes to the progeny. This lessens the ethical considerations for experiments that do not alter germline characteristics.

Bibliography

Prokaryotic Gene Expression

Cohen, S. N., and Shapiro, J. A. Transposable genetic elements. *Sci. Am.* 242:40, 1980.

Miller, J. H. The *lac* gene: its role in *lac* operon control and its use as a genetic system. In: J. H. Miller and W. S. Resnikoff (Eds.), *The Operon.* Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press, 1978, p. 31.

Platt, T. Regulation of gene expression in the tryptophan operon of *Escherichia coli.* In: J. H. Miller and W. S. Resnikoff (Eds.), *The Operon.* Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press, 1978, p. 263.

Simon, M., Zieg, J., Silverman, M., Mandel, G., and Doolittle, R. Phase variation: evolution of a controlling element. *Science* 209:1370, 1980.

Eukaryotic Gene Expression

Ahn, A. H., and Kunkel, L. M. The structural and functional diversity of dystrophin. *Nature Genetics* 3:283, 1993.

Chen, E. Y., Liao, Y.-C., Smith, D. H., Barrera-Saldaña, H. A., Gelinas, R. E., and Seeburg, P. H. The human growth hormone locus: nucleotide sequence, biology and evolution. *Genomics* 4:479, 1989.

Enver, T., and Greaves, D. R. Globin gene switching: a paradigm or what? *Curr. Opin. Biotech.* 2:787, 1991.

Johns, D. R. Mitochondrial DNA and disease. *N. Engl. J. Med.* 333:638, 1995.

Karlsson, S., and Nienhuis, A. W. Developmental regulation of human globin genes. *Annu. Rev. Biochem.* 54:1071, 1985.

Maniatis, T., Fritsch, E. F., Laurer, J., and Lawn, R. M. The molecular genetics of human hemoglobins. *Annu. Rev. Genet.* 14:145, 1980.

Mitchell, P. J., and Tjian, R. Transcription regulation in mammalian cells by sequence-specific DNA binding proteins. *Science* 245:371, 1989.

Palmiter, R. D., Brinster, R. L., Hammer, R. E., et al. Dramatic growth of mice that develop from eggs microinjected with metallothionein-growth hormone fusion genes. *Nature* 300:611, 1982.

Singh, G., Lott, M. T., and Wallace, D. C. A mitochondrial DNA mutation as a cause of Leber's hereditary optic neuropathy. *N. Engl. J. Med.* 320:1300, 1989.

Struhl, K. Chromatin structure and RNA polymerase II connection: implications for transcription. *Cell* 84:179, 1996.

The Huntington's Disease Collaborative Research Group. A novel gene containing a trinucleotide repeat that is expanded and unstable on Huntington's disease chromosomes. *Cell* 72:971, 1993.

Questions

J. Baggott and C. N. Angstadt

1. Full expression of the *lac* operon requires:

   A. lactose and cAMP.

   B. allolactose and cAMP.

   C. lactose alone.

   D. allolactose alone.

   E. absence or inactivation of the *lac* corepressor.

2. In an operon:

   A. each gene of the operon is regulated independently to achieve levels of expression required by the cell.

   B. control may be exerted via induction or via repression.

   C. operator and promoter may be trans to the genes they regulate.

D. the structural genes are either not expressed at all or are fully expressed.

E. control of gene expression consists exclusively of induction and repression.

3. The *E. coli lacZYA* region will be upregulated if:

A. there is a defect in binding of the inducer to the product of the *lacI* gene.

B. glucose and lactose are both present in the growth medium, but there is a defect in the cell's ability to bind the CAP protein.

C. glucose and lactose are both readily available in the growth medium.

D. the operator has mutated so it can no longer bind repressor.

E. the *lac* corepressor is not present.

4. All of the following describe an operon EXCEPT:

A. control mechanism for eukaryotic genes.

B. includes structural genes.

C. expected to code for polycistronic mRNA.

D. contains control sequences such as an operator.

E. can have multiple promoters.

Refer to the following for Questions 5–9:

A. repression

B. corepression

C. attenuation

D. stringent response

E. RNA splicing

5. Associated with guanosine tetraphosphate and guanosine pentaphosphate.

6. Not found in prokaryotes.

7. Involves rho-independent transcription termination.

8. Involves a leader peptide containing several occurrences of the same amino acid.

9. The only regulatory mechanism for the *his* operon.

10. Ribosomal operons:

A. all contain genes for proteins of just one ribosomal subunit.

B. all contain genes for proteins of both ribosomal subunits.

C. all contain genes for only ribosomal proteins.

D. can have their expression regulated at the level of translation.

E. are widely separated in the *E. coli* chromosome.

11. All of the following phrases describe transposons EXCEPT:

A. a means for the permanent incorporation of antibiotic resistance into the bacterial chromosome.

B. contain short inverted terminal repeat sequences.

C. code for an enzyme that synthesizes guanosine tetraphosphate and guanosine pentaphosphate, which inhibit further transposition.

D. include at least one gene that codes for a transposase.

E. contain varying numbers of genes, from two to several.

12. Introns:

A. are of approximately uniform size.

B. are skipped over during translation.

C. are found in all eukaryotic genes.

D. function to separate functional domains of proteins.

E. are smaller and shorter in unicellular lower eukaryotes than in higher, more complicated eukaryotes.

13. Repetitive DNA:

A. is common in bacterial and mammalian systems.

B. is uniformly distributed throughout the genome.

C. includes DNA that codes for rRNA.

D. consists mostly of DNA that codes for enzymes catalyzing major metabolic processes.

E. is resistant to the action of restriction endonucleases.

14. The β-gene cluster contains:

A. one haploid gene.

B. one haploid β gene.

C. one haploid γ gene.

D. two haploid genes.

E. two haploid genes.

15. The number of α genes in the haploid α-gene cluster is

A. one.

B. two.

C. three.

D. four.

E. five.

16. In designing a recombinant DNA for the purpose of synthesizing an active eukaryotic polypeptide in bacteria all of the following should be true EXCEPT:

A. the eukaryotic gene may contain its usual complement of introns.

B. the foreign polypeptide should be resistant to degradation by bacterial proteases.

C. glycosylation of the polypeptide should be unnecessary.

D. the foreign polypeptide should be nontoxic to the bacteria.

E. bacterial controlling elements are necessary.

Answers

1. B A: The true inducer is allolactose, not lactose (p. 806). C: Lactose is converted in the cell to allolactose. D: In addition, cAMP must bind to the CAP protein, and the cAMP-CAP complex serves as a positive control of transcription (p. 807). E: The *lac* operon does not involve corepression.

2. B Induction and repression are among the mechanisms used to control operons. A: In an operon the structural genes are under coordinate control. C: The operator and promoter are elements of the same strand of DNA as the operon they control; they are not diffusible. D: Typically, regulation of operators is somewhat leaky; some gene product is produced even in the repressed state. E: Another mechanism for regulation of an operon is attenuation (p. 810).

3. D If the operator is unable to bind repressor, the rate of transcription is greater than the basal level (p. 804). A: The product of the *lac*I gene is the repressor protein. When this protein binds an inducer, it changes its conformation, no longer binds to the operator site of DNA, and transcription occurs at an increased rate. Failure to bind an inducer prevents this sequence. B and C: In the presence of glucose catabolite repression occurs. Glucose lowers the intracellular level of cAMP. The catabolite activator protein (CAP) then cannot complex with cAMP, so there is no CAP–cAMP complex to activate transcription. The same would occur if the cell had lost its capacity to synthesize cAMP (p. 808). E: The *lac* operon does not involve corepression.

4. A Operons are prokaryotic mechanisms. B–D: An operon is the complete regulatory unit of a set of clustered genes, including the structural genes (which are transcribed together to form a polycistronic mRNA), regulatory genes, and control elements, such as the operator (p. 801). E: An operon may have more than one promoter, as does the tryptophan operon of *E. coli* (p. 810).

5. D The exact functions of these species are not yet known, but their production is very rapid after the onset of amino acid starvation (p. 815).

6. E Splicing is a eukaryotic phenomenon (p. 820).

7. C The hairpin loop that forms between regions 3 and 4 (Figure 19.11) is followed by an oligo-U region (Figure 19.10). This constellation compromises the signal for rho-independent termination of transcription. (See pp. 811–812.)

8. C Synthesis of the leader peptide depends strongly on availability of this amino acid, since it must be incorporated several times. When it is insufficiently available, the ribosome stalls, in region 1 (Figure 19.11), allowing the 2–3 hairpin to form. This in turn prevents formation of the 3–4 hairpin, which would signal termination of transcription.

9. C In this operon the stalled ribosome acts rather like a positive regulator protein, that is, the cAMP–CAP complex (p. 813).

10. D Excess ribosomal protein binds to its own mRNA, preventing initiation of further synthesis of that protein (p. 814). A, B, C, and E: The genes for one half of the ribosomal proteins are in two major clusters. There is no pattern to the distribution of genes for the proteins of the two ribosomal subunits, and they are intermixed with genes for other proteins involved in protein synthesis.

11. C These guanosine phosphates are synthesized by the product of the *relA* gene; they inhibit initiation of transcription of the rRNA and tRNA genes, shutting off protein synthesis in general. This is the stringent response (p. 815).

12. E A: Introns are of various sizes. B: They are excised during splicing, not skipped over during translation. C: Although they are common, some genes do not have them, for example, the histone and interferon genes (p. 821). D: Sometimes they occur between functional domains of proteins, but not always.

13. C This makes sense, since many copies of these structural elements are needed (p. 823). A and B: Highly repetitive and moderately repetitive DNA are found only in eukaryotes. Highly repetitive sequences tend to be clustered, as are some moderately repetitive sequences (p. 822). D: Most repetitive DNA does not code for a stable gene product (p. 822). E: The Alu family of moderately repetitive DNA is named for the restriction endonuclease that cleaves them (p. 823).

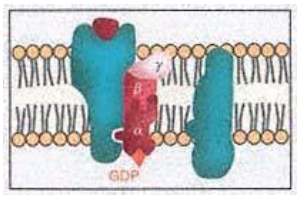
14. B This means that there are only two β genes per diploid cell. As a result, in β-thalassemia, one defective β-globin gene gives rise to a minor form of the disease, while two defective genes cause the major form. (See p. 828.)

15. B As a result, α-thalassemia is more complicated than β-thalassemia because there are four α-globin genes per diploid cell, and anywhere between zero and four of them can be defective. (See p. 828.)

16. A A and C: The bacterial system has no mechanism for posttranscriptional modification of mRNA or for posttranslational (or cotranslational) modification of protein. E: Bacterial systems need bacterial promoters, and so on (p. 832).

Chapter 20—
Biochemistry of Hormones I: Polypeptide Hormones

Gerald Litwack and Thomas J. Schmidt



20.1— Overview

Cells are regulated by many hormones, growth factors, neurotransmitters, and certain toxins through interactions of these diverse ligands with their **cognate receptors** located at the cell surface. This collection of receptors is the major mechanism through which **peptide hormones** and **amino acid-derived hormones** exert their effects at the cellular level. Another important mechanism involves permeation of the cell membrane by **steroid hormones** that subsequently interact with their intracellular cognate receptors (Chapter 21). These two sites, the plasma membrane and the intracellular milieu, represent the principal locations of the initial interaction between ligands and cellular receptors and are diagrammed in Figure 20.1. Polypeptide hormones and several amino acid-derived hormones bind to cognate receptors in the plasma membrane. One exception is thyroid hormone, which binds to a receptor that resides in the nucleus much like certain steroid hormone receptors.

The **hormonal cascade system** is applicable to many, but not all, hormones. It begins with signals in the central nervous system (CNS), followed by hormone secretion by the hypothalamus, pituitary, and end target organ. In this chapter major polypeptide hormones are summarized and the synthesis of specific hormones is described. Synthesis of the amino acid-derived hormones, epinephrine and triiodo-L-thyronine, is also outlined. Examples of hormone inactivation and degradation are presented. The remainder of this chapter focuses on receptors, **signal transduction**, and **second messenger pathways.** Receptor internalization is described and examples of cyclic hormonal cascade systems are introduced. Finally, a discussion of oncogenes and receptor function is presented.

In terms of receptor mechanisms, aspects of hormone–receptor interactions are presented with a brief mathematical analysis. Signal transduction is considered, especially in reference to GTP-binding proteins. Second messenger systems discussed include cAMP and the protein kinase A pathway, inositol triphosphate–diacylglycerol and the Ca^{2+}–protein kinase C pathway, and cGMP and

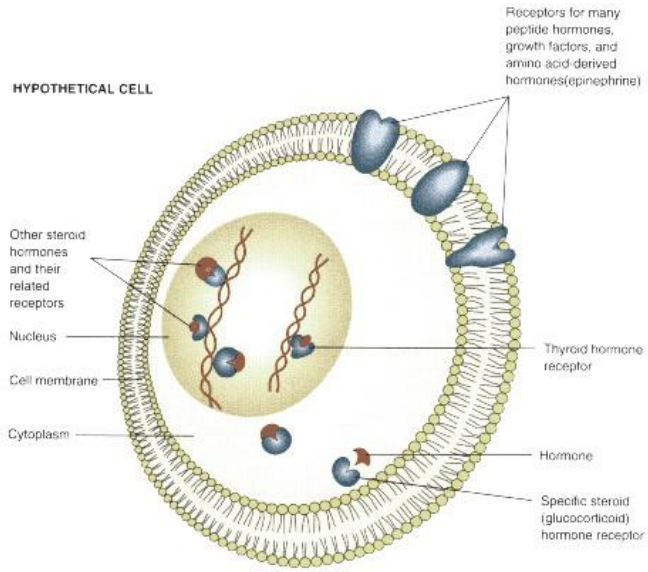


Figure 20.1
Diagram showing the different locations of classes of receptors expressed by a target cell.

the protein kinase G pathway. These pathways are discussed in the context of representative hormone action. Newly identified components of these signal transduction pathways are defined in terms of the kinase system(s) involved. In addition, the insulin receptor and its tyrosine kinase and second messenger pathways are considered.

20.2— Hormones and the Hormonal Cascade System

The definition of a hormone has been expanded over the last several decades. Hormones secreted by endocrine glands were originally considered to represent all of the physiologically relevant hormones. Today, the term **hormone** refers to any substance in an organism that carries a signal to generate some sort of alteration at the cellular level. Thus **endocrine hormones** represent a class of hormones that arise in one tissue, or "gland," and travel a considerable distance through the circulation to reach a target cell expressing cognate receptors. **Paracrine hormones** arise from a cell and travel a relatively small distance to interact with their cognate receptors on another neighboring cell. **Autocrine hormones** are produced by the same cell that functions as the target for that hormone (neighboring cells may also be targets). Thus we can classify hormones based on their radii of action. Often, endocrine hormones that travel long distances to their target cells may be more stable than autocrine hormones that exert their effects over very short distances.

Cascade System Amplifies a Specific Signal

For many hormonal systems in higher animals, the signal pathway originates with the brain and culminates with the ultimate target cell. Figure 20.2 outlines the sequence of events in this cascade. A stimulus may originate in the external environment or within the organism in this cascade. This signal may be transmitted as an electrical pulse (action potential) or as a chemical signal or both. In many cases, but not all, such signals are forwarded to the limbic system and subsequently to the hypothalamus, the pituitary, and the target gland that secretes the final hormone. This hormone then affects various target cells to a degree that is frequently proportional to the number of cognate receptors

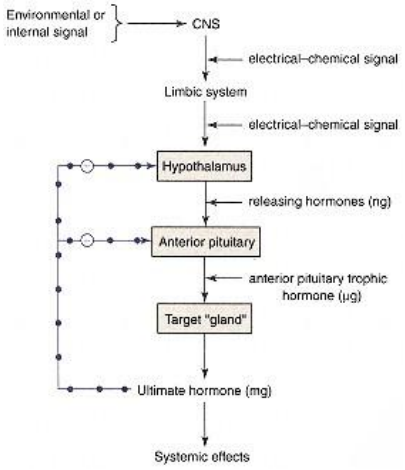


Figure 20.2
Hormonal cascade of signals from CNS to ultimate hormone.
The target "gland" refers to the last hormone-producing tissue in the cascade, which is stimulated by an appropriate anterior pituitary hormone. Examples would be thyroid gland, adrenal cortex, ovary, and testis. Ultimate hormone feeds back negatively on sites producing intermediate hormones in the cascade. Amounts [nanogram (ng), microgram (mg), and milligram (mg)] represent approximate quantities of hormone released.
Redrawn from Norman, A. W., and Litwack, G. *Hormones.* New York: Academic Press, 1987, p. 38.

expressed by that cell. This may be a true **cascade** in the sense that increasing amounts of hormones are generated at successive levels (hypothalamus, pituitary, and target gland) and also because the half-lives of these blood-borne hormones tend to become longer in progression from the hypothalamic hormone to the ultimate hormone. In the case of environmental stress, for example, there is a single stressor (change in temperature, noise, trauma, etc.). This stress results in a signal to the hippocampal structure in the limbic system that signals the hypothalamus to release a hypothalamic releasing hormone, corticotropin-releasing hormone (CRH), which is usually secreted in nanogram amounts and may have a $t_{1/2}$ in the bloodstream of several minutes. This hormone travels down a closed portal system to gain access to the **anterior pituitary**, where it binds to its cognate receptor in the cell membrane of corticotropic cells and initiates a set of metabolic changes resulting in the release of adrenocorticotropic hormone (ACTH) as well as β-lipotropin. This hormone, which is released in microgram amounts and has a longer $t_{1/2}$ than CRH, circulates in the bloodstream until it binds to its cognate receptors expressed in the membranes of cells located in the inner layer of the cortex of the adrenal gland (target gland). Here it affects metabolic changes leading to the synthesis and release in 24 h of the ultimate hormone, cortisol, in multimilligram amounts and this active glucocorticoid hormone has a substantial $t_{1/2}$ in blood. Cortisol is taken up by a wide variety of cells that express varying amounts of the intracellular glucocorticoid receptor. The ultimate hormone, in this case cortisol, feeds back negatively on cells of the anterior pituitary, hypothalamus, and perhaps higher levels to shut down the overall pathway in a process that is also mediated by the glucocorticoid receptor. At the **target cell** level these cortisol-receptor complexes mediate specific transcriptional responses and the individual hormonal effects summate to produce the systemic effects of the hormone. The cascade is represented in this example by a single environmental stimulus generating a series of hormones in progressively larger amounts and with increasing stabilities, and by the ultimate hormone that affects most of the cells in the body. Many other systems operate similarly, there being different specific **releasing hormones, anterior pituitary tropic hormones,** and ultimate hormones involved in the process. Clearly, the final number of target cells affected may be large or small depending on the distribution of receptors for each ultimate hormone.

A related system involves the **posterior pituitary hormones,** oxytocin and vasopressin (antidiuretic hormone), which are stored in the posterior pituitary gland but are synthesized in neuronal cell bodies located in the hypothalamus. This system is represented in Figure 20.3; elements of Figure 20.2 appear in the central vertical pathway. The posterior pituitary system branches to the right from the hypothalamus. Oxytocin and vasopressin are synthesized in separate cell bodies of hypothalamic neurons. More cell bodies dedicated to synthesis of vasopressin are located in the supraoptic nucleus and more cell bodies dedicated to synthesis of oxytocin are located in the paraventricular nucleus. Their release from the posterior pituitary gland along with **neurophysin**, a stabilizing protein, occurs separately via specific stimuli impinging on each of these types of neuronal cells.

There are highly specific signals dictating the release of polypeptide hormones along the cascade of this system. Thus there are a variety of **aminergic neurons** (secreting amine-containing substances like dopamine and serotonin) which connect to neurons involved in the synthesis and release of the **releasing hormones** of the hypothalamus. Releasing hormones are summarized in Table 20.1. These aminergic neurons fire depending on various types of internal or external signals and their activities account for **pulsatile release patterns** of certain hormones, such as the gonadotropin-releasing hormone (GnRH), and the **rhythmic cyclic release** of other hormones like cortisol.

Another prominent feature of the hormonal cascade (Figure 20.3) is the **negative feedback** system operating when sufficiently high levels of the ultimate hormone have been secreted into the circulation. Generally, there are three feedback loops—the **long feedback,** the **short feedback,** and the **ultra-**

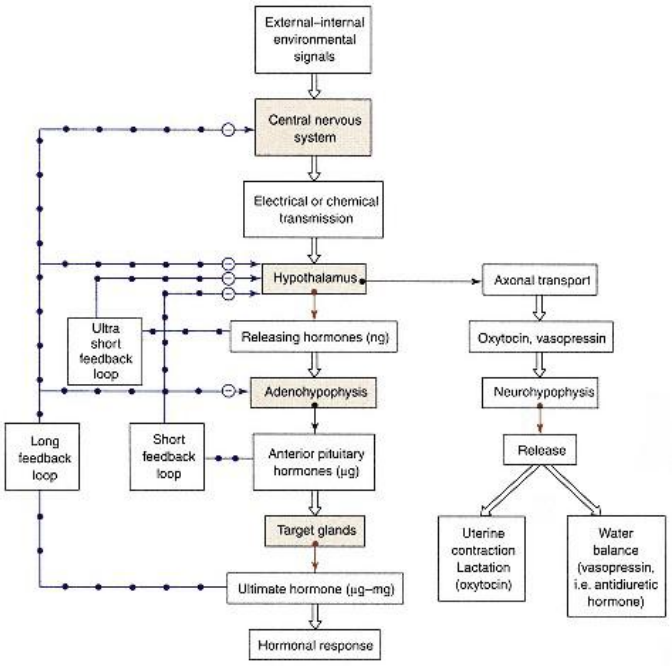


Figure 20.3
Many hormonal systems involve the hypothalamus.
Cascade of hormonal signals starting with an external or internal environmental signal. This is transmitted first to the CNS and may involve components of the limbic system, such as the hippocampus and amygdala. These structures innervate the hypothalamus in a specific region, which responds with secretion of a specific releasing hormone, usually in nanogram amounts. Releasing hormones are transported down a closed portal system connecting the hypothalamus and anterior pituitary, bind to cell membrane receptors and cause the secretion of specific anterior pituitary hormones, usually in microgram amounts. These access the general circulation through fenestrated local capillaries and bind to specific target gland receptors. The interactions trigger release of an ultimate hormone in microgram to milligram daily amounts, which generate the hormonal response by binding to receptors in several target tissues. In effect, this overall system is an amplifying cascade. Releasing hormones are secreted in nanogram amounts and they have short half-lives on the order of a few minutes. Anterior pituitary hormones are produced often in microgram amounts and have longer half-lives than releasing hormones. Ultimate hormones can be produced in daily milligram amounts with much longer half-lives. Thus the products of mass × half-life constitute an amplifying cascade mechanism. With respect to differences in mass of hormones produced from hypothalamus to target gland, the range is nanograms to milligrams, or as much as one million-fold. When the ultimate hormone has receptors in nearly every cell type, it is possible to affect the body chemistry of virtually every cell by a single environmental signal. Consequently, the organism is in intimate association with the external environment, a fact that we tend to underemphasize. Solid arrows indicate a secretory process. Long arrows studded with open or closed circles indicate negative feedback pathways (ultra-short, short, and long feedback loops).
Redrawn from Norman, A. W., and Litwack, G. *Hormones*.
New York: Academic Press, 1987, p. 102.

TABLE 20.1 Hypothalamic Releasing Hormones[a]

Releasing Hormone	*Number of Amino Acids in Structure*	*Anterior Pituitary Hormone Released or Inhibited*
Thyrotropin-releasing hormone (TRH)	3	Thyrotropin (TSH); can also release prolactin (PRL) experimentally
Gonadotropin-releasing hormone (GnRH)	10	Luteinizing and follicle-stimulating hormones (LH and FSH) from the same cell type; leukotriene C_4 (LTC_4) can also release LH and FSH by a different mechanism
Gonadotropin release-inhibiting factor (GnRIF)	12.2 kDa molecular weight	LH and FSH release inhibited
Corticotropin-releasing hormone (CRH)	41	ACTH, β-lipotropin (β-LPH), and some β-endorphin
Arginine vasopressin (AVP)	9	Stimulates CRH action in ACTH release
Angiotensin II (AII)	8	Stimulates CRH action in ACTH release; releases ACTH weakly
Somatocrinin (GRH)	44	Growth hormone (GH) release
Somatostatin (GIH)	14	GH release inhibited
Hypothalamic gastrin-releasing peptide		Inhibits release of GH and PRL
Prolactin-releasing factor (PRF)		Releases prolactin (PRL)
Prolactin release-inhibiting factor (PIF)		Evidence that a new peptide may inhibit PRL release; dopamine also inhibits PRL release and was thought to be PIF for some time; dopamine may be a secondary PIF: oxytocin may inhibit PRL release

[a] Melanocyte-stimulating hormone (MSH) is a major product of the *pars intermedia* (Figure 20.5) in the rat and is under the control of aminergic neurons. Humans may also secrete α-MSH from *pars intermedia*-like cells although this structure is anatomically indistinct in the human.

CLINICAL CORRELATION 20.1

Testing Activity of the Anterior Pituitary

Releasing hormones and chemical analogs, particularly of the smaller peptides, are now routinely synthesized. The gonadotropin-releasing hormone, a decapeptide, is available for use in assessing the function of the anterior pituitary. This is of importance when a disease situation may involve either the hypothalamus, the anterior pituitary, or the end organ. Infertility is an example of such a situation. What needs to be assessed is which organ is at fault in the hormonal cascade. Initially, the end organ, in this case the gonads, must be considered. This can be accomplished by injecting the anterior pituitary hormone LH or FSH. If sex hormone secretion is elicited, then the ultimate gland would appear to be functioning properly. Next, the anterior pituitary would need to be analyzed. This can be done by i.v. administration of synthetic GnRH; by this route GnRH can gain access to the gonadotropic cells of the anterior pituitary and elicit secretion of LH and FSH. Routinely, LH levels are measured in the blood as a function of time after the injection. These levels are measured by radioimmunoassay (RIA) in which radioactive LH or hCG is displaced from binding to an LH-binding protein by LH in the serum sample. The extent of the competition is proportional to the amount of LH in the serum. In this way a progress of response is measured that will be within normal limits or clearly deficient. If the response is deficient, the anterior pituitary cells are not functioning normally and are the cause of the syndrome. On the other hand, normal pituitary response to GnRH would indicate that the hypothalamus was nonfunctional. Such a finding would prompt examination of the hypothalamus for conditions leading to insufficient availability/production of releasing hormones. Obviously, the knowledge of hormone structure and the ability to synthesize specific hormones permit the diagnosis of these disease states.

Marshall, J. C., and Barkan, A. L. Disorders of the hypothalamus and anterior pituitary. In: W. N. Kelley (Ed.), *Internal Medicine*. New York: Lippincott, 1989, p. 2159; and Conn, P. M. The molecular basis of gonadotropin-releasing hormone action. *Endocr. Rev*. 7:3, 1986.

short feedback loops. In the long feedback loop, the final hormone binds a cognate receptor in/on cells of the anterior pituitary, hypothalamus, and CNS to prevent further elaboration of hormones from those cells that are involved in the cascade. The short feedback loop is accounted for by the pituitary hormone that feeds back negatively on the hypothalamus operating through a cognate receptor. In ultra-short feedback loops the hypothalamic releasing factor feeds back at the level of the hypothalamus to inhibit further secretion of this releasing factor. These mechanisms provide tight controls on the operation of the cascade, responding to stimulating signals as well as negative feedback, and render this system highly responsive to the hormonal milieu. Clinical Correlation 20.1 describes approaches for testing the responsiveness of the anterior pituitary gland.

Polypeptide Hormones of the Anterior Pituitary

The polypeptide hormones of the anterior pituitary are shown in Figure 20.4 together with their controlling hormones from the hypothalamus. The major

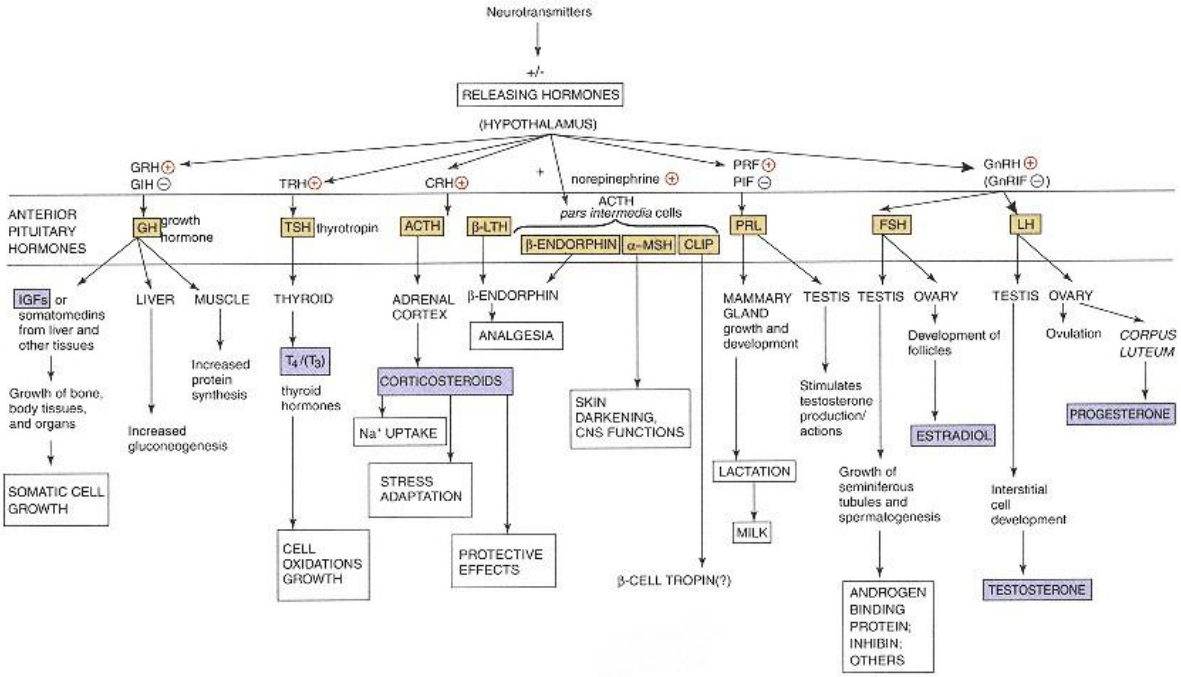


Figure 20.4
Overview of anterior pituitary hormones with hypothalamic releasing hormones and their actions.

CLINICAL CORRELATION 20.2

Hypopituitarism

The hypothalamus is connected to the anterior pituitary by a delicate stalk that contains the portal system through which releasing hormones, secreted from the hypothalamus, gain access to the anterior pituitary cells. In the cell membranes of these cells are specific receptors for releasing hormones. In most cases, different cells express different releasing hormone receptors. The connection between the hypothalamus and anterior pituitary can be disrupted by trauma or tumors. Trauma can occur in automobile accidents or other local damaging events that may result in severing of the stalk and preventing the releasing hormones from reaching their target anterior pituitary cells. When this happens, the anterior pituitary cells no longer have their signaling mechanism for the release of anterior pituitary hormones. In the case of tumors of the pituitary gland, all of the anterior pituitary hormones may not be shut off to the same degree or the secretion of some may disappear sooner than others. In any case, if hypopituitarism occurs this condition may result in a life-threatening situation in which the clinician must determine the extent of loss of pituitary hormones, especially ACTH. Posterior pituitary hormones—oxytocin and vasopressin—may also be lost, precipitating a problem of excessive urination (vasopressin deficiency) that must be addressed. The usual therapy involves administration of the end organ hormones, such as thyroid hormone, cortisol, sex hormones, and progestin; with female patients it is also necessary to maintain the ovarian cycle. These hormones can easily be administered in oral form. Growth hormone deficiency is not a problem in the adult but would be an important problem in a growing child. The patient must learn to anticipate needed increases of cortisol in the face of stressful situations. Fortunately, these patients are usually maintained in reasonably good condition.

Marshall, J. C., and Barkan, A. L. Disorders of the hypothalamus and anterior pituitary. In: W. N. Kelley (Ed.), *Internal Medicine*. New York: Lippincott, 1989, p. 2159; and Robinson, A. G. Disorders of the posterior pituitary. In: W. N. Kelley (Ed.), *Internal Medicine*, New York: Lippincott, 1989, p. 2172.

hormones of the anterior pituitary are growth hormone (GH), thyrotropin or thyroid-stimulating hormone (TSH), adrenocorticotropic hormone (ACTH), β-lipotropin (β-LTH), β-endorphin (from *pars intermedia*-like cells), α-MSH (from *pars intermedia*-like cells), β-MSH (from *pars intermedia*-like cells), corticotropin-like intermediary peptide (CLIP; from *pars intermedia*-like cells), prolactin (PRL), follicle-stimulating hormone (FSH), and luteinizing hormone (LH). Of these, all are single polypeptide chains, except TSH, FSH, and LH, which are dimers that share a similar or identical subunit, the α subunit. Since the intermediate lobe in humans is rudimentary, the circulating levels of free α- and β-MSH are relatively low. It is of interest, particularly in the human, that MSH receptors recognize and are activated by ACTH, since the first 13 amino acids of ACTH contain the α-MSH sequence. For this reason, ACTH may be an important contributing factor to skin pigmentation and may exceed the importance of MSH, especially in conditions where the circulating level of ACTH is high. The clinical consequences of hypopituitarism are presented in Clin. Corr. 20.2.

20.3— Major Polypeptide Hormones and Their Actions

Since cellular communication is so specific, it is not surprising that there are a large number of hormones in the body and new hormones continue to be discovered. Limitations of space permit a summary of only a few of the well-characterized hormones. Table 20.2 presents some major polypeptide hormones and their actions. By inspection of Table 20.2 it becomes evident that many hormones cause the release of other substances, some of which may themselves be hormones. This is particularly the case for hormonal systems that are included in cascades like that presented in Figures 20.2 and 20.3. Other activities of receptor-hormone complexes located in cell membranes are to increase the flux of ions into cells, particularly calcium ions, and to activate or suppress activities of enzymes in contact with the receptor or a transducing protein with which the receptor interacts. Examples of these kinds of activities are discussed later in this chapter. In the functioning of most membrane–receptor complexes,

TABLE 20.2 Important Polypeptide Hormones in the Body and Their Actions[a]

Source	*Hormone*	*Action*
Hypothalamus	Thyrotropin-releasing hormone (TRH)	Acts on thyrotrope to release TSH
	Gonadotropin-releasing hormone (GnRH)	Acts on gonadotrope to release LH and FSH from the same cell
	Growth hormone-releasing hormone or somatocrinin (GRH)	Acts on somatotrope to release GH
	Growth hormone release inhibiting hormone or somatostatin (GIH)	Acts on somatotrope to prevent release of GH
	Corticotropin-releasing hormone (CRH) Vasopressin is a helper hormone to CRH in releasing ACTH; angiotensin II also stimulates CRH action in releasing ACTH	Acts on corticotrope to release ACTH and β-lipotropin
	Prolactin-releasing factor (PRF) (not well established)	Acts on lactotrope to release PRL
	Prolactin release inhibiting factor (PIF) (not well established; may be a peptide hormone under control of dopamine or may be dopamine itself)	Acts on lactotrope to inhibit release of PRL
Anterior pituitary	Thyrotropin (TSH)	Acts on thyroid follicle cells to bring about release of T_4 (T_3)
	Luteinizing hormone (LH) (human chorionic gonadotropin, hCG, is a similar hormone from the placenta)	Acts on Leydig cells of testis to increase testosterone synthesis and release; acts on corpus luteum of ovary to increase progesterone production and release
	Follicle-stimulating hormone (FSH)	Acts on Sertoli cells of seminiferous tubule to increase proteins in sperm and other proteins; acts on ovarian follicles to stimulate maturation of ovum and production of estradiol
	Growth hormone (GH)	Acts on a variety of cells to produce IGFs (or somatomedins), cell growth, and bone sulfation
	Adrenocorticotropic hormone (ACTH)	Acts on cells in the adrenal gland to increase cortisol production and secretion
	β-Endorphin	Acts on cells and neurons to produce analgesic and other effects
	Prolactin (PRL)	Acts on mammary gland to cause differentiation of secretory cells (with other hormones) and to stimulate synthesis of components of milk
	Melanocyte-stimulating hormone (MSH)	Acts on skin cells to cause the dispersion of melanin (skin darkening)
Ultimate gland hormones	Insulin-like growth factors (IGF)	Respond to GH and produce growth effects by stimulating cell mitosis
	Thyroid hormone (T_4/T_3) (amino acid-derived hormone)	Responds to TSH and stimulates oxidation in many cells
	Opioid peptides	May derive as breakdown products of γ-lipotropin or β-endorphin or from specific gene products; can respond to CRH or dopamine and may produce analgesia and other effects
	Inhibin	Responds to FSH in ovary and in Sertoli cell; regulates secretion of FSH from anterior pituitary. Second form of inhibin (activin) may stimulate FSH secretion
	Corticotropin-like intermediary peptide (CLIP)	Derives from intermediate pituitary by degradation of ACTH; contains β-cell tropin activity, which stimulates insulin release from β cells in presence of glucose

(continued)

TABLE 20.2 *(Continued)*

Source	*Hormone*	*Action*
Peptide hormones responding to other signals than anterior pituitary hormones	Arginine vasopressin (AVP; antidiuretic hormone, ADH)	Responds to increase in osmoreceptor, which senses extracellular [Na^+]; increases water reabsorption from distal kidney tubule
	Oxytocin	Responds to suckling reflex and estradiol; causes milk "let down" or ejection in lactating female, involved in uterine contractions of labor; luteolytic factor produced by *corpus luteum*; decreases steroid synthesis in testis
β Cells of pancreas respond to glucose and other blood constituents to release insulin	Insulin	Increases tissue utilization of glucose
α Cells of pancrease respond to low levels of glucose and falling serum calcium	Glucagon	Decreases tissue utilization of glucose to elevate blood glucose
Derived from circulating blood protein by actions of renin and converting enzyme	Angiotensin II and III (AII and AIII)	Renin initially responds to decreased blood volume or decreased [Na^+] in the *macula densa* of the kidney. AII/AIII stimulate outer layer of adrenal cells to synthesize and release aldosterone
Released from heart atria in response to hypovolemia; regulated by other hormones	Atrial natriuretic factor (ANF) or atriopeptin	Acts on outer adrenal cells to decrease aldosterone release; has other effects also
Generates from plasma, gut, or other tissues	Bradykinin	Modulates extensive vasodilation resulting in hypotension
Hypothalamus and intestinal mucosa	Neurotensin	Effects on gut; may have neurotransmitter actions
Hypothalamus, CNS, and intestine	Substance P	Pain transmitter, increases smooth muscle contractions of the GI tract
Nerves and endocrine cells of gut; hypothermic hormone	Bombesin	Increases gastric acid secretion
	Cholecystokinin (CCK)	Stimulates gallbladder contraction and bile flow; increases secretion of pancreatic enzymes
Stomach antrum	Gastrin	Increases secretion of gastric acid and pepsin
Duodenum at pH values below 4.5	Secretin	Stimulates pancreatic acinar cells to release bicarbonate and water to elevate duodenal pH
Hypothalamus and GI tract	Vasointestinal peptide (VIP)	Acts as a neurotransmitter in peripheral autonomic nervous system; relaxes smooth muscles of circulation; increases secretion of water and electrolytes from pancreas and gut
Kidney	Erythropoietin	Acts on bone marrow for terminal differentiation and initiates hemoglobin synthesis
Ovarian corpus luteum	Relaxin	Inhibits myometrial contractions; its secretion increases during gestation
	Human placental lactogen (hPL)	Acts like PRL and GH because of large amount of hPL produced
Salivary gland	Epidermal growth factor	Stimulates proliferations of cells derived from ectoderm and mesoderm together with serum; inhibits gastric secretion
Thymus	Thymopoietin (α-thymosin)	Stimulates phagocytes; stimulates differentiation of precursors into immune competent T cells

(table continued on next page)

TABLE 20.2 *(Continued)*

Source	*Hormone*	*Action*
Parafollicular C cells of thyroid gland	Calcitonin (CT)	Lowers serum calcium
Parathyroid glands	Parathyroid hormone (PTH)	Stimulates bone resorption; stimulates phosphate excretion by kidney; raises serum calcium levels
Endothelial cells of blood vessels	Endothelin	Vasoconstriction

Source: Part of this table is reproduced from Norman, A. W., and Litwack, G. *Hormones*. Orlando, FL: Academic Press, 1987.

[a] This is only a partial list of polypeptide hormones in humans. TSH, thyroid-stimulating hormone or thyrotropin; LH, luteinizing hormone, FSH, follicle-stimulating hormone; GH, growth hormone; ACTH, adenocorticotropic hormone; PRL, prolactin; T_4, thyroid hormone (also T); IGF, insulin-like growth factor. For the releasing hormones and for some hormones in other categories, the abbreviation may contain "H" at the end when the hormone has been well characterized, and "F" in place of H to refer to "Factor" when the hormone has not been well characterized. Names of hormones may contain "tropic" or "trophic" endings, tropic is mainly used here. Tropic refers to a hormone generating a change, whereas trophic refers to growth promotion. Both terms can refer to the same hormone at different stages of development. Many of these hormones have effects in addition to those listed here.

stimulation of enzymes or flux of ions is followed by a chain of events, which may be described as intracellular cascades, during which a high degree of amplification is obtained.

20.4— Genes and Formation of Polypeptide Hormones

Genes for polypeptide hormones contain the information for the hormone and the control elements upstream of the transcriptionally active sequence. In some cases, more than one hormone is encoded in a gene. One example is **proopiomelanocortin,** a hormone precursor that encodes the following hormones: ACTH, β-lipotropin, and other hormones like γ-lipotropin, γ-MSH, α-MSH, CLIP, β-endorphin, and potentially β-MSH and enkephalins. In the case of the posterior pituitary hormones, oxytocin and vasopressin, information for these hormones are each encoded on a separate gene together with information for each respective **neurophysin,** a protein that binds to the completed hormone and stabilizes it.

Proopiomelanocortin Is a Precursor Polypeptide for Eight Hormones

Proopiomelanocortin, as schematized in Figure 20.5, can generate at least eight hormones from a single gene product. All products are not expressed simultaneously in a single cell type, but occur in separate cells based on their content of specific proteases required to cleave the propeptide, specific metabolic controls, and the presence of different positive regulators. Thus, while proopiomelanocortin is expressed in both the corticotropic cell of the anterior pituitary and the *pars intermedia* cell, the stimuli and products are different as summarized in Table 20.3. The *pars intermedia* is a discrete anatomical structure located between the anterior and posterior pituitary in the rat (Figure 20.6). In the human, however, the *pars intermedia* is not a discrete anatomical structure, although the cell type may be present in the equivalent location.

Many Polypeptide Hormones Are Encoded Together in a Single Gene

An example of another gene and gene products encoding more than one peptide are the genes for vasopressin and oxytocin and their accompanying neurophysin proteins, products that are released from the posterior pituitary upon specific stimulation. In much the same manner that ACTH and β-lipotropin

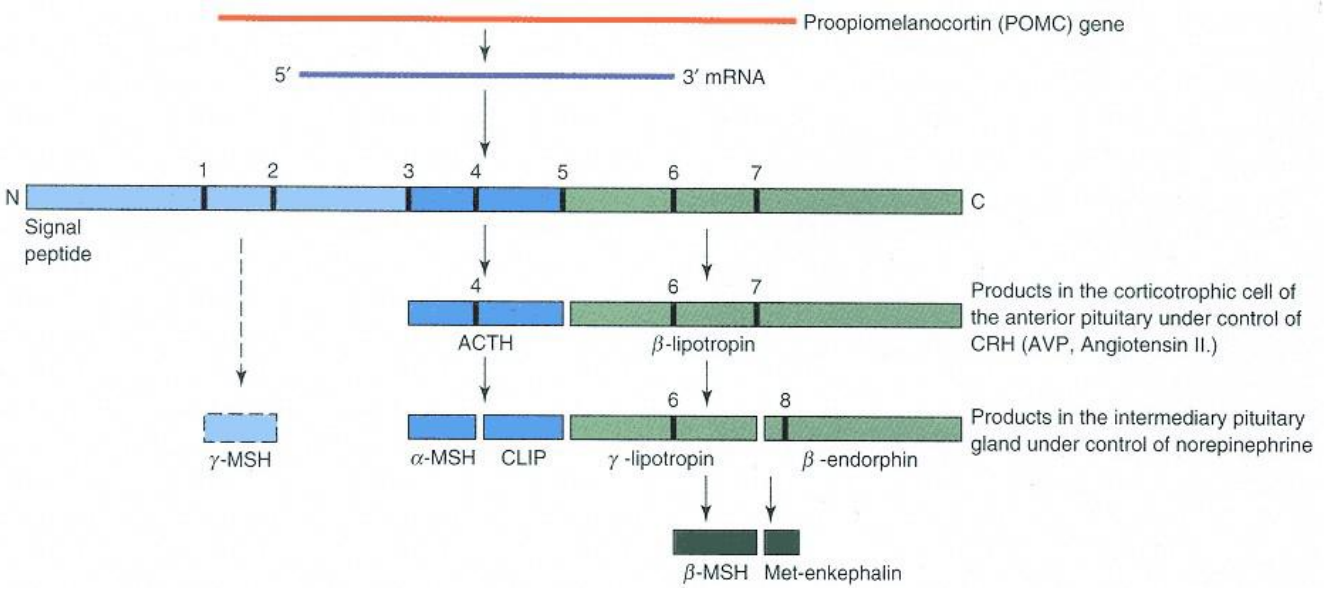


Figure 20.5
Proopiomelanocortin is a polypeptide product encoded by a single gene.
The dark vertical bars represent proteolytic cleavage sites for specific enzymes. The cleavage sites are Arg-Lys, Lys-Arg, or Lys-Lys. Some specificity also may be conferred by neighboring amino acid residues. In the corticotropic cell of the anterior pituitary, enzymes are present that cleave at sites 3 and 5, releasing the major products, ACTH and β-lipotropin, into the general circulation. In the pars intermedia, especially in vertebrates below humans, these products are further cleaved at major sites 4, 6, and 7 to release α-MSH, CLIP, γ-lipotropin, and β-endorphin into the general circulation. Some β-lipotropin arising in the corticotroph may be further degraded to form β-endorphin. These two cell types appear to be under separate controls. The corticotropic cell of the anterior pituitary is under the positive control of the CRH and its auxiliary helpers, arginine vasopressin (AVP) and angiotensin II. AVP by itself does not release ACTH but enhances the action of CRH in this process. The products of the intermediary pituitary, α-MSH, CLIP (corticotropin-like intermediary peptide), γ-lipotropin, and β-endorphin, are under the positive control of norepinephrine, rather than CRH, for release. Obviously there must exist different proteases in these different cell types in order to generate a specific array of hormonal products. β-Endorphin also contains a pentapeptide, enkephalin, which potentially could be released at some point (hydrolysis at 8).

TABLE 20.3 Summary of Stimuli and Products of Proopiomelanocortin[a]

Cell type	Corticotroph	Pars intermedia
Stimulus	CRH (+) (Cortisol (-))	Dopamine (-) Norepinephrine (+)
Auxiliary stimulus	AVP, AII	
Major products	ACTH, β-lipotropin (β-endorphin)	αMSH, CLIP, γ-lipotropin, β-endorphin

[a] CRH, corticotropin-releasing hormone; AVP, arginine vasopressin; AII, angiotensin II; ACTH, adrenocorticotropin; α-MSH, α melanocyte-stimulating hormone; CLIP, corticotropin-like intermediary peptide.

Note: Although there are pars intermedia cells in the human pituitary gland, they do not represent a distinct lobe.

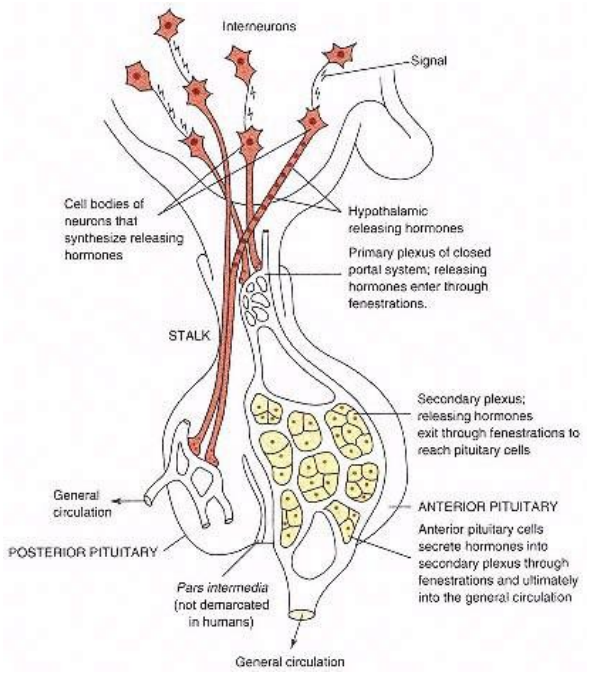


Figure 20.6
The hypothalamus with nuclei in various locations in which the hypothalamic releasing hormones are synthesized.
Shown is the major vascular network consisting of a primary plexus where releasing hormones enter its circulation through fenestrations and the secondary plexus in the anterior pituitary where the releasing hormones are transported out of the circulation, again through fenestrations in the vessels, to the region of the anterior pituitary target cells. Also shown are the resultant effects of the actions of the hypothalamic releasing hormones causing the secretion into the general circulation of the anterior pituitary hormones.
Adapted from Norman, A. W., and Litwack, G. *Hormones*. New York: Academic Press, 1987, p. 104.

(β-LPH) are split out of the proopiomelanocortin precursor peptide, so are the products vasopressin, neurophysin II, and a glycoprotein of as yet unknown function split out of the vasopressin precursor. A similar situation exists for oxytocin and neurophysin I (Figure 20.7).

Vasopressin and **neurophysin II** are released by the activity of baroreceptors and osmoreceptors, which sense a fall in blood pressure or a rise in extracellular sodium ion concentration, respectively. Generally, **oxytocin** and

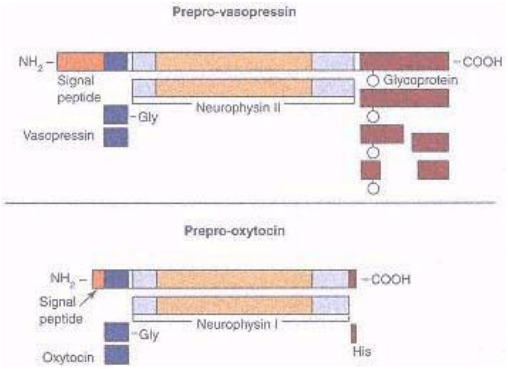


Figure 20.7
Prepro-vasopressin and prepro-oxytocin.
Proteolytic maturation proceeds from top to bottom for each precursor. The organization of the gene translation products is similar in either case except that a glycopeptide is included on the proprotein of vasopressin in the C-terminal region. Orange bars of the neurophysin represent conserved amino acid regions; gray bars represent variable C and N termini.
Redrawn with permission from Richter, D. VP and OT are expressed as polyproteins. *Trends* Biochem. 8:278, 1983.

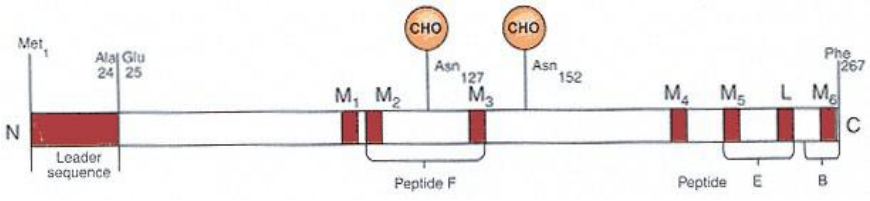


Figure 20.8
Model of enkephalin precursor.
The distribution of Met-enkephalin sequences (M_1–M_6) and Leu-enkephalin (L) sequences within the protein precursor of bovine adrenal medulla. CHO, potential carbohydrate attachment sites. Redrawn from Comb, M., Seeburg, P. H., Adelman, J., Eiden, L., and Herbert, E. *Nature* 295:663, 1982.

neurophysin I are released from the posterior pituitary by the suckling response in lactating females or by other stimuli mediated by a specific cholinergic mechanism. Oxytocin–neurophysin I release can be triggered by injection of estradiol. Release of vasopressin–neurophysin II can be stimulated by administration of nicotine. The two separate and specific releasing agents, estradiol and nicotine, prove that oxytocin and vasopressin, together with their respective neurophysins, are synthesized and released from different cell types. Although oxytocin is well known for its milk-releasing action in the lactating female, in the male it seems to have a separate role associated with an increase in testosterone synthesis in the testes.

Other polypeptide hormones are being discovered that are co-encoded together by a single gene. An example is the discovery of the gene encoding GnRH, a decapeptide that appears to reside to the left of a gene for the GnRH-associated peptide (GAP), which, like dopamine, may be capable of inhibiting prolactin release. Thus both hormones—GnRH and the prolactin release inhibiting factor GAP—appear to be co-secreted by the same hypothalamic cells.

Multiple Copies of a Hormone Can Be Encoded on a Single Gene

An example of multiple copies of a single hormone encoded on a single gene is the gene product for enkephalins located in the chromaffin cell of the adrenal medulla. **Enkephalins** are pentapeptides with opioid activity; methionine-enkephalin (Met-ENK) and leucine-enkephalin (Leu-ENK) have the structures:

Tyr-Gly-Gly-Phe-Met (Met-ENK)

Tyr-Gly-Gly-Phe-Leu (Leu-ENK)

A model of enkephalin precursor in adrenal medulla is presented in Figure 20.8, which encodes several Met-ENK (M) molecules and a molecule of Leu-

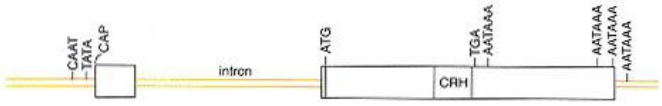


Figure 20.9
Nucleic acid sequence for rat proCRH genes.
Representation of the rat proCRH gene. Exons are shown as blocks and the intron by a double red line. The TATA and CAAT sequence, putative cap site, translation initiation ATG, translation terminator TGA, and poly(A) addition signals (AATAAA) are indicated. The location of the CRH peptide is indicated by CRH. Redrawn from Thompson, R. D., Seasholz, A. F., and Herbert, E. *Mol. Endocrinol.* 1:363, 1987.

ENK (L). Again, the processing sites to release enkephalin molecules from the protein precursor involve Lys–Arg, Arg–Arg, and Lys–Lys bonds.

Many genes for hormones are constructed to encode only one hormone and this may be the general situation. An example of a single hormone gene is shown in Figure 20.9. In this case the information for the hormone CRH is contained in the second exon and the information in the first exon is not expressed. Having cDNAs for use as probes that contain the information for expression of CRH allows for the localization of the hormone in tissues. Previously it was thought that the hormone should be restricted to the hypothalamus, the anterior pituitary, and the stalk, which contains the closed vascular transporting system (Figure 20.6). However, RNA extracts from different tissues probed with this DNA reveal the location of CRH mRNA in testis, brain stem, and adrenal gland in addition to pituitary and hypothalamus. The presence of the hormone in extrahypothalamic–pituitary axis tissues and its functions there are subjects of active investigation.

20.5— Synthesis of Amino Acid-Derived Hormones

Many hormones and neurotransmitters are derived from amino acids, principally from tyrosine and phenylalanine. Glutamate, aspartate, and other compounds are important neurotransmitter substances as well. Although there may be some confusion about which compounds are neurotransmitters and which are hormones, it is clear that epinephrine from the adrenal medulla is a hormone, whereas norepinephrine is a neurotransmitter. This section considers epinephrine and thyroxine or triiodothyronine. The other biogenic amines, such as dopamine, which are considered to be neurotransmitters, are discussed in Chapter 22.

Epinephrine Is Synthesized from Phenylalanine/Tyrosine

The synthesis of epinephrine occurs in the adrenal medulla. A number of steroid hormones, including aldosterone, cortisol, and dehydroepiandrosterone (sulfate), are produced in the adrenal cortex and are discussed in Chapter 21. The biochemical reactions leading to the formation of **epinephrine** from tyrosine or phenylalanine are presented in Figure 20.10. Epinephrine is a principal hormone secreted from the adrenal medulla chromaffin cell along with some norepinephrine, enkephalins, and some of the enzyme *dopamine-β*-hydroxylase. Secretion of epinephrine is signaled by the neural response to stress, which is transmitted to the adrenal medulla by way of a preganglionic acetylcholinergic neuron (Figure 20.11). Release of acetylcholine by the neuron increases the availability of intracellular calcium ion, which stimulates exocytosis and release of the material stored in the **chromaffin granules** (Figure 20.11b). This overall system of epinephrine synthesis, storage, and release from the adrenal medulla is regulated by neuronal controls and also by glucocorticoid hormones synthesized in and secreted from the adrenal cortex in response to stress. Since the products of the adrenal cortex are transported through the adrenal medulla on their way out to the general circulation, cortisol becomes elevated in the medulla and induces **phenylethanolamine *N*-methyltransferase (PNMT),** a key enzyme catalyzing the conversion of norepinephrine to epinephrine. Thus, in biochemical terms, the stress response at the level of the adrenal cortex ensures the production of epinephrine from the adrenal medulla (Figure 20.12). Presumably, epinephrine once secreted into the bloodstream not only affects α receptors of hepatocytes to ultimately increase blood glucose levels as indicated, but also interacts with α receptors on vascular smooth muscle cells and on pericytes to cause cellular contraction and increase blood pressure.

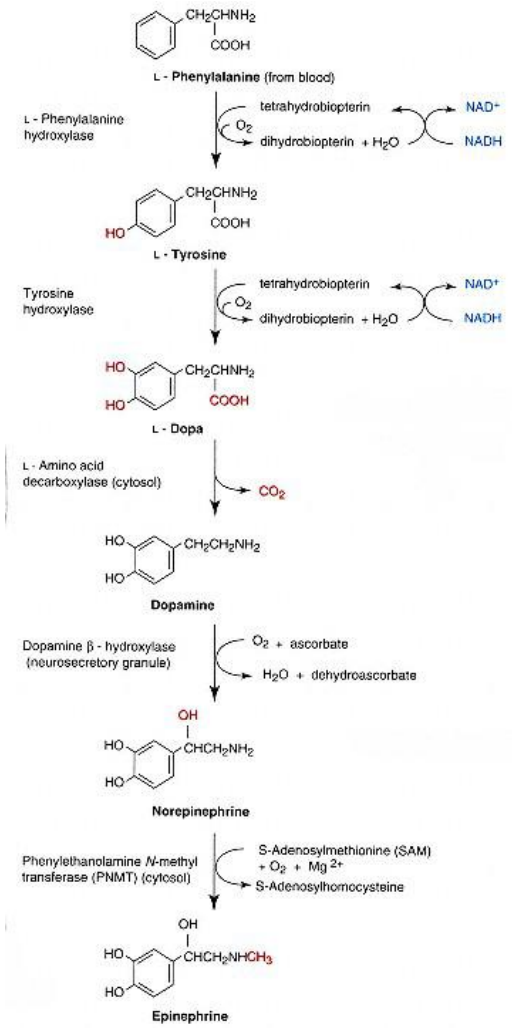


Figure 20.10
Biochemical steps in synthesis of epinephrine by chromaffin cell of adrenal medulla.

Synthesis of Thyroid Hormone Requires Incorporation of Iodine into a Tyrosine of Thyroglobulin

An outline of the biosynthesis and secretion of thyroid hormone, **tetraiodo-L-thyronine (T_4),** also called **thyroxine,** and its active cellular counterpart, **triiodo-L-thyronine (T_3)** (structures presented in Figure 20.13) is presented in Figure 20.14. The thyroid gland is differentiated to concentrate iodide from the blood and through the series of reactions shown in Figures 20.13 and 20.14, monoiodotyrosine (MIT), diiodotyrosine (DIT), T_4, and T_3 are produced within **thyroglobulin** (TG). Thus the iodinated amino acids and thyronines are stored in the thyroid follicle as part of thyroglobulin. Recent work indicates that there

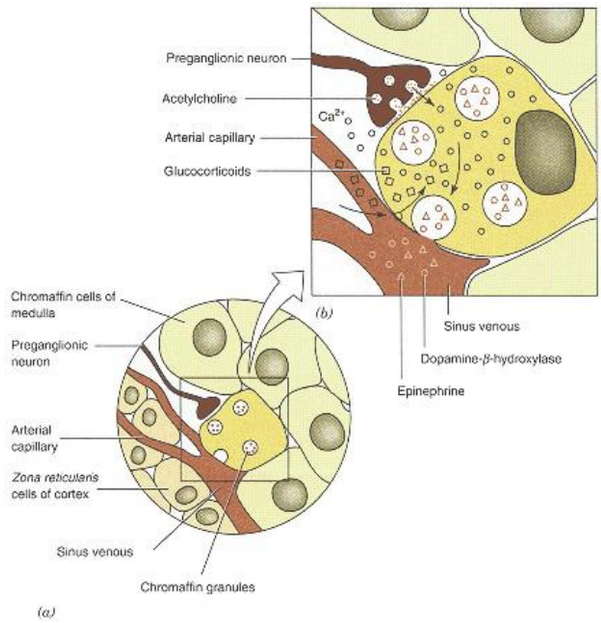


Figure 20.11
Relationship of adrenal medulla chromaffin cells to preganglionic neuron innervation and the structural elements involved in the synthesis of epinephrine and the discharge of catecholamines in response to acetylcholine.
(a) Functional relationship between cortex and medulla for control of synthesis of adrenal catecholamines. Glucocorticoids that stimulate enzymes catalyzing the conversion of norepinephrine to epinephrine reach the chromaffin cells from capillaries shown in (b).
(b) Discharge of catecholamines from storage granules in chromaffin cells after nerve fiber stimulation, resulting in the release of acetylcholine. Calcium enters the cells as a result, causing the fusion of granular membranes with the plasma membrane and exocytosis of the contents.
Reprinted with permission from Krieger, D. T., and Hughes, J. C. (Eds.). *Neuroendocrinology*. Sunderland, MA: Sinauer Associates, 1980.

are hot spots (regions for very active iodination) in the thyroglobulin sequence for the incorporation of iodine. Apparently, the sequences around iodotyrosyls occur in three consensus groups: Glu/Asp-Tyr, associated with the synthesis of thyroxine or iodotyrosines; Ser/Thr-Tyr-Ser, associated with the synthesis of

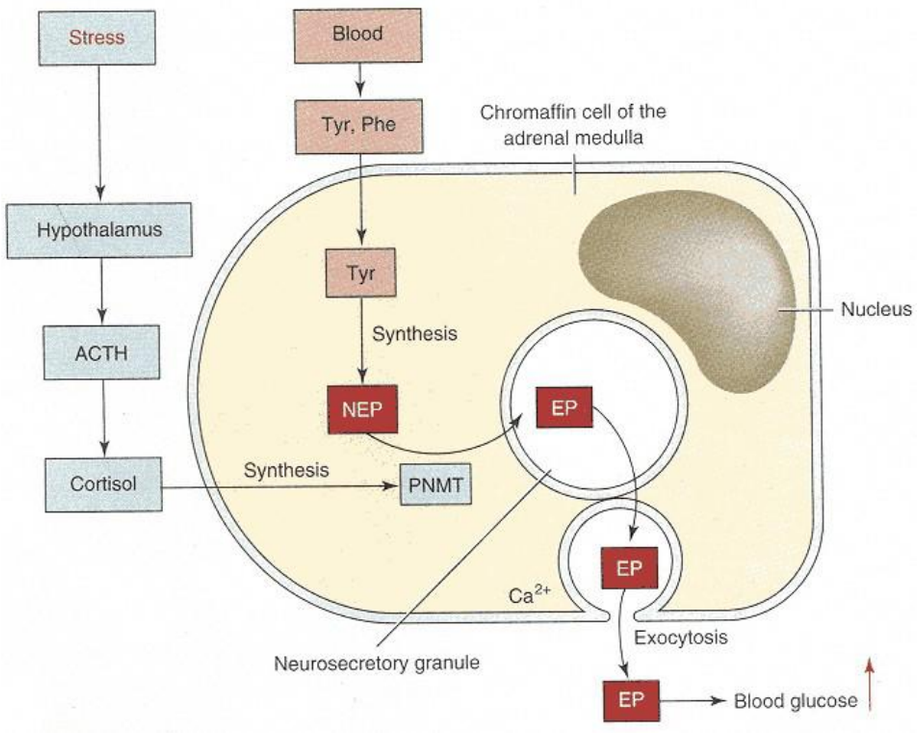


Figure 20.12
Biosynthesis, packaging, and release of epinephrine in the adrenal medulla chromaffin cell.
PNMT, phenylethanolamine N-methyltransferase; EP, epinephrine; NEP, norepinephrine. Neurosecretory granules contain epinephrine, dopamine β-hydroxylase, ATP, Met- or Leu-enkephalin, as well as larger enkephalin-containing peptides or norepinephrine in place of epinephrine. Epinephrine and norepinephrine are stored in different cells. Enkephalins could also be contained in separate cells, although that is not completely clear.
Adapted from Norman, A. W., and Litwack, G. *Hormones*. New York: Academic Press, 1987, p. 464.

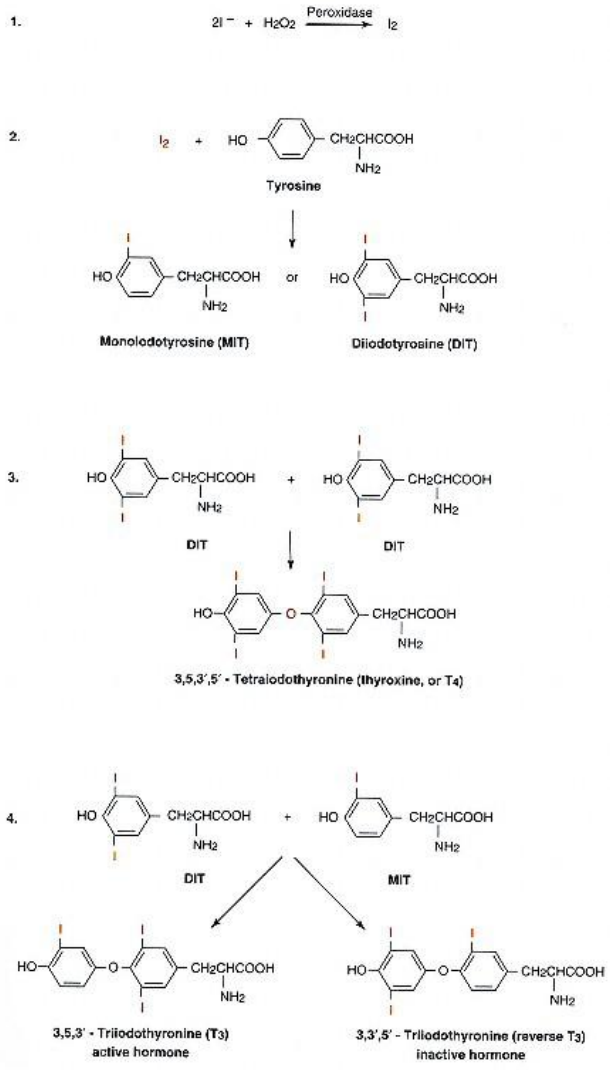


Figure 20.13
Synthesis and structures of thyroid hormones T_4, T_3, and reverse T_3.
Step 1, oxidation of iodide:
Step 2, iodination of tyrosine residues;
Step 3, coupling of DIT to DIT;
Step 4, coupling of DIT to MIT
(coupling may be intramolecular or intermolecular).

iodothyronine and iodotyrosine; and Glu-X-Tyr, associated with the remaining iodotyrosyls in the sequence. As depicted in Figure 20.14, secretion of T_3 and T_4 into the bloodstream requires endocytosis of the thyroglobulin stored in the follicle and subsequent proteolysis within the epithelial cell. The released DIT and MIT are then deiodinated and the released iodide ions are recycled and reutilized for hormone synthesis.

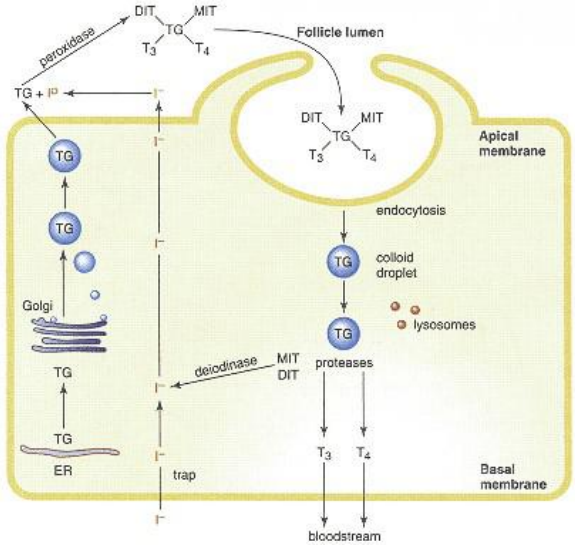


Figure 20.14
Cellular mechanisms for T_3 and T_4 release into the bloodstream.
Iodide trapping at basal membrane of thyroid epithelium concentrates iodide approximately 30-fold. Secretion of T_3 and T_4 into bloodstream requires endocytosis of thyroglobulin and subsequent proteolysis. DIT and MIT are deiodinated and the released iodide ions are reutilized for hormone synthesis.

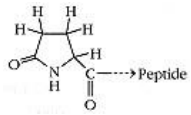
20.6— Inactivation and Degradation of Hormones

Most polypeptide hormones are degraded to amino acids by hydrolysis, which presumably occurs in the **lysosome.** Partial hydrolysis by proteinases is a principal pathway for degradation. Certain hormones, however, contain modified amino acids; for example, among the hypothalamic releasing hormones, the N-terminal amino acid can be **cycloglutamic acid** (or pyroglutamic acid) (Table 20.4) and a C-terminal amino acid amide. Some of the releasing hormones

TABLE 20.4 Hypothalamic Releasing Hormones Containing an N-Terminal Pyroglutamate, [a] a C-Terminal Amino Acid Amide, or Both

Hormone	*Sequence*[b]
Thyrotropin-releasing hormone (TRH)	pGlu-H-Pro-NH_2
Gonadotropin-releasing hormone (GnRH)	pGlu-HWSYGLRP-Gly-NH_2
Corticotropin-releasing hormone (CRH)	SQEPPISLDLTFHLLREVLEMTKADQLAQQAHSNRKLLDI-Ala-NH_2
Somatocrinin (GRH)	YADAIFTNSYRKVLGQLSARKLLQDIMSRQQGESNQERG-ARAR- Leu-NH_2

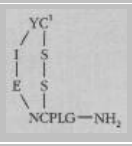
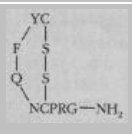
[a] The pyroglutamate structure is



[b] Single-letter abbreviations used for amino acids: Ala, A; Arg, R; Asn, N; Asp, D; Cys, C; Glu, E, Gln, Q; Gly, G; His, H; Ile, I; Leu, L; Lys, K; Met, M; Phe, F; Pro, P; Ser, S; Thr, T; Trp, W; Tyr, Y; Val, V.

that have either or both of these amino acid derivatives are listed in Table 20.4. Apparently, breakage of the cyclic glutamate ring or cleavage of the C-terminal amide can lead to inactivation of many of these hormones and such enzymic activities have been reported in blood. This activity probably accounts, in part, for the short half-life of many of these hormones.

TABLE 20.5 Examples of Hormones Containing a Cystine Disulfide Bridge Structure

Hormone	*Sequence*[a]
Somatostatin	
Oxytocin	
Arginine vasopressin	

[a] Letters refer to single-letter amino acid abbreviations (see Table 20.4)

Some hormones contain a ring structure joined by a cystine disulfide bond. A few examples are given in Table 20.5. Peptide hormones, such as those shown in Table 20.5, may be degraded initially by the random action of **cystine aminopeptidase** and **glutathione transhydrogenase** as shown in Figure 20.15. Alternatively, as has been suggested in the case of oxytocin, the peptide may be broken down through partial proteolysis to shorter peptides, some of which may have hormonal actions on their own. Maturation of **prohormones** in many cases involves proteolysis, which may be considered as a degradation process in the sense that the prohormone is degraded to active forms (e.g., Figure 20.5), although degradation is usually thought of as the reduction of active peptides to inactive ones.

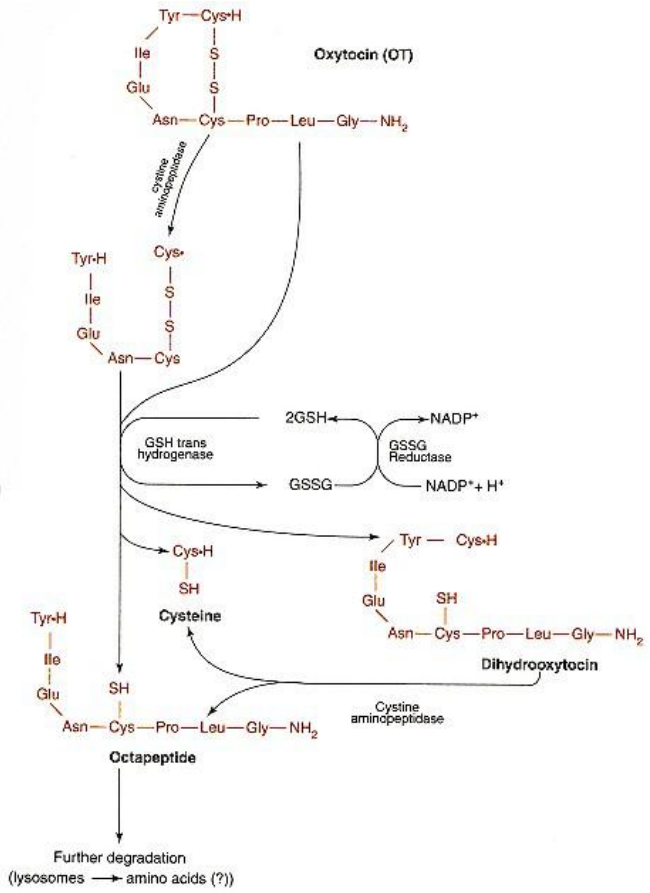


Figure 20.15
Degradation of posterior pituitary hormones.
Oxytocin transhydrogenase is similar to degrading enzymes for insulin; presumably, these enzymes also degrade vasopressin.
Redrawn from Norman, A. W., and Litwack, G. *Hormones*. New York: Academic Press, 1987, p. 167.

20.7— Cell Regulation and Hormone Secretion

Hormonal secretion is under specific control. In the cascade system displayed in Figures 20.2 and 20.3, hormones must emanate from one source, cause hormonal release from the next cell type in line, and so on, down the cascade system. The correct responses must follow from a specific stimulus. The precision of these signals is defined by the hormone and the receptor as well as by the activities of the CNS, which precedes the first hormonal response in many cases. Certain generalizations can be made. Polypeptide hormones generally bind to their cognate receptors located in cell membranes. The receptor recognizes structural features of the hormone to generate a high degree of specificity and affinity. The affinity constants for these interactions are in the range of 10^9–10^{11} M^{-1}, representing tight binding. This interaction usually activates or complexes with a transducing protein in the membrane, such as a **G-protein** (GTP-binding protein), or other transducer and causes an activation of some enzymatic function on the cytoplasmic side of the membrane. In some cases receptors undergo **internalization** to the cell interior; these receptors may or may not (e.g., the insulin receptor) be coupled to transducing proteins in the cell membrane. A discussion of internalization of receptors is presented in Section 20.11. The "activated" receptor complex could physically open a membrane ion channel or have other profound impacts on membrane structure and function. For example, binding of the hormone to the receptor may cause conformational changes in the receptor molecule, enabling it to associate with transducer in which further conformational changes may occur to permit interaction with an enzyme on the cytoplasmic side of the plasma membrane. This interaction may cause conformational changes in an enzyme so that its catalytic site becomes active.

G-Proteins Serve as Cellular Transducers of Hormone Signals

Most transducers of receptors in the plasma membrane are GTP-binding proteins and are referred to as **G-proteins.** G-Proteins consist of three types of **subunits—α, β, and γ**. The **α subunit** is the guanine nucleotide-binding component and is thought to interact with the receptor indirectly through the β and γ subunits and then directly with an enzyme, such as adenylate cyclase, resulting in enzyme activation. Actually there are two forms of the α subunit, designated $_s$ for a stimulatory subunit and $_i$ for an inhibitory subunit. Two types of receptors, and thus hormones, control the adenylate cyclase reaction: hormone–receptors that lead to a stimulation of the adenylate cyclase and those that lead to an inhibition of the cyclase. This is depicted in Figure 20.16 with an indication of the role of $_s$ and $_i$ and some of the hormones that interact with the stimulatory and inhibitory receptors.

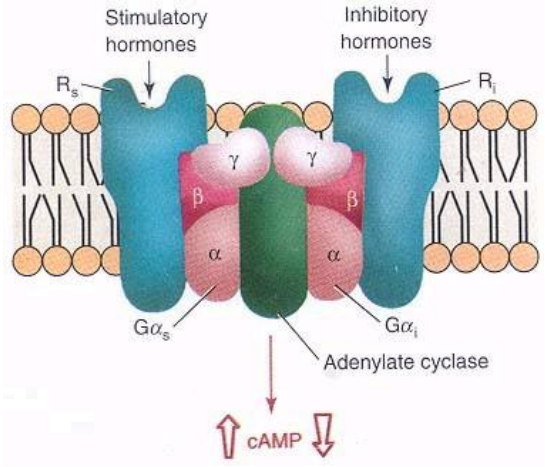


Figure 20.16
Components that constitute a hormone-sensitive adenylate cyclase system and the subunit composition.
Adenylate cyclase is responsible for conversion of ATP to cAMP. The occupancy of R_s by stimulatory hormones stimulates adenylate cyclase via formation of an active dissociated $G\alpha_s$ subunit. The occupancy of R_i by inhibitory hormones results in the formation of an "active" $G\alpha_i$ complex and concomitant reduction in cyclase activity. The fate of β and γ subunits in these dissociation reactions is not yet known. R_s, stimulatory hormone receptor; R_i, inhibitory hormone receptor.

The sequence of events that occurs when hormone and receptor interact is presented in Figure 20.17 and is as follows: receptor binds hormone in the membrane (Step 1); which produces a conformational change in receptor to expose a site for G-protein (β, γ subunit) attachment (Step 2); G-protein can be either stimulatory, G_s, or inhibitory, G_i, referring to the ultimate effects on the activity of adenylate cyclase; the receptor interacts with β, γ subunit of G-protein, enabling the α subunit to exchange GTP for bound GDP (Step 3); dissociation of GDP causes separation of G-protein α subunit from β, γ subunit and the α-binding site for interaction with adenylate cyclase appears on the surface of the G-protein α subunit (Step 4); α subunit binds to adenylate cyclase and activates the catalytic center, so that ATP is converted to cAMP (Step 5); GTP is hydrolyzed to GDP by the GTPase activity of the α subunit, returning it to its original conformation and allowing its interaction with β, γ subunit once again (Step 6); GDP associates with the α subunit and the system is returned to the unstimulated state awaiting another cycle of activity. It is important to note that there is also evidence suggesting that the β, γ complexes may play important roles in regulating certain effectors including adenylate cyclase.

In the case where an inhibitory G-protein is coupled to the receptor, the events are similar but inhibition of adenylate cyclase activity may arise by direct interaction of the inhibitory α subunit with adenylate cyclase or, alternatively, the inhibitory α subunit may interact directly with the stimulatory α subunit on the other side and prevent the stimulation of adenylate cyclase activity indirectly. Immunochemical evidence suggests multiple G_i subtypes and molecular cloning of complementary DNAs encoding putative α subunits has also provided evidence for multiple $_i$ subtypes.

Purification and biochemical characterization of G-proteins (G_s as well as G_i) have revealed somewhat unanticipated diversity in this subfamily. Polymerase chain reaction-based cloning has now brought the number of distinct genes encoding mammalian α subunits to at least 15. With regard to α subunits, further diversity is achieved by alternative splicing of the $_s$ (four forms) gene. There also appears to be diversity among the mammalian β and γ subunits. At least four distinct β subunit cDNAs and probably as many γ subunits have been described. What is not clear is how these complexes combine to form distinct β, γ complexes. Some data suggest that different β, γ complexes may have distinct properties with respect to α subunit and receptor interactions, but additional research will be required to fully describe these unique interactions.

Table 20.6 lists some activities transduced by G-protein subfamilies.

TABLE 20.6 Activities Transduced by G-Protein Subfamilies

α Subunit	***Expression***		***Effector***
G_s	Ubiquitous		Adenylate cyclase, Ca^{2+} channel
G_{olf}	Olfactory		Adenylate cyclase
G_{t-1} (transducin)	Rod photoreceptors		cGMP-phosphodiesterase
G_{t-2} (transducin)	Cone photoreceptors		cGMP-phosphodiesterase
G_{i1}	Neural > other tissues	}	Adenylate cyclase
G_{i2}	Ubiquitous		
G_{i3}	Other tissues > neural		
G_o	Neural, endocrine		Ca^{2+} channel
G_q	Ubiquitous	}	Phospholipase C
G_{11}	Ubiquitous		
G_{14}	Liver, lung, kidney		
$G_{15/16}$	Blood cells		

Source: Adapted from Spiegel, A. M., Shenker, A., and Weinstein, L. S. *Endocr. Rev.* 13:536, 1992.

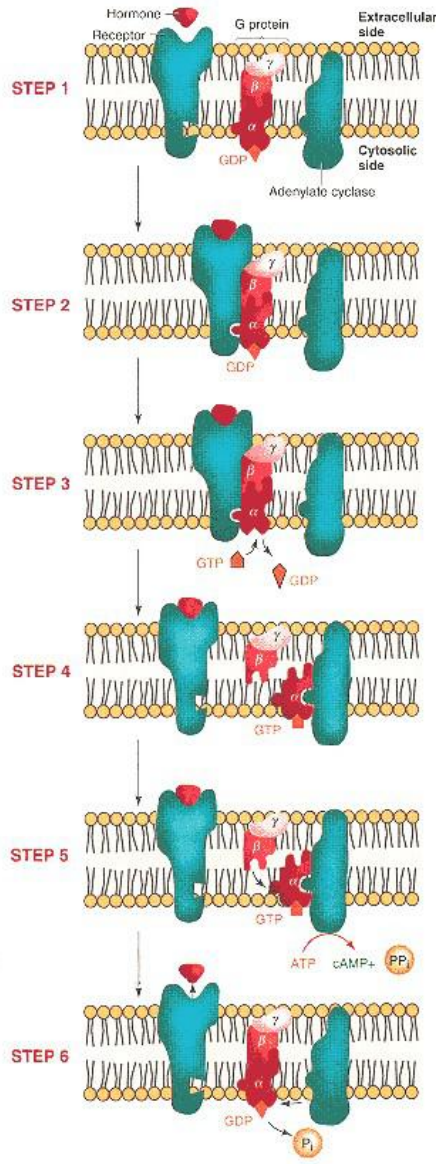


Figure 20.17
Activation of adenylate cyclase by binding of a hormone to its receptor.
The cell membrane is depicted, which contains on its outer surface a receptor protein for a hormone. On the inside surface of the membrane is adenylate cyclase protein and the transducer protein G. In the resting state GDP is bound to the α subunit of the G-protein. When a hormone binds to the receptor, a conformational change occurs (Step 1). The activated receptor binds to the G-protein (Step 2), which activates the latter so that it releases GDP and binds GTP (Step 3), causing the α and the complex of β and γ-subunits to dissociate (Step 4). Free G_α subunit binds to the adenylate cyclase and activates it so that it catalyzes the synthesis of cAMP from ATP (Step 5); this step may involve a conformational change in G_α. In some cases the β,γ complex may play an important role in regulation of certain effectors including adenylate cyclase. When GTP is hydrolyzed to GDP, a reaction most likely catalyzed by G_α itself, G_α is no longer able to activate adenylate cyclase (Step 6), and G_α and $G_{\beta\gamma}$ reassociate. The hormone dissociates from the receptor and the system returns to its resting state.
Redrawn from Darnell, J., Lodish, H., and Baltimore, D. *Molecular Cell Biology*. New York: Scientific American Books, 1986, p. 682.

Cyclic AMP Activates Protein Kinase A Pathway

The generation of cAMP in the cell usually activates protein kinase A, referred to as the **protein kinase A pathway.** The overall pathway is presented in Figure 20.18. Four cAMP molecules are used in the reaction to complex two regulatory subunits (R) and liberating two protein kinase catalytic subunits (C). The liberated catalytic subunits are able to phosphorylate proteins to produce a cellular effect. In many cases the cellular effect leads to the release of preformed hormones. For example, ACTH binds to membrane receptors, elevates intracellular **cAMP** levels, and releases cortisol from the *zona fasciculata* cells of the adrenal gland by this general mechanism. Part of the mechanism of release of thyroid hormones from the thyroid gland involves the cAMP pathway as outlined in Figure 20.19. TSH has been shown to stimulate numerous key steps in this secretory process, including iodide uptake and endocytosis of thyroglobulin (Figure 20.14). The protein kinase A pathway is also responsible for the release of testosterone by testicular Leydig cells as presented in Figure 20.20. There are many other examples of hormonal actions mediated by cAMP and the protein kinase A pathway.

Inositol Triphosphate Formation Leads to Release of Calcium from Intracellular Stores

Uptake of calcium from the cell exterior through calcium channels may be affected directly by hormone-receptor interaction at the cell membrane. In some cases, ligand-receptor interaction is thought to open calcium channels directly in the cell membrane (Chapter 5, Section 5.5). Another system to increase intracellular Ca^{2+} concentration derives from hormone-receptor activation of **phospholipase C** activity transduced by a G-protein (Figure 20.21).

A hormone operating through this system binds to a specific cell membrane receptor, which interacts with a G-protein in a mechanism similar to that of the protein kinase A pathway and transduces the signal, resulting in stimulation of phospholipase C. This enzyme catalyzes the hydrolysis of **phosphatidylinositol-4,5-bisphosphate** (PIP_2) to form two **second messengers, diacylglycerol (DAG)** and **inositol 1,4,5-triphosphate (IP_3).**

Inositol 1,4,5-triphosphate diffuses to the cytosol and binds to an IP_3 receptor on the membrane of a particulate **calcium** store, either separate from or

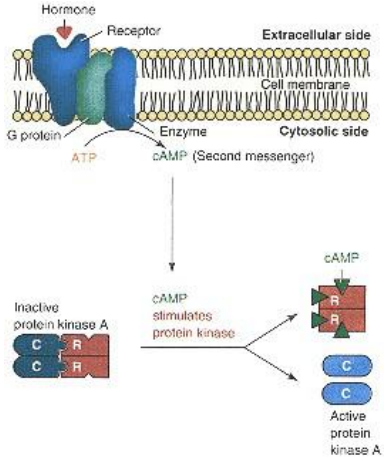


Figure 20.18
Activation of protein kinase A.
Hormone–receptor mediated stimulation of adenylate cyclase and subsequent activation of protein kinase A. C, catalytic subunit; R, regulatory subunit.

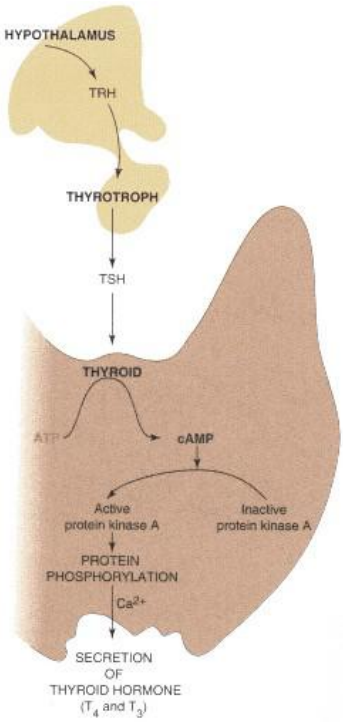


Figure 20.19
Overview of secretion controls of thyroid hormone.

CLINICAL CORRELATION 20.3

Lithium Treatment of Manic–Depressive Illness: The Phosphatidylinositol Cycle

Lithium has been used for years in the treatment of manic depression. Our newer knowledge suggests that lithium therapy involves the phosphatidylinositol (PI) pathway. This pathway generates the second messengers inositol 1,4,5-triphosphate (IP_3) and diacylglycerol following the hormone/neurotransmitter–membrane receptor interaction and involves the G-protein complex and activation of phospholipase C. IP_3 and its many phosphorylated derivatives are ultimately dephosphorylated in a stepwise fashion to generate free inositol. Inositol is then used for the synthesis of phosphatidylinositol monophosphate. The phosphatase that dephosphorylates IP to inositol is inhibited by Li^+. In addition, Li^+ may also interfere directly with G-protein function. The result of Li^+ inhibition is that the PI cycle is greatly slowed even in the face of continued hormonal/neurotransmitter stimulation and the cell becomes less sensitive to these stimuli. Manic–depressive illness may occur through the overactivity of certain CNS cells, perhaps as a result of abnormally high levels of hormones or neurotransmitters whose actions are to stimulate the PI cycle. The chemotherapeutic effect of the Li^+ could be to decrease the cellular responsiveness to elevated levels of agents that might promote high levels of PI cycle and precipitate manic-depressive illness.

Avissar, S., and Schreiber, G. Muscarinic receptor subclassification and G-proteins: significance for lithium action in affective disorders and for the treatment of the extrapyramidal side effects of neuroleptics. *Biol. Psychiatry* 26:113, 1989; Hallcher, L. M., and Sherman, W. R The effects of lithium ion and other agents on the activity of myoinositol 1-phosphatase from bovine brain. *J. Biol. Chem.* 255:896, 1980; and Pollack, S. J., Atack, J. R., Knowles, M. R., McAllister, G., Ragan, C. I., Baker, R., Fletcher, S. R., Iversen, L. L., and Broughton, H. B. Mechanism of inositol monophosphatase, the putative target of lithium therapy. *Proc. Natl. Acad. Sci. USA* 91:5766, 1994.

part of the endoplasmic reticulum. IP_3 binding results in the release of calcium ions contributing to the large increase in cytosolic Ca^{2+} levels. Calcium ions may be important to the process of exocytosis by taking part in the fusion of secretory granules to the internal cell membrane, in microtubular aggregation or in the function of contractile proteins, which may be part of the structure of the exocytotic mechanism, or all of these.

The IP_3 is metabolized by stepwise removal of phosphate groups (Figure 20.21) to form inositol. This combines with phosphatidic acid (PA) to form phosphatidylinositol (PI) in the cell membrane. PI is phosphorylated twice by a kinase to form PIP_2, which is ready to undergo another round of hydrolysis and formation of second messengers (DAG and IP_3) upon hormonal stimulation. If the receptor is still occupied by hormone, several rounds of the cycle could occur before the hormone–receptor complex dissociates or some other feature of the cycle becomes limiting. It is interesting that the conversion of inositol phosphate to inositol is inhibited by **lithium ion** (Li^+) (Figure 20.21). This could be the metabolic basis for the beneficial effects of Li^+ in manic-depressive illness (see Clin. Corr. 20.3). Finally, it is important to note that not all of the generated IP_3 is dephosphorylated during hormonal stimulation. Some of the IP_3 is phosphorylated via IP_3 kinase to yield inositol 1,3,4,5-tetraphosphate (IP_4), which may mediate some of the slower or more prolonged hormonal responses or facilitate replenishment of intracellular Ca^{2+} stores from the extracellular fluid, or both.

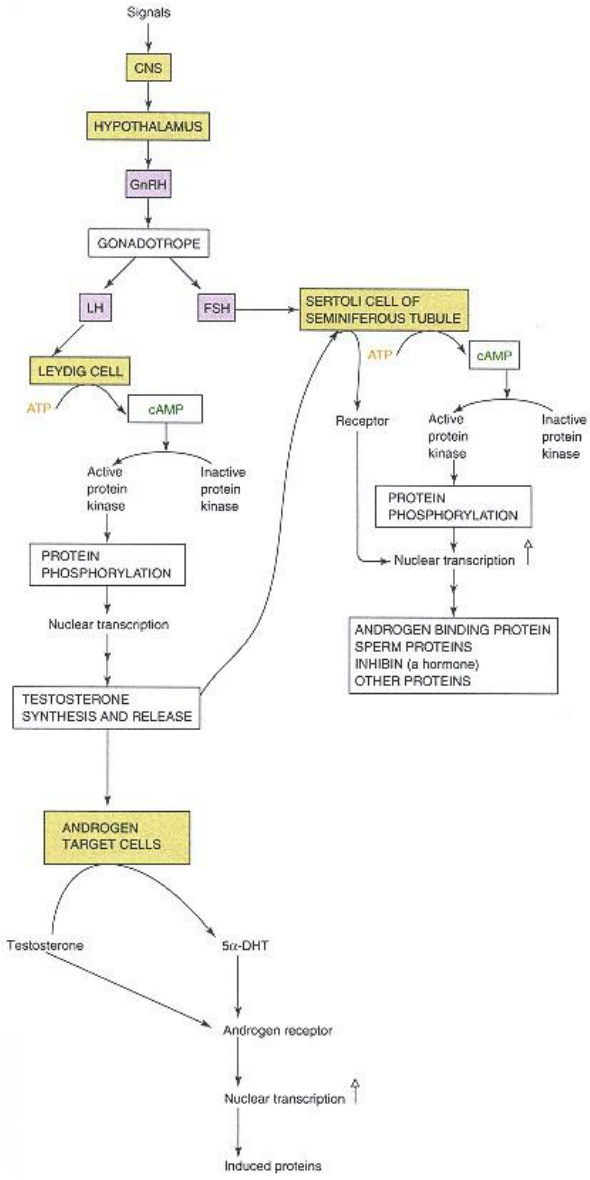


Figure 20.20
Overview of the secretion controls and some general actions of the gonadotropes and testosterone release in males.
In some, but not all, androgen target cells, testosterone is reduced to the more potent androgen, 5α- di-hydrotestosterone (5α-DHT).

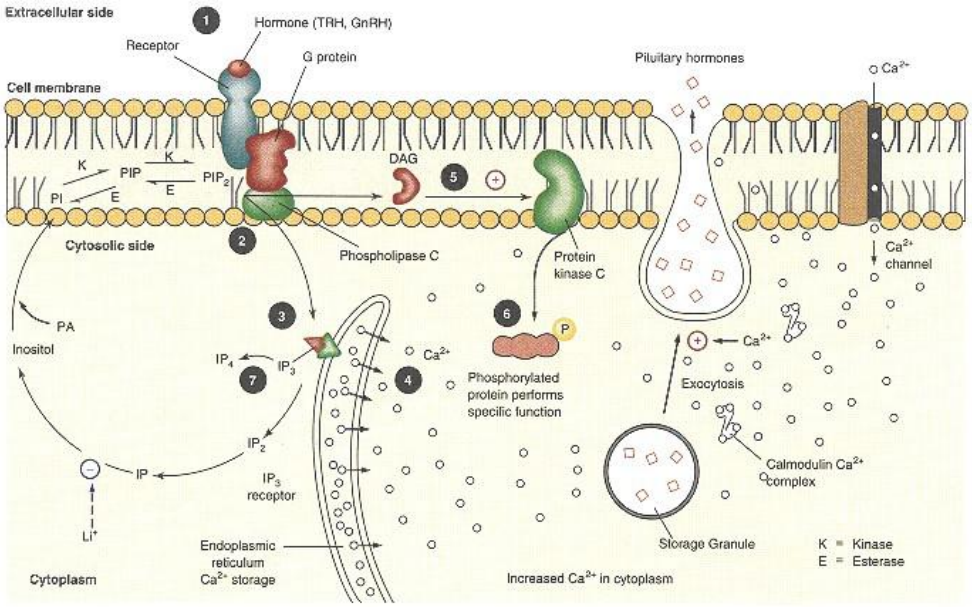


Figure 20.21
Overview of hormonal signaling through the phosphatidylinositol system generating the second messengers, inositol 1,4,5-trisphosphate (IP_3) and diacylglycerol (DAG).

The action of IP_3 is to increase cytosol Ca^{2+} levels by a receptor-mediated event in the cellular calcium store. Steps in pathway: (1) binding of hormone to cell membrane receptor; (2) production of IP_3 from PIP_2; (3) binding of IP_3 to receptor on calcium storage site; (4) release of free calcium to the cytosol; (5) release of DAG and subsequent binding to protein kinase C; (6) phosphorylation of protein substrates by protein kinase C activated by DAG and Ca^{2+}; and (7) phosphorylation of IP_3 to yield IP_4. DAG, diacylglycerol; PA, phosphatidic acid; IP, inositol phosphate; IP_2, inositol bisphosphate; IP_3, inositol 1,4,5-triphosphate; IP_4, inositol 1,3,4,5-tetrakisphosphate; PIP, phosphatidylinositol phosphate; PIP_2, phosphatidylinositol 4,5-bisphosphate; K, kinase; E, esterase.

Diacylglycerol Activates Protein Kinase C Pathway

At the same time that the IP_3 produced by hydrolysis of PIP_2 is increasing the concentration of Ca^{2+} in the cytosol, the other cleavage product, DAG, mediates different effects. Importantly, DAG activates a crucial serine/threonine protein kinase called **protein kinase C** because it is Ca^{2+} dependent (details of protein kinase C discussed on p. 883). The initial rise in cytosolic Ca^{2+} induced by IP_3 is believed to somehow alter kinase C so that it translocates from the cytosol to the cytoplasmic face of the plasma membrane. Once translocated, it is activated by a combination of Ca^{2+}, DAG, and the negatively charged membrane phospholipid, phosphatidylserine. Once activated, protein kinase C then phosphorylates specific proteins in the cytosol or, in some cases, in the plasma membrane. These phosphorylated proteins perform specific functions that they could not mediate in their nonphosphorylated states. For example, a phosphorylated protein could potentially migrate to the nucleus and stimulate mitosis and

growth. It is also possible that a phosphorylated protein could play a role in the secretion of preformed hormones.

20.8— Cyclic Hormonal Cascade Systems

Hormonal cascade systems can be generated by external signals as well as by internal signals. Examples of this are the **diurnal variations** in levels of cortisol secreted from the adrenal gland probably initiated by serotonin and vasopressin, the day and night variations in the secretion of **melatonin** from the pineal gland and the internal regulation of the **ovarian cycle.** Some of these biorhythms operate on a cyclic basis, often dictated by daylight and darkness, and are referred to as **chronotropic control** of hormone secretion.

Melatonin and Serotonin Synthesis Are Controlled by Light and Dark Cycles

The release of melatonin from the pineal gland, presented in overview in Figure 20.22, is an example of a biorhythm. Here, as in other such systems, the internal signal is provided by a neurotransmitter, in this case norepinephrine produced by an adrenergic neuron. In this system, control is exerted by light entering the eyes and is transmitted to the pineal gland by way of the CNS. The adrenergic neuron innervating the pinealocyte is inhibited by light transmitted through the eyes. Norepinephrine released as a neurotransmitter in the dark stimulates cAMP formation through a β receptor in the pinealocyte cell membrane, which leads to the enhanced synthesis of ***N*-acetyltransferase.** The increased activity of this enzyme causes the conversion of **serotonin** to ***N*-acetylserotonin,** and **hydroxyindole-*O*-methyltransferase (HIOMT)** then catalyzes the conversion of *N*-acetylserotonin to **melatonin,** which is secreted in the dark hours but not during light hours. Melatonin is circulated to cells containing receptors that generate effects on reproductive and other functions. For example, melatonin has been shown to exert an antigonadotropic effect, although the physiological significance of this effect is unclear.

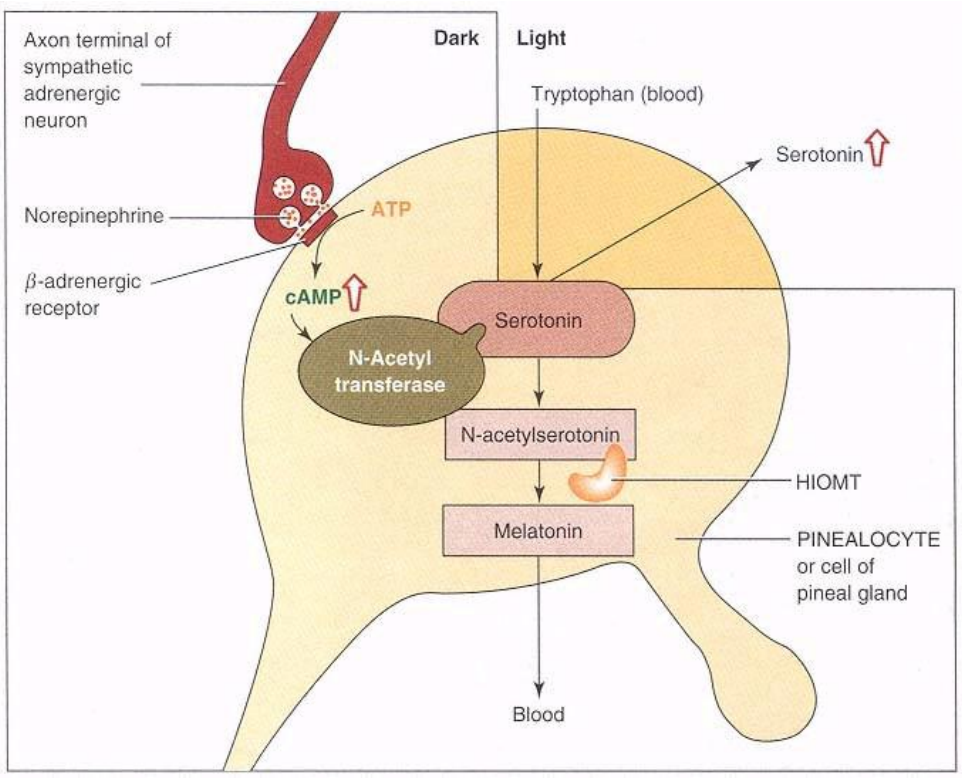


Figure 20.22
Biosynthesis of melatonin in pinealocytes.
HIOMT, hydroxyindole-O-methyltransferase.
Redrawn from Norman, A. W., and Litwack, G. *Hormones.* New York: Academic Press, 1987, p. 710.

Ovarian Cycle Is Controlled by Gonadotropin-Releasing Hormone

An example of a pulsatile release mechanism is regulation of the periodic release of GnRH. A periodic control regulates the release of this substance at definitive periods (of about 1 h in higher animals) and is controlled by aminergic neurons, which may be adrenergic (norepinephrine secreting) in nature. The initiation of this function occurs at puberty and is important in both the male and female. While the male system functions continually, the female system is periodic and known as the **ovarian cycle.** This system is presented in Figure 20.23. In the

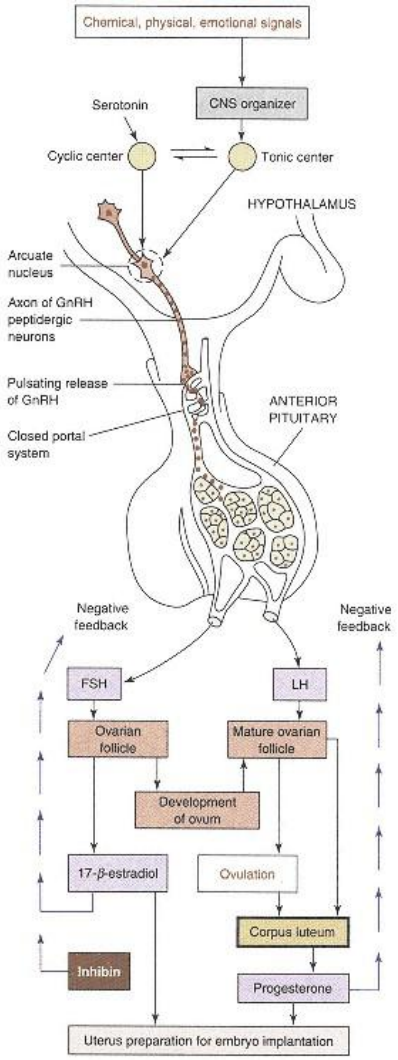


Figure 20.23
Ovarian cycle in terms of generation of hypothalamic hormone, pituitary gonadotropic hormones, and sex hormones.
To begin the cycle at puberty, several centers in the CNS coordinate with the hypothalamus so that hypothalamic GnRH can be released in a pulsatile fashion. This causes the release of the gonadotropic hormones, LH and FSH, which in turn affect the ovarian follicle, ovulation, and the corpus luteum. The hormone inhibin selectively inhibits FSH secretion. Products of the follicle and corpus luteum, respectively, are β-estradiol and progesterone. GnRH, gonadotropin-releasing hormone; FSH, follicle-stimulating hormone; LH, luteinizing hormone.

male, the cycling center in the CNS does not develop because its development is blocked by androgens before birth.

In the female, a complicated set of signals needs to be organized in the CNS before the initial secretion of GnRH occurs at puberty. The higher centers (CNS organizer) must harmonize with the tonic and cycling centers and these interact with each other to prime the hypothalamus. The pulsatile system, which innervates the arcuate nucleus of the hypothalamus, must also function for GnRH to be released, and this system apparently must be functional throughout life for these cycles to be maintained. Release of GnRH from the axon terminals of the cells that synthesize this hormone is followed by entry of the hormone into the primary plexus of the closed portal system connecting the hypothalamus and the anterior pituitary (Figure 20.23). The blood–brain barrier preventing peptide transport is overcome in this process by allowing GnRH to enter the vascular system through fenestrations, or openings in the blood vessels, that permit such transport. The GnRH is then carried down the **portal system** and leaves the secondary plexus through fenestrations, again, in the region of the target cells **(gonadotropes)** of the anterior pituitary. The hormone binds to its cognate membrane receptor and the signal, mediated by the phosphatidylinositol metabolic system, causes the release of both FSH and LH from the same cell. The **FSH** binds to its cognate membrane receptor on the ovarian follicle and, operating through the protein kinase A pathway via cAMP elevation, stimulates synthesis and secretion of 17β-estradiol, the female sex hormone, and maturation of the follicle and ovum. Other proteins, such as **inhibin,** are also synthesized. Inhibin is a negative feedback regulator of FSH production in the gonadotrope. When the follicle reaches full maturation and the ovum also is matured, LH binds to its cognate receptor and plays a role in ovulation together with other factors, such as prostaglandin $F_{2\alpha}$. The residual follicle remaining after ovulation becomes the functional *corpus luteum* under primary control of LH (Figure 20.23). The **LH** binds to its cognate receptor in the *corpus luteum* cell membrane and, through stimulation of the protein kinase A pathway, stimulates synthesis of progesterone, the progestational hormone. **Estradiol** and **progesterone** bind to intracellular receptors (Chapter 21) in the uterine endometrium and cause major changes resulting in the thickening of the wall and vascularization in preparation for implantation of the fertilized egg. Estradiol, which is synthesized in large amount prior to production of progesterone, induces the progesterone receptor as one of its inducible phenotypes. This induction of progesterone receptors primes the uterus for subsequent stimulation by progesterone secreted by the *corpus luteum*.

Absence of Fertilization

If fertilization of the ovum does not occur, the *corpus luteum* involutes as a consequence of diminished LH supply. Progesterone levels fall sharply in the blood with the regression of the *corpus luteum*. Estradiol levels also fall due to the cessation of its production by the *corpus luteum*. Thus the stimuli for a thickened and vascularized uterine endometrial wall are lost. *Menstruation* occurs through a process of programmed cell death of the uterine endometrial cells until the endometrium reaches its unstimulated state. Ultimately, the fall in blood steroid levels releases the negative feedback inhibition on the gonadotropes and hypothalamus and the cycle starts again with release of FSH and LH by the gonadotropes in response to GnRH.

The course of the ovarian cycle is shown in Figure 20.24 with respect to the relative blood levels of hormones released from the hypothalamus, anterior pituitary, ovarian follicle, and corpus luteum. In addition, changes in the maturation of the follicle and ovum as well as the uterine endometrium are shown. Aspects of the steroid hormones, estradiol and progesterone, are discussed in Chapter 21.

The cycle first begins at puberty when GnRH is released, corresponding

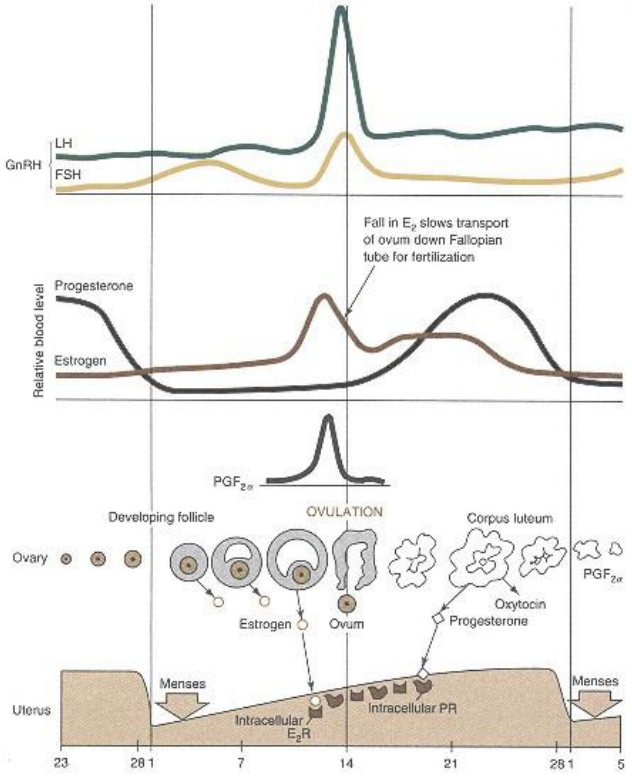


Figure 20.24
The ovarian cycle.
In the upper diagram, relative blood levels of GnRH, LH, FSH, progesterone, estrogen, and $PGF_{2\alpha}$ are shown. In the lower diagram, events in the ovarian follicle, corpus luteum, and uterine endometrium are diagrammed. GnRH, gonadotropin-releasing hormone; LH, luteinizing hormone; FSH, follicle-stimulating hormone; $PGF_{2\alpha}$, prostaglandin $F_{2\alpha}$; E_2, estradiol; E_2R, intracellular estrogen receptor; PR, intracellular progesterone receptor.

to day 1 in Figure 20.24. GnRH is released in a pulsatile fashion, causing the gonadotrope to release FSH and LH; there is a rise in the blood levels of these gonadotropic hormones in subsequent days. Under the stimulation of FSH the follicle begins to mature (lower section of Figure 20.24) and estradiol (E_2) is produced. In response to estradiol the uterine endometrium begins to thicken (there would have been no prior menstruation in the very first cycle). Eventually, under the continued action of FSH, the follicle matures with the maturing ovum, and extraordinarily high levels of estradiol are produced (around day 13 of the cycle). These levels of estradiol, instead of causing feedback inhibition, now generate, through **feedback stimulation,** a huge release of LH and to a lesser extent FSH from the gonadotrope. The FSH responds to a smaller extent due to the ovarian production of the hormone inhibin under the influence of FSH. Inhibin is a specific negative feedback inhibitor of FSH, but not of LH, and probably suppresses the synthesis of the β subunit of FSH. The high midcycle peak of LH is referred to as the "LH spike." Ovulation then occurs at about day 14 (midcycle) through the effects of high LH concentration together with other factors, such as $PGF_{2\alpha}$. Both LH and $PGF_{2\alpha}$ act on cell membrane receptors. After ovulation, the function of the follicle declines as reflected by the fall in blood estrogen levels. The spent follicle now differentiates into the functional corpus luteum driven by the still high levels of blood LH (Figure 20.23, top).

Under the influence of prior high levels of estradiol (estrogen) and the high levels of progesterone produced by the now functional *corpus luteum*, the uterine endometrial wall reaches its greatest development in preparation for implantation of the fertilized egg, should fertilization occur. Note that the previous availability of estradiol in combination with the estrogen receptor (E_2R) produces elevated levels of progesterone receptor (PR) within the cells of the uterine wall. The blood levels of estrogen fall with the loss of function of the follicle but some estrogen is produced by the *corpus luteum* in addition to the much greater levels of progesterone. In the absence of fertilization the *corpus luteum* continues to function for about 2 weeks, then involutes because of the loss of high levels of LH. The production of oxytocin by the *corpus luteum* itself and the production or availability of $PGF_{2\alpha}$ cause inhibition of progesterone synthesis and enhances luteolysis by a process of programmed cell death (Chapter 21). With the death of the *corpus luteum* there is a profound decline in blood levels of estradiol and progesterone so that the thickened endometrial wall can no longer be maintained and menstruation occurs, followed by the start of another cycle with a new developing follicle.

Fertilization

The situation changes if fertilization occurs as shown in Figure 20.25. The *corpus luteum*, which would have ceased function by 28 days, remains viable due to the production of **chorionic gonadotropin,** which resembles and acts like LH, from the trophoblast. Eventually, the production of **human chorionic gonadotropin** (hCG) is taken over by the placenta, which continues to produce the hormone at very high levels throughout most of the gestational period. Nevertheless, the *corpus luteum*, referred to as the "*corpus luteum* of pregnancy," eventually dies and, by about 12 weeks of pregnancy, the placenta has taken over the production of progesterone, which is secreted at high levels throughout pregnancy. Although both progesterone and estrogen are secreted in progressively greater quantities throughout pregnancy, from the seventh month onward estrogen secretion continues to increase while progesterone secretion remains constant or may even decrease slightly (Figure 20.25). The increased production of a progesterone-binding protein may also serve to lower the effective concentration of free progesterone in the myometrium. Thus the estrogen/progesterone ratio increases toward the end of pregnancy and may

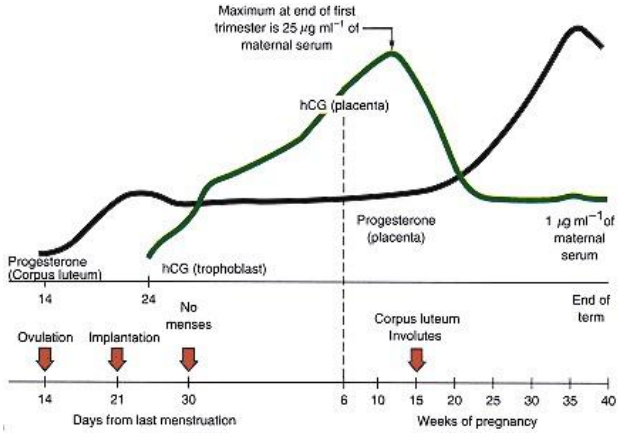


Figure 20.25
Effect of fertilization on ovarian cycle in terms of progesterone and secretion of human chorionic gonadotropin (hCG).

be partly responsible for the increased uterine contractions. Oxytocin secreted by the posterior pituitary also contributes to these uterine contractions. The fetal membranes also release prostaglandins ($PGF_{2\alpha}$) at the time of parturition and they also increase the intensity of uterine contractions. Finally, the fetal adrenal glands secrete cortisol, which not only stimulates fetal lung maturation by inducing surfactant but may also stimulate uterine contractions.

As mentioned before, the system in the male is similar, but less complex in that cycling is not involved, and it progresses much as outlined in Figure 20.25. This is only one example of biorhythmic and pulsatile systems.

20.9— Hormone–Receptor Interactions

Receptors are proteins and differ by their specificity for ligands and by their location in the cell (see Figure 20.1). The interaction of ligand with receptor essentially resembles a semienzymatic reaction:

$$\text{Hormone} + \text{receptor} \rightleftharpoons \text{hormone–receptor complex}$$

The **hormone–receptor complex** usually undergoes conformational changes resulting from interaction with the hormonal ligand. These changes allow for a subsequent interaction with a transducing protein (G-protein) in the membrane or for activation to a new state in which active domains become available on the surface of the receptor. The mathematical treatment of the interaction of hormone and receptor is a function of the concentrations of the reactants, hormone [H] and receptor [R], in the formation of the hormone–receptor complex [RH], and the rates of formation and reversal of the reaction:

$$[H] + [R] \underset{k-1}{\overset{k+1}{\rightleftharpoons}} [RH]$$

The reaction can be studied under conditions, such as low temperature, that will further reduce reactions involving the hormone–receptor complex. The equilibrium can thus be expressed in terms of the association constant, K_a, which is equal to the inverse of the dissociation constant, K_d:

$$K_a = \frac{[RH]}{[H][R]} = \frac{k_{+1}}{k_{-1}} = \frac{1}{K_d}$$

The concentrations are equilibrium concentrations that can be restated in terms of the forward and reverse velocity constants, k_{+1} being the on-rate and k_{-1} being the off-rate (**on** refers to hormone association with the receptor and **off** refers to hormone dissociation). Experimentally, equilibrium under given conditions is determined by a progress curve of binding that reaches saturation. A saturating amount of hormone is determined using variable amounts of free hormone and measuring the amount bound with some convenient assay. The half-maximal value of a plot of receptor-bound hormone (ordinate) versus total free-hormone concentration (abscissa) approximates the dissociation constant, which will have a specific hormone concentration in molarity as its value. Hormone bound to receptor is corrected for nonspecific binding of the hormone to the membrane or other nonreceptor intracellular proteins. This can be measured conveniently if the hormone is radiolabelled, by measuring receptor binding using labeled hormone ("hot" or "uncompeted") and receptor binding using labeled hormone after the addition of an excess (100–1000 times) of unlabeled hormone ("hot" + "cold" or competed). The excess of unlabeled hormone will displace the high-affinity hormone-binding sites but not the low-affinity nonspecific binding sites. Thus when the "competed" curve is subtracted from the "uncompeted" curve, as seen in Figure 20.26, an intermediate curve will represent specific binding of labeled hormone to receptor. This is of critical

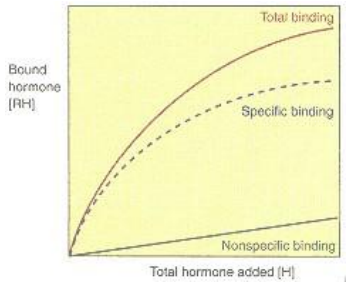


Figure 20.26
Typical plot showing specific hormone binding.

importance when receptor is measured in a system containing other proteins. As an approximation, 20 times the K_d value of hormone is usually enough to saturate the receptor.

Scatchard Analysis Permits Determination of the Number of Receptor-Binding Sites and Association Constant for Ligand

Most measurements of K_d are made using **Scatchard analysis,** which is a manipulation of the equilibrium equation. The equation can be developed by a number of routes but can be envisioned from mass action analysis of the equation presented above. At equilibrium the total possible number of binding sites (B_{max}) equals the unbound plus the bound sites, so that B_{max} = R + RH, and the unbound sites (R) will be equal to R = B_{max} – RH. To consider the sites left unbound in the reaction the equilibrium equation becomes

$$K_a = \frac{[RH]}{[H]([B_{max}] - [RH])}$$

Thus

$$\frac{\text{bound}}{\text{free}} = \frac{[RH]}{[H]} = K_a([B_{max}] - ([RH]) = \frac{1}{K_d}([B_{max}] - [RH])$$

The Scatchard plot of bound/free = [RH]/[H] on the ordinate versus bound on the abscissa yields a straight line, as shown in Figure 20.27. When the line is extrapolated to the abscissa, the intercept gives the value of B_{max} (the total number of specific receptor-binding sites). The slope of the negative straight line is $-K_a$ or $-1/K_d$.

These analyses are sufficient for most systems but become more complex when there are two components in the Scatchard plot. In this case the straight line usually curves as it approaches the abscissa and a second phase is observed somewhat asymptotic to the abscissa while still retaining a negative slope (Figure 20.28*a*). In order to obtain the true value of K_d for the steeper, higher-affinity sites, the low-affinity curve must be subtracted from the first set, which also corrects the extrapolated value of B_{max}. From these analyses information is obtained on K_d, the number of classes of binding sites (usually one or two), and the maximal number of high-affinity receptor sites (receptor number) in the system (see Figure 20.28*b*). These curvilinear Scatchard plots can result not only from the existence of more than one distinct binding component but also as a consequence of what is referred to as **negative cooperativity.** This term refers to the fact that in some systems the affinity of the receptor for its ligand is gradually decreased as more and more ligand binds. From application to a wide variety of systems it appears that K_d values for many hormone receptors range from 10^{-9} to 10^{-11} M, indicating very tight binding. These interactions are generally marked by a high degree of specificity so that both parameters describe interactions of a high order, indicating the uniqueness of receptors and the selectivity of signal reception.

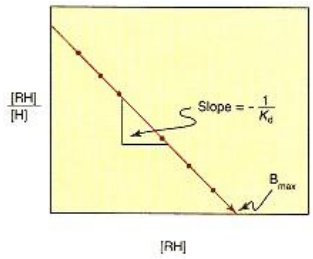


Figure 20.27
Typical plot of Scatchard analysis of specific binding of ligand to receptor.

Some Hormone–Receptor Interactions Involve Multiple Hormone Subunits

Interaction of hormone and receptor can be exemplified by the anterior pituitary hormones, **thyrotropin (TSH), luteinizing hormone (LH),** and **follicle-stimulating hormone (FSH).** These hormones each contain two subunits, an α and a β subunit. The α subunit for all three hormones is nearly identical and the α subunit of any of the three can substitute for the other two. Consequently, the α subunit performs some function in common to all three hormones in their interaction with receptor but is obviously not responsible for the specificity

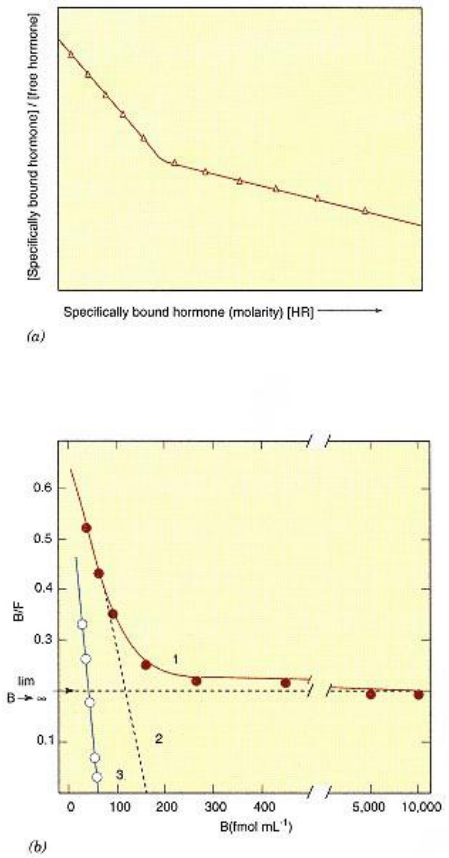


Figure 20.28
Scatchard analysis of curves representing two components.
(a) Scatchard curve showing two components.
(b) Scatchard plot with correction of high-affinity component by subtraction of nonspecific binding attributable to the low-affinity component. Curve 1: total binding. Curve 2: Linear extrapolation of high-affinity component that includes contribution from low-affinity component. Curve 3: Specific binding of high-affinity component after removal of nonspecific component.
Redrawn from Chamness, G. C., and McGuire, W. L. *Steroids* 26:538, 1975.

required for each cognate receptor. The hormones cannot replace each other in binding to their specific receptor. Thus the specificity of receptor recognition is imparted by the β subunit, whose structure is unique for the three hormones.

On the basis of topological studies with monoclonal antibodies, a picture of the interaction of LH with its receptor has been suggested as shown in Figure 20.29. In this model, the receptor recognizes both subunits of the hormonal ligand, but the β subunit is specifically recognized by the receptor to lead to a response. With the TSH–receptor complex there may be more than one second messenger generated. In addition to the stimulation of adenylate cyclase and the increased intracellular level of cAMP, the phosphatidylinositol pathway (Figure 20.21) is also turned on. The preferred model is one in which there is a single receptor whose interaction with hormone activates both the adenylate cyclase and the phospholipid second messenger systems, as shown in Figure 20.30. Thus a variety of reactions could follow the hormone–receptor interaction through the subsequent stimulation of cAMP levels (protein kinase A pathway) and stimulation of phosphatidylinositol turnover (protein kinase C pathway).

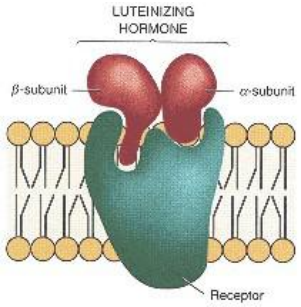


Figure 20.29
Interaction of the α and β subunits of LH with the LH receptor of rat Leydig cells.
The interaction was determined by topological analysis with monoclonal antibodies directed against epitopes on the α and β subunits of the hormone. Both α and β subunits participate in LH receptor binding. Adapted from Alonoso-Whipple, C., Couet, M. L., Doss, R., Koziarz, J., Ogunro, E. A., and Crowley, W. E. Jr. *Endocrinology* 123:1854, 1988.

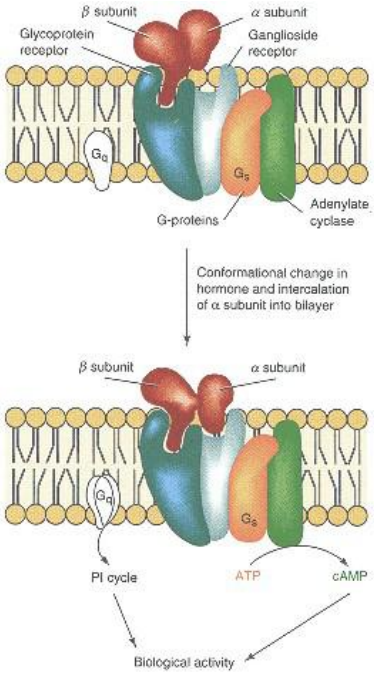


Figure 20.30
Model of TSH receptor, which is composed of glycoprotein and ganglioside component.
After the TSH β subunit interacts with receptor, the hormone changes its conformation and the α subunit is brought into the bilayer, where it interacts with other membrane components. The β subunit of TSH may carry primary determinants recognized by the glycoprotein receptor component. It is suggested that the TSH signal to adenylate cyclase is via the ganglioside; the glycoprotein component appears more directly linked to phospholipid signal system. PI, phosphatidylinositol; G_s, G-protein linked to activation of adenylate cyclase; G_q G-protein linked to PI cycle. Adapted with modifications from L. D. Kohn, et al. *Biochemical Actions of Hormones*, 12. G. Litwack (Ed.). Academic Press, 1985, p. 466.

20.10— Structure of Receptors: β-Adrenergic Receptor

Structures of receptors are conveniently discussed in terms of functional domains. Consequently, for membrane receptors there will be functional **ligand-binding domains** and the **transmembrane domains,** which for many membrane receptors involve protein kinase activities. In addition, specific **immuno-logical domains** contain primary epitopes of antigenic regions. Several membrane receptors have been cloned and studied with regard to structure and function, including the β receptors ($_1$ and $_2$), which recognize catecholamines, principally norepinephrine, and stimulate adenylate cyclase. The $_1$ and $_2$ receptors are subtypes that differ in affinities for norepinephrine and for synthetic antagonists. Thus $_1$-adrenergic receptor binds norepinephrine with a higher affinity than epinephrine, whereas the order of affinities is reversed for the $_2$-adrenergic receptor. The drug isoproterenol has a greater affinity for both receptors than the two hormones. In Figure 20.31 the amino acid sequence is shown (with single letter abbreviations for amino acids; see Table 20.4 for list) for the $_2$-adrenergic receptor. A polypeptide stretch extending from α helix I extends to the extracellular space. There are seven membrane-spanning domains and these appear also in the $_1$ receptor, where there is extensive homology with the $_2$ receptor. Cytosolic peptide regions extend to form loops between I and II, III and IV, and V and VI and an extended chain from VII.

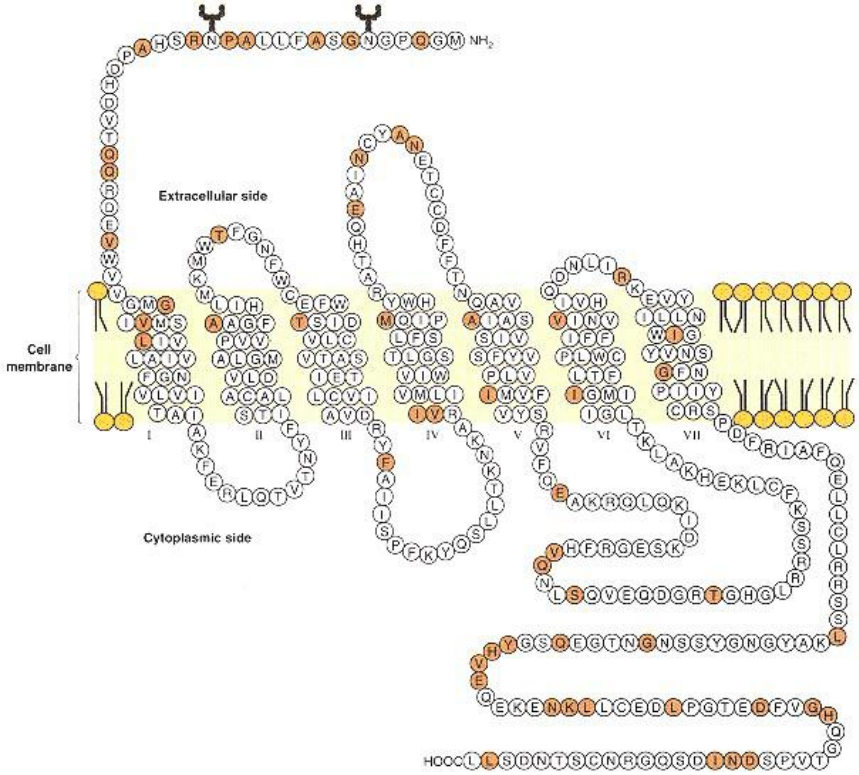


Figure 20.31

Proposed model for insertion of the β_2-adrenergic receptor (AR) in the cell membrane.

The model is based on hydropathicity analysis of the human $_2$-AR. Standard one-letter codes for amino acid residues are used. Hydrophobic domains are represented as transmembrane helices. Pink circles with black letters indicate residues in the human sequence that differ from those in hamster. Also noted are the potential sites of N-linked glycosylation.

Redrawn from Kobilka, B. K., Dixon, R. A., Frielle, T., Dohlman, H. G., et al. *Proc. Natl. Acad. Sci. USA* 84:46, 1987.

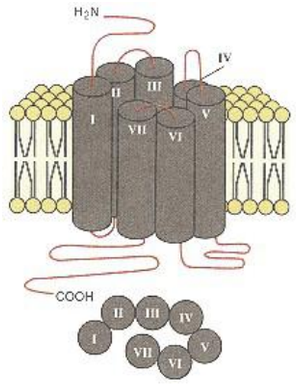


Figure 20.32
Proposed arrangement of β-adrenergic receptor helices in the membrane.
Lower portion of the figure is a view from above the plane of the plasma membrane. It is proposed that helices IV, VI, and VII are arranged in the membrane in such a way as to delineate a ligand-binding pocket, with helix VII centrally located.
Adapted from Frielle, T., Daniel, K. W., Caron, M. G., and Lefkowitz, R. J. *Proc. Natl. Acad. Sci. USA* 85:9494, 1988.

The long extended chain from VII may contain phosphorylation sites (serine and threonine residues) of the receptor, which are important in terms of the receptor regulatory process involving receptor desensitization. Phosphorylation of these residues within the cytoplasmic tail of the receptor results in the binding of an inhibitory protein, called β arrestin, which blocks the receptor's ability to activate G_s. Cell exterior peptide loops extend from II to III, IV to V, and VI to VII, but mutational analysis suggests that the external loops do not take part in ligand binding. It appears that ligand binding may occur in a pocket arranged by the location of the membrane-spanning cylinders I–VII, which for the $_1$ receptor appear to form a ligand pocket, as shown from a top view in Figure 20.32. Recently reported work suggests that transmembrane domain VI may play a role in the stimulation of adenylate cyclase activity. By substitution of a specific cysteine residue in this transmembrane domain, a mutant was generated that displays normal ligand-binding properties but a decreased ability to stimulate the cyclase.

20.11— Internalization of Receptors

Up to now we have described receptor systems that transduce signals through other membrane proteins, such as G-proteins, which move about in the fluid

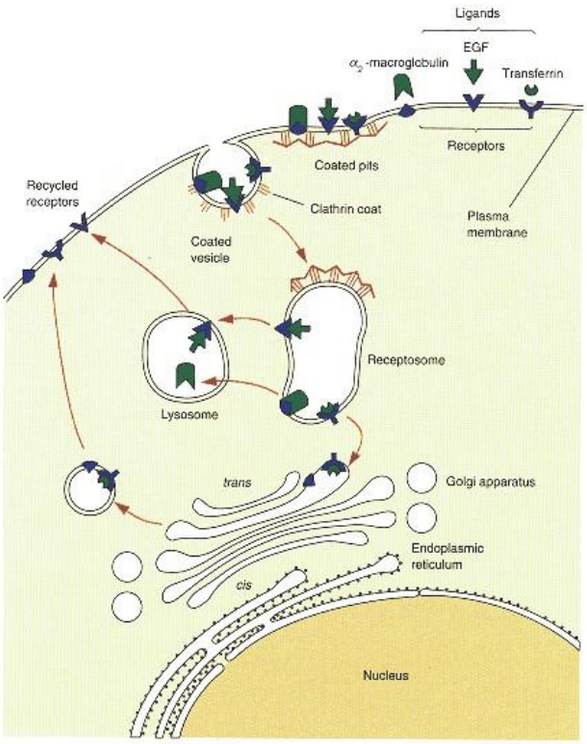


Figure 20.33
Diagrammatic summary of the morphological pathway of endocytosis in cells.
The morphological elements of the pathway of endocytosis are not drawn to scale. The ligands shown as examples are EGF, transferrin, and $_2$-macroglobulin. EGF is an example of a receptor system in which both ligand and receptor are delivered to lysosomes; transferrin is shown as an example of a system in which both the ligand and receptor recycle to the surface; $_2$-macroglobulin is shown as an example of a system in which the ligand is delivered to lysosomes but the receptor recycles efficiently back to the cell surface via the Golgi apparatus.
Adapted from Pastan, I., and Willingham, M. C. (Eds.). *Endocytosis*. New York: Plenum Press, 1985, p. 3.

cell membrane. However, many types of cell membrane hormone–receptor complexes are internalized, that is, moved from the cell membrane to the cell interior by a process called **endocytosis.** This would represent the opposite of exocytosis in which components within the cell are moved to the cell exterior. The process of endocytosis as presented in Figure 20.33 involves the polypeptide–receptor complex bound in **coated pits,** which are indentations in the plasma membrane that invaginate into the cytosol and pinch off from the membrane to form **coated vesicles.** The vesicles shed their coats, fuse with each other, and form vesicles called receptosomes. The receptors and ligands on the inside of these **receptosomes** can have different fates. Receptors can be recycled to the cell surface following fusion with the Golgi apparatus. Alternatively, the vesicles can fuse with lysosomes for degradation of both the receptor and hormone. In addition, some hormone–receptor complexes are dissociated in the lysosome and only the hormone is degraded, while the receptor is returned intact to the membrane. In some systems, the receptor may also be concentrated in coated pits in the absence of exogenous ligand and cycle in and out of the cell in a constitutive, nonligand-dependent manner.

Clathrin Forms a Lattice Structure to Direct Internalization of Hormone–Receptor Complexes from the Plasma Membrane

The major protein component of the coated vesicle is **clathrin,** a nonglycosylated protein of mol wt 180,000 whose amino acid sequence is highly conserved. The coated vesicle contains 70% clathrin, 5% polypeptides of about 35 kDa, and 25% polypeptides of 50–100 kDa. Aspects of the structure of a coated vesicle are shown in Figure 20.34. Coated vesicles have a lattice-like surface

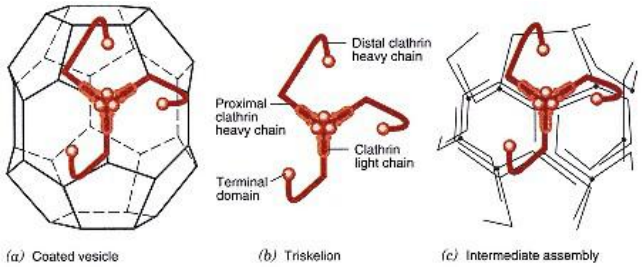


Figure 20.34
Structure and assembly of a coated vesicle.
(a) A typical coated vesicle contains a membrane vesicle about 40 nm in diameter surrounded by a fibrous network of 12 pentagons and 8 hexagons. The fibrous coat is constructed of 36 clathrin triskelions. One clathrin triskelion is centered on each of the 36 vertices of the coat. Coated vesicles having other sizes and shapes are believed to be constructed similarly: each vesicle contains 12 pentagons but a variable number of hexagons.
(b) Detail of a clathrin triskelion. Each of three clathrin heavy chains is bent into a proximal arm and a distal arm. A clathrin light chain is attached to each heavy chain, most likely near the center.
(c) An intermediate in the assembly of a coated vesicle, containing 10 of the final 36 triskelions, illustrates the packing of the clathrin triskelions. Each of the 54 edges of a coated vesicle is constructed of two proximal and two distal arms intertwined. The 36 triskelions contain 36 × 3 = 108 proximal and 108 distal arms, and the coated vesicle has precisely 54 edges.
See Crowther, R. A., and Pearse, B. M. F. *J. Cell Biol*. 91:790, 1981.
Redrawn from Nathk, I. S., Heuser, J., Lupas, A., Stock, J., Turck, C. W., and Brodsky, E. M. *Cell* 68:899, 1992.
Redrawn from Darnell, J., Lodish, H., and Baltimore, D. *Molecular Cell Biology*. New York: Scientific American Books, 1986, p. 647.

structure comprised of hexagons and pentagons. Three clathrin molecules generate each polyhedral vertex and two clathrin molecules contribute to each edge. The smallest such structure would contain 12 pentagons with 4–8 hexagons and 84 or 108 clathrin molecules. A 200-nm diameter coated vesicle contains about 1000 clathrin molecules. Clathrin can form flexible lattice structures that can act as scaffolds for vesicular budding. Completion of the budding process results in the mature vesicle being able to enter the cycle.

The events following endocytosis are not always clear with respect to a specific membrane receptor system. This process can be a means to introduce the intact receptor or ligand to the cell interior in cases where the nucleus is thought to contain a receptor or ligand-binding site. Consider, for example, growth factors that are known to bind to a cell membrane receptor but trigger events leading to mitosis. It is possible that signal transmission occurs by the alteration of a specific cytosolic protein, perhaps by membrane growth factor receptor-associated protein kinase activity, resulting in the nuclear translocation of the covalently modified cytosolic protein. In the case of internalization, delivery of an intact ligand (or portion of the ligand) could interact with a nuclear receptor. Such mechanisms are speculative. Nevertheless, these ideas could constitute a rationale for the participation of endocytosis in signal transmission to intracellular components.

Endocytosis renders a cell less responsive to hormone. Removal of the receptor to the interior, or cycling of membrane components, alters responsiveness or metabolism (e.g., glucose receptors can be shuttled between the cell interior and the cell membrane under the control of hormones in certain cells). In another type of downregulation, a hormone–receptor complex translocated to the nucleus can repress its own receptor mRNA levels by interacting with a specific DNA sequence. More about this form of receptor downregulation is mentioned in Chapter 21.

20.12— Intracellular Action: Protein Kinases

Many amino acid-derived hormones or polypeptides bind to cell membrane receptors (except for thyroid hormone) and transmit their signal by (1) elevation of cAMP and transmission through the **protein kinase A pathway;** (2) triggering of the hydrolysis of phosphatidylinositol 4,5-bisphosphate and stimulation of the **protein kinase C** and IP_3–Ca^{2+} pathways; or (3) stimulation of intracellular levels of cGMP and activation of the **protein kinase G pathway.** There are also other less prevalent systems for signal transfer, which, for example, affect molecules in the membrane like phosphatidylcholine. As previously discussed in the case of TSH–receptor signaling, it may be possible that two of these pathways are activated.

The cAMP system operating through protein kinase A activation has been described. Specific proteins are expected to be phosphorylated by this kinase compared to other protein kinases, such as protein kinase C. Both protein kinase A and C phosphorylate proteins on **serine** or **threonine** residues. An additional protein kinase system involves phosphorylation of **tyrosine,** which occurs in cytoplasmic domains of some membrane receptors especially growth factor receptors. This system is important for the insulin receptor, IGF receptor, and certain oncogenes discussed below. The cellular location of these protein kinases is presented in Figure 20.35.

The catalytic domain in the protein kinases is similar in amino acid sequence, suggesting that they have all evolved from a common primordial kinase. The three **tyrosine-specific kinases** shown in Figure 20.35 are transmembrane receptor proteins that, when activated by the binding of specific extracellular ligands, phosphorylate proteins (including themselves) on tyrosine residues inside the cell. Both chains of the insulin receptor are encoded by a single

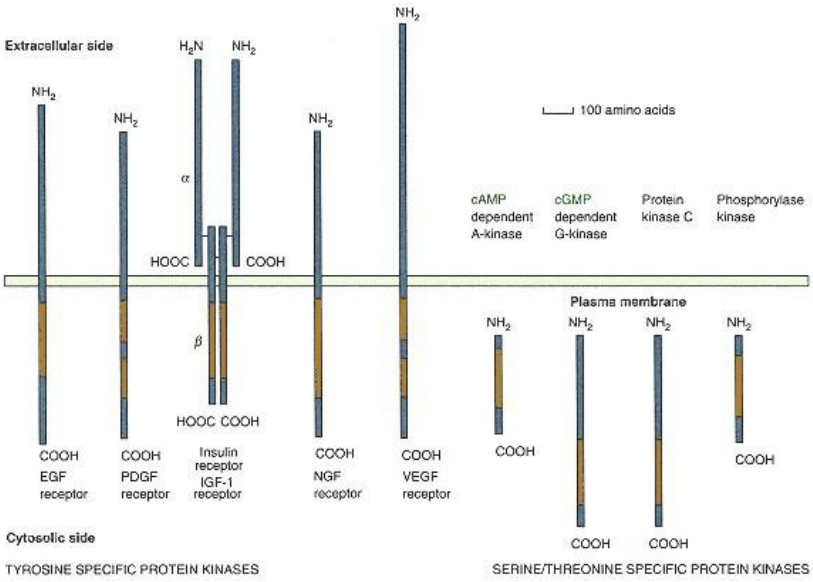


Figure 20.35
Protein kinases showing the size and location of their catalytic domain.
In each case the catalytic domain (red region) is about 250 amino acid residues long. The regulatory subunits normally associated with A-kinase and with phosphorylase kinase are not shown. EGF, epidermal growth factor; NGF, nerve growth factor; VEGF, vascular endothelial growth factor.
Redrawn from Alberts, B., Bray, D., Lewis, J., Raff, M., Roberts, K., and Watson, J. D. *Molecular Biology of the Cell*, 3rd ed. New York: Garland Publishing, 1994, p. 760.

gene, which produces a precursor protein that is cleaved into the two disulfide-linked chains. The extracellular domain of the PDGF receptor is thought to be folded into five immunoglobulin (Ig)-like domains, suggesting that this protein belongs to the Ig superfamily.

Proteins that are regulated by phosphorylation–dephosphorylation can have multiple phosphorylation sites and may be phosphorylated by more than one class of protein kinase.

Insulin Receptor:
Transduction through Tyrosine Kinase

From Figure 20.35 it is seen that the α subunits of the **insulin receptor** are located outside the cell membrane and apparently serve as the insulin-binding site. The insulin–receptor complex undergoes an activation sequence probably involving conformational changes and phosphorylation **(autophosphorylation)** of tyrosine residues located in the cytoplasmic portion of the receptor β subunits). This results in activation of the tyrosine kinase activity located in the β subunit, which is now able to phosphorylate cytosolic proteins that may carry the insulin signal to the interior of the cell. The net results of these phosphorylation events include a series of short-term metabolic effects, such as increased uptake of glucose, as well as longer-term effects of insulin on cellular differentiation and growth. Although, as already indicated, the insulin receptor itself is a tyrosine kinase that is activated upon hormone binding, the

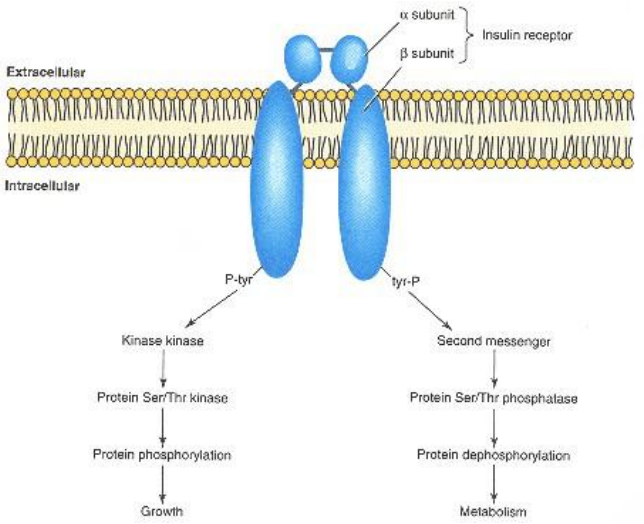


Figure 20.36
Hypothetical model depicting two separate biochemical pathways to explain paradoxical effects of insulin on protein phosphorylation.
Insulin simultaneously produces increases in the serine/threonine phosphorylation of some proteins and decreases in others. This paradoxical effect may result from activation of both kinases and phosphatases. Model explains (1) the generation of a soluble second messenger that directly or indirectly activates serine/threonine phosphatase and (2) the stimulation of a cascade of protein kinases, resulting in phosphorylation of cellular proteins.
Redrawn from Saltiel, A. R. The paradoxical regulation of protein phosphorylation in insulin action. *FASEB J.* 8:1034, 1994.

subsequent changes in phosphorylation occur predominantly on serine and threonine residues, as indicated in Figure 20.36. As also shown, insulin can simultaneously stimulate the phosphorylation of some proteins and the dephosphorylation of other proteins. Either of these biochemical events can lead to activation or inhibition of specific enzymes involved in mediating the effects of insulin. These opposite effects (phosphorylation and dephosphorylation) mediated by insulin suggest that perhaps separate signal transduction pathways may originate from the insulin receptor to produce these pleiotropic actions. A hypothetical scheme for this bifurcation of signals in insulin's action is presented in Figure 20.37. The substrates of the insulin–receptor tyrosine kinase are an important current research effort since phosphorylated proteins could produce the long-term effects of insulin. On the other hand, there is evidence that an insulin second messenger may be developed at the cell membrane to account for the short-term metabolic effects of insulin. The substance released as a result of insulin–insulin receptor interaction may be a glycoinositol derivative that, when released from the membrane into the cytosol, could be a stimulator of phosphoprotein phosphatase. This activity would dephosphorylate a variety of enzymes, either activating or inhibiting them, and produce effects already known to be associated with the action of insulin. In addition, this second messenger, or the direct phosphorylating activity of the receptor tyrosine kinase, might explain the movement of glucose receptors (transporters) from the cell interior to the surface to account for enhanced cellular glucose utilization in cells that utilize this mechanism to control glucose uptake. These possibilities are reviewed in Figure 20.37. Activation of the enzymes indicated in this figure leads to increased metabolism of glucose while inhibition of the enzymes indicated leads to decreased breakdown of glucose or fatty acid stores.

Activity of Vasopressin:
Protein Kinase A

An example of the activation of the **protein kinase A** pathway by a hormone is the activity of arginine vasopressin (AVP) on the distal kidney cell. Here the action of **vasopressin (VP),** also called the antidiuretic hormone (Table 20.5),

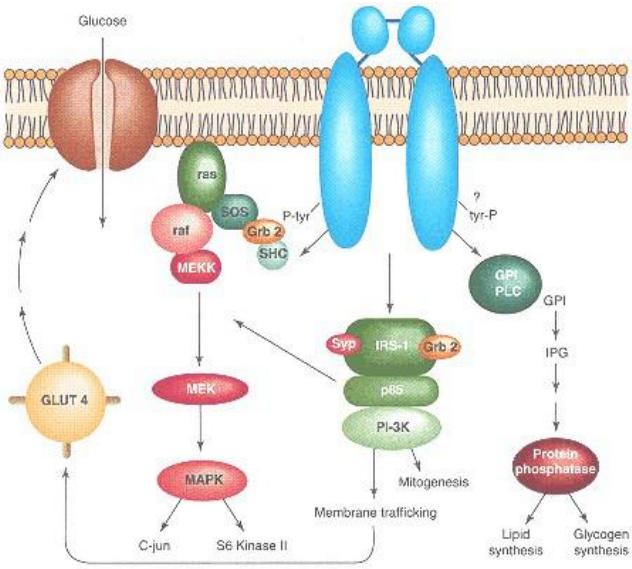


Figure 20.37
Hypothetical scheme for signal transduction in insulin action.
The insulin receptor undergoes tyrosine autophosphorylation and subsequent kinase activation upon hormone binding. The receptor phosphorylates intracellular substrates including IRS-1 and Shc proteins, which associate with SH2-containing proteins like p85, SYP, or Grb2 upon phosphorylation. Formation of the IRS-1–p85 complex activates PI 3-kinase; the IRS-1–SYP complex activates SYP, leading to MEK activation. Formation of the Shc–Grb2 complex mediates the stimulation of $P21^{Ras}$ GTP binding, leading to a cascade of phosphorylations. These phosphorylations probably occur sequentially and involve *raf* proto-oncogene, MEK, MAP kinase, and S6 kinase II. The receptor is probably separately coupled to activation of a specific phospholipase C that catalyzes the hydrolysis of the glycosyl-PI molecules in the plasma membrane. A product of this reaction, inositol phosphate glycan (IPG), may act as a second messenger, especially with regard to activation of serine/threonine phosphatases and the subsequent regulation of lipid and glucose metabolism. Abbreviations: IRS-1, insulin receptor substrate-1; SH, *src* homology; MAP kinase, mitogen-activated protein kinase; MEK, MAP kinase kinase; GPI, glycosylphosphatidylinositol; PLC, phospholipase; SOS, son of sevenless. Redrawn from Saltiel, A. R. The paradoxical regulation of protein phosphorylation in insulin action. *FASEB J*. 8:1034, 1994.

is to cause increased water reabsorption from the urine in the distal kidney. A mechanism for this system is shown in Figure 20.38. Neurons synthesizing AVP (vasopressinergic neurons) are signaled to release AVP from their nerve endings by interneuronal firing from a **baroreceptor** responding to a fall in blood pressure or from an **osmoreceptor** (probably an interneuron), which responds to an increase in extracellular salt concentration. The high extracellular salt concentration apparently causes shrinkage of the osmoreceptor cell and generates an electrical signal transmitted down the axon of the osmoreceptor to the cell body of the VP neuron generating an action potential. This signal is then transmitted down the long axon from the VP cell body to its nerve ending where, by depolarization, the VP–neurophysin II complex is released in to the

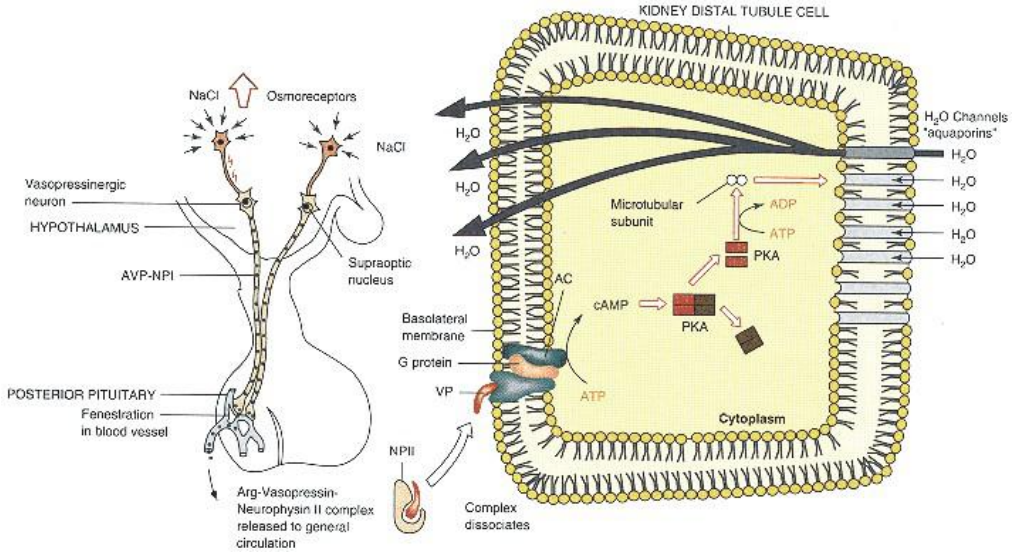


Figure 20.38
Secretion and action of arginine vasopressin in the distal kidney.
The release of arginine vasopressin (AVP or VP) from the posterior pituitary begins with a signal from the osmoreceptor, or baroreceptor (not shown), in the upper right-hand corner of figure. The signal can be an increase in the extracellular concentration of sodium chloride, which causes the osmoreceptor neuron to shrink and send an electrical message down its axon, which interfaces with the vasopressinergic cell body. This signal is transmitted down the long axon of the vasopressinergic neuron and depolarizes the nerve endings causing the release, by exocytosis, of the VP–neurophysin complex stored there. They enter the local circulation through fenestrations in the vessels and perfuse the general circulation. Soon after release, neurophysin dissociates from VP and VP binds to its cognate receptor in the cell membrane of the kidney distal tubule cell (other VP receptors are located on the corticotrope of the anterior pituitary and on the hepatocytes and their mechanisms in these other cells are different from the one for the kidney tubule cell). NPII, neurophysin II; VP, vasopressin; R, receptor; AC, adenylate cyclase; MF, myofibril; GP, glycogen phosphorylase; PK_i, inactive protein kinase; PK_a, active protein kinase; R-Ca, regulatory subunit–cyclic AMP complex; TJ, tight junction; PD, phosphodiesterase. Vasopressin–neurophysin complex dissociates at some point and free VP binds to its cell membrane receptor in the plasma membrane surface. Through a G-protein adenylate cyclase is stimulated on the cytoplasmic side of the cell membrane, generating increased levels of cAMP from ATP. Cyclic AMP-dependent protein kinases are stimulated and phosphorylate various proteins (perhaps including microtubular subunits) which, through aggregation, insert as water channels (aquaporins) in the luminal plasma membrane, thus increasing the reabsorption of water by free diffusion.
Redrawn in part from Dousa, T. P., and Valtin, H. Cellular actions of vasopressin in the mammalian kidney. *Kidney Int*. 10:45, 1975.

extracellular space. The complex enters local capillaries through fenestrations and progresses to the general circulation. The complex dissociates and free VP is able to bind to its cognate membrane receptors in the distal kidney, anterior pituitary, hepatocyte, and perhaps other cell types. After binding to the kidney receptor, VP causes stimulation of adenylate cyclase through the stimulatory G-protein and activates protein kinase A. The protein kinase phosphorylates

TABLE 20.7 Examples of Hormones that Operate Through the Protein Kinase A Pathway

Hormone	*Location of Action*
CRH	Corticotrope of anterior pituitary
TSH (also phospholipid metabolism?)	Thyroid follicle
LH	Leydig cell of testis
	Mature follicle at ovulation and *corpus luteum*
FSH	Sertoli cell of seminiferous tubule Ovarian follicle
ACTH	Inner layers of cells of adrenal cortex
Opioid peptides	Some in CNS function on inhibitory pathway through G_i
AVP	Kidney distal tubular cell (the AVP hepatocyte receptor causes phospholipid turn-over and calcium ion uptake; the AVP receptor in anterior pituitary causes phospholipid turnover)
PGI_2 (prostacyclin)	Blood platelet membrane
Norepinephrine/epinephrine	β-Receptor

microtubular subunits that aggregate to form specific water channels, referred to as aquaporins, which are inserted into the luminal membrane for admission of larger volumes of water than would occur by free diffusion. Water is transported across the kidney cell to the basolateral side and to the general circulation, causing a dilution of the original high salt concentration (signal) and an increase in blood pressure. These aquaporins, which are a family of integral membrane proteins that function as selective water channels, consist of six transmembrane α helical domains. Although aquaporin monomers function as water channels or pores, their stability and proper functioning may require a tetrameric assembly. Specific mutations in the amino acid sequences of the intracellular and extracellular loops of these proteins result in nonfunctional aquaporins and the development of diabetes insipidus, which is characterized by increased thirst and production of a large volume of urine.

Some hormones that operate through the protein kinase A pathway are listed in Table 20.7.

Gonadotropin-Releasing Hormone (GnRH): Protein Kinase C

Table 20.8 presents examples of polypeptide hormones that stimulate the phosphatidylinositol pathway. An example of a system operating through stimulation of the phosphatidylinositol pathway and subsequent activation of the **protein kinase C** system is **GnRH** action, shown in Figure 20.39. Probably under aminergic interneuronal controls, a signal is generated to stimulate the cell body of the GnRH-ergic neuron where GnRH is synthesized. The signal is transmitted down the long axon to the nerve ending where the hormone is stored. The hormone is released from the nerve ending by exocytosis resulting from depolarization caused by signal transmission. The GnRH enters the primary plexus of the closed portal system connecting the hypothalamus and anterior pituitary through fenestrations. Then GnRH exits the closed portal system through fenestrations in the secondary plexus and binds to cognate receptors in the cell membrane of the gonadotrope (see enlarged view in Figure 20.39). The signal from the hormone–receptor complex is transduced (through a G-protein) and phospholipase C is activated. This enzyme catalyzes the hydrolysis of PIP_2 to form DAG and IP_3. Diacylglycerol activates protein kinase C, which phosphoryl-

TABLE 20.8 Examples of Polypeptide Hormones that Stimulate the Phosphatidylinositol Pathway

Hormone	*Location of Action*
TRH	Thyrotrope of the anterior pituitary releasing TSH
GnRH	Gonadotrope of the anterior pituitary releasing LH and FSH
AVP	Corticotrope of the anterior pituitary; assists CRH in releasing ACTH; hepatocyte: causes increase in cellular Ca^{2+}
TSH	Thyroid follicle: releasing thyroid hormones causes increase in phosphatidylinositol cycle as well as increase in protein kinase A pathway
Angiotensin II/III	*Zona glomerulosa* cell of adrenal cortex: releases aldosterone
Epinephrine (thrombin)	Platelet: releasing ADP/serotonin; hepatocyte via α receptor: releasing intracellular Ca^{2+}

ates specific proteins, some of which may participate in the resulting secretory process to transport LH and FSH to the cell exterior. The product IP_3, which binds to a receptor on the membrane of the calcium storage particle, probably located near the cell membrane, stimulates the release of calcium ion. Elevated cytosolic Ca^{2+} causes increased stimulation of protein kinase C and participates in the exocytosis of LH and FSH from the cell.

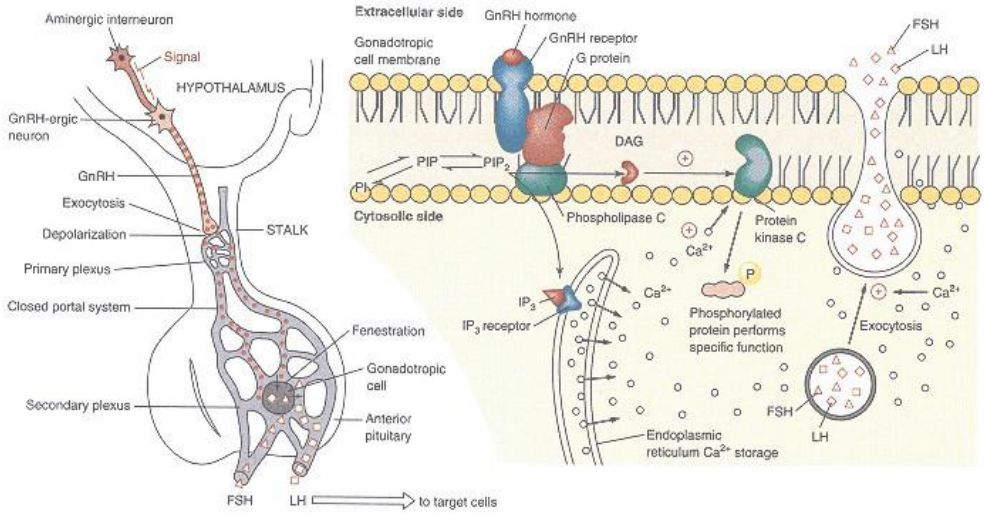


Figure 20.39
Overview of regulation of secretion of LH and FSH.
A general mode of action of GnRH to release the gonadotropes from the gonadotropic cell of the anterior pituitary is presented. GnRH, gonadotropin-releasing hormone; FSH, follicle-stimulating hormone; LH, luteinizing hormone; DAG, diacylglycerol.

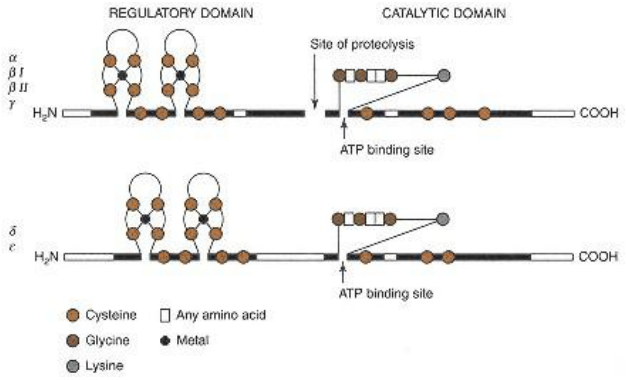


Figure 20.40
Common structure of protein kinase C subspecies.
Modified from U. Kikkawa, A. Kishimoto, and Y. Nishizuka, *Annu. Rev. Biochem.* 58:31, 1989.

Much recent work has focused on protein kinase C. It has been shown to have a number of subspecies; such heterogeneity may indicate that there are multiple functions for this critical enzyme (Figure 20.40). The enzyme consists of two domains, a regulatory and a catalytic domain, which can be separated by proteolysis at a specific site. The free catalytic domain, formerly called **protein kinase M,** can phosphorylate proteins free of the regulatory components. The free catalytic subunit, however, may be degraded. More needs to be learned about the dynamics of this system and the translocation of the enzyme from one compartment to another. The regulatory domain contains two Zn^{2+} fingers usually considered to be hallmarks of DNA-binding proteins (see Chapter 3). This DNA-binding activity has not yet been demonstrated for protein kinase C and metal fingers may participate in other types of interactions. The ATP-binding site in the catalytic domain contains the G box, GXGXXG, which is a consensus sequence for ATP binding with a downstream lysine residue.

Activity of Atrial Natriuretic Factor (ANF): Protein Kinase G

The third system is the **protein kinase G** system, which is stimulated by the elevation of cytosolic cGMP (Figure 20.41). **Cyclic GMP** is synthesized by guanylate cyclase from GTP. Like adenylate cyclase, guanylate cyclase is linked to a specific biological signal through a membrane receptor. The guanylate cyclase extracellular domain may serve the role of the hormone receptor. This is directly coupled to the cytosolic domain through one membrane-spanning domain (Figure 20.42), which may be applicable to the **atrial natriuretic factor** (ANF; also referred to as atriopeptin) **receptor–guanylate cyclase system.** Thus the hormone-binding site, transmembrane domain, and guanylate cyclase activities are all served by a single polypeptide chain.

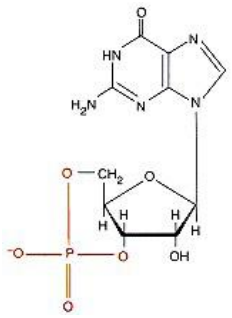


Figure 20.41
Structure of cGMP.

This hormone is a family of peptides, as shown in Figure 20.43; a sequence of human ANF is shown at the bottom. The functional domains of the ANF receptor are illustrated in Figure 20.44. Atrial natriuretic factor is released from atrial cells of the heart under control of several hormones. Data from atrial cell culture suggest that ANF secretion is stimulated by activators of protein kinase C and decreased by activators of protein kinase A. These opposing actions may be mediated by the actions of α- and β-adrenergic receptors, respectively. An overview of the secretion of ANF and its general effects is shown in Figure 20.45. ANF is released by a number of signals, such as blood volume expansion, elevated blood pressure directly induced by vasoconstrictors, high salt intake,

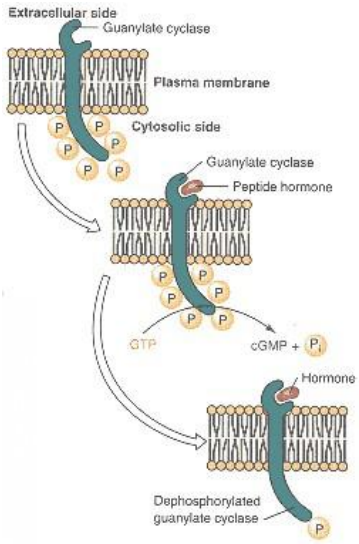


Figure 20.42
Model for the regulation of guanylate cyclase activity after peptide hormone binding.
The enzyme exists in a highly phosphorylated state under normal conditions. Binding of hormone markedly enhances enzyme activity, followed by a rapid dephosphorylation of guanylate cyclase and a return of activity to basal state despite continued presence of hormonal peptide.
Redrawn from Schultz, S., Chinkers, M., and Garbers, D. L. *FASEB J.* 3:2026, 1989.

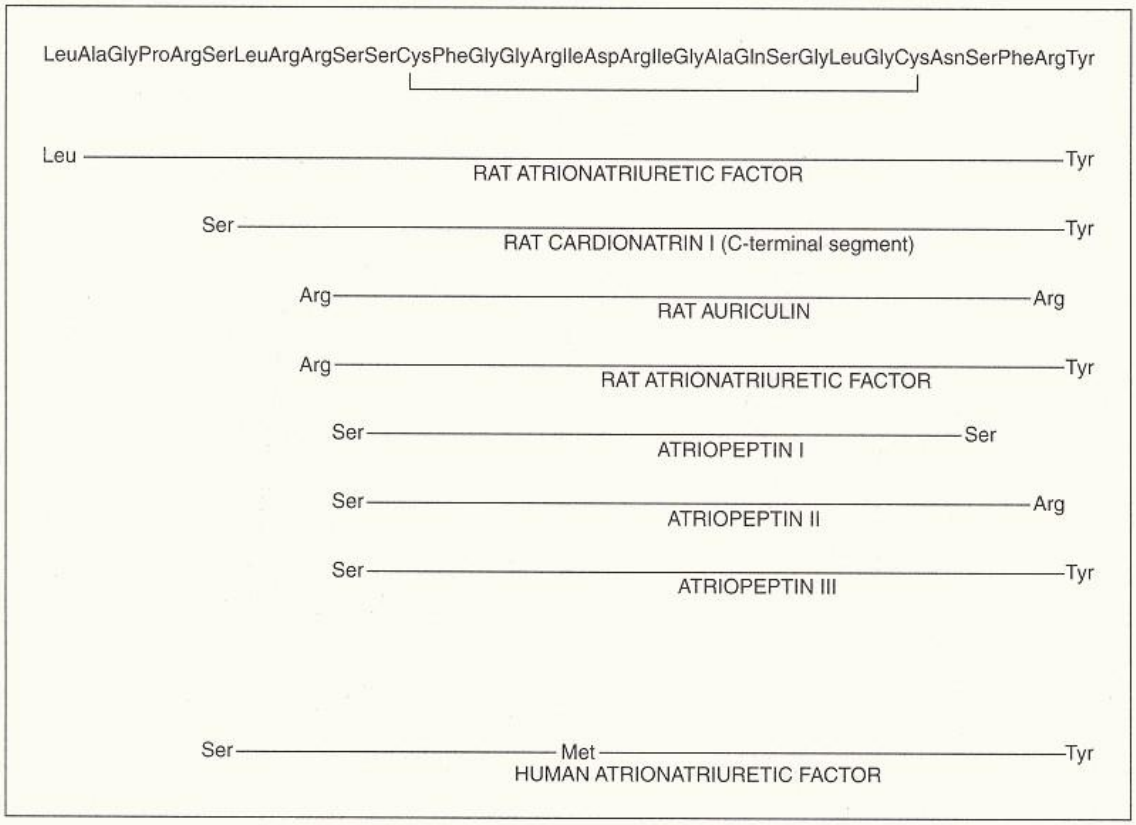


Figure 20.43
Atrial natriuretic peptides.
These active peptides relax vascular smooth muscle and produce vasodilation and natriuresis as well as other effects discussed in the text.
Adapted from Carlin, M., and Genest, J. The heart and the atrial natriuretic factor. *Endocr. Rev.* 6:107, 1985.

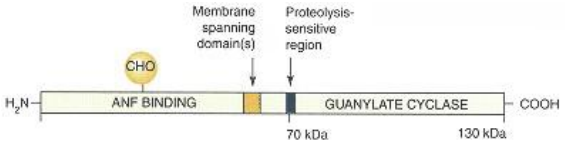


Figure 20.44
Functional domains of ANF-R_1 receptor.
Hypothetical model shows the sequence of an ANF-binding domain, a membrane-spanning domain(s), a proteolysis-sensitive region, a guanylate cyclase catalytic domain, glucosylation sites (CHO), and amino (H_2N) and carboxyl terminals (COOH) of receptor.
Redrawn from Liu, B., Meloche, S., McNicoll, N., Lord, C., and DeLéan, A. *Biochemistry* 28:5599, 1989.

and increased heart pumping rate. ANF is secreted as a dimer that is inactive for receptor interaction and is converted in plasma to a monomer capable of interacting with receptor. The actions of ANF (Figure 20.45) are to increase the glomerular filtration rate without increasing renal blood flow, leading to increased urine volume and excretion of sodium ion. Renin secretion is also reduced and aldosterone secretion by the adrenal cortex is lowered. This action reduces aldosterone-mediated sodium reabsorption. ANF inhibits the vasoconstriction produced by angiotensin II and relaxes the constriction of the renal vessels, other vascular beds, and large arteries. ANF operates through its mem-

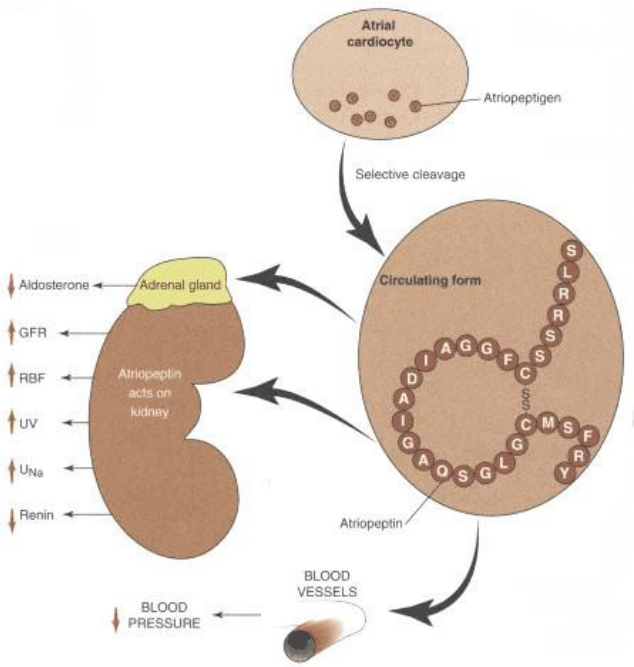


Figure 20.45
Schematic diagram of atrial natriuretic factor–atriopeptin hormonal system.
Prohormone is stored in granules located in perinuclear atrial cardiocytes. An elevated vascular volume results in cleavage and release of atriopeptin, which acts on the kidney (glomeruli and papilla) to increase the glomerular filtration rate (GFR), to increase renal blood flow (RBF), to increase urine volume (UV) and sodium excretion (U_{Na}), and to decrease plasma renin activity. Natriuresis and diuresis are also enhanced by the suppression of aldosterone secretion by the adrenal cortex and the release from the posterior pituitary of arginine vasopressin. Vasodilatation of blood vessels also results in a lowering of blood pressure (BP). Diminution of vascular volume provides a negative feedback signal that suppresses circulating levels of atriopeptin.
Redrawn from Needleman, P., and Greenwald, J. E. Atriopeptin: a cardiac hormone intimately involved in fluid, electrolyte, and blood pressure homeostasis. *N. Engl. J. Med*. 314:828, 1986.

brane receptor, which appears to be the extracellular domain of guanylate cyclase. The cGMP produced activates protein kinase G, which further phosphorylates cellular proteins to express many of the actions of this pathway. More needs to be learned about protein kinase G. Using analogs of ANF it has been shown that the majority of receptors expressed in the kidney are biologically silent, since they fail to elicit a physiological response. This new class of receptors may serve as specific peripheral storage–clearance binding sites and as such act as a hormonal buffer system to modulate plasma levels of ANF.

20.13— Oncogenes and Receptor Functions

Oncogenes are genes that are expressed by cancerous transformed cells. A cancer cell may express few or many oncogenes that dictate the aberrant uncontrolled behavior of the cell. There are three mechanisms by which oncogenes allow a cell to escape dependence on exogenous growth factors; these are presented in Figure 20.46. Some oncogenes are genes for parts of receptors, most often related to growth factor hormone receptors, which can function in the absence of the hormonal ligand. Thus an oncogene may represent a truncated gene where the ligand-binding domain is missing. This would result in production of the receptor protein, insertion into the cell membrane, and continuous constitutive function in the absence or presence of ligand (Figure 20.46*b,c*). In this situation the second messengers would be produced constitutively at a high rate, instead of being regulated by ligand, and the result would be uncontrolled growth of the cell. Some oncogenes may have tyrosine protein kinase activity and therefore function like tyrosine kinase normally related to certain cell membrane receptors. Other oncogenes relate to thyroid and steroid hormone receptors (see Chapter 21) while still others are DNA-binding proteins, some of which may be transactivating factors or related to such factors. Oncogene-encoded proteins that bind to DNA may be identical with or related to transactivating factors. The oncogene *Jun*, for example, is a component of activator protein 1 (AP1), a transactivating factor that regulates transcription. Table 20.9 reviews some of the oncogenes, or cancer-causing genes, together with the functions of their proto-oncogenes (normal proliferation gene).

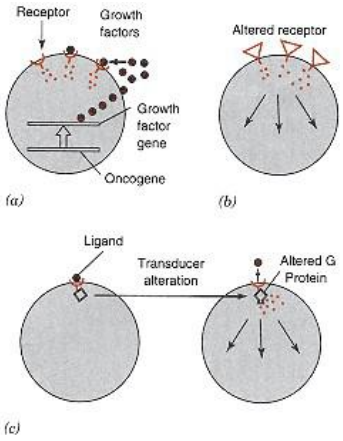


Figure 20.46
Mechanisms by which oncogenes can allow a cell to escape dependence on exogenous growth factors.
(a) By autocrine mechanism, where the cytosolic oncogene indirectly stimulates expression of growth factor gene and oversecretion of growth factors, which then over-stimulates receptors on same cell;
(b) by receptor alteration so that receptor is "permanently turned on" without a requirement for growth factor binding; and
(c) by transducer alteration, where the intermediate between the receptor and its resultant activity, that is, the GTP-stimulatory protein, is permanently activated, uncoupling the normal requirement of ligand–receptor binding.
Redrawn from Weinberg, R. A. The action of oncogenes in the cytoplasm and nucleus. *Science* 230:770, 1985.

TABLE 20.9 Known Oncogenes, Their Products and Functions[a]

			Oncogenic Protein	
Name of Oncogene	***Retrovirus***	***Virus-Induced Tumor***	***Cellular Location***	***Proto-oncogene Function***
src	Chicken sarcoma	Chicken sarcoma	Plasma membrane	Tyrosine-specific protein kinase
yes	Chicken sarcoma		Plasma membrane (?)	
fgr	Cat sarcoma		(?)	
abl	Mouse leukemia	Human leukemia	Plasma membrane	Tyrosine-specific protein kinase
fps	Chicken sarcoma		Cytoplasm (plasma membrane?)	
fes	Cat sarcoma	Sarcoma	Cytoplasm (cytoskeleton?)	Tyrosine-specific protein kinase
ros	Chicken sarcoma		(?)	
erb-B	Chicken leukemia	Erythroleukemia, fibrosarcoma	Plasma and cytoplasmic membranes	EGF receptor's cytoplasmic tyrosine-specific protein kinase domain
fms	Cat sarcoma	Sarcoma	Plasma and cytoplasmic membranes	Tyrosine-specific protein kinase; macrophage colony-stimulating factor receptor
mil	Chicken carcinoma		Cytoplasm	(?)
raf	Mouse sarcoma	Sarcoma	Cytoplasm	Protein kinase (serine/threonine) activated by Ras
mos	Mouse sarcoma	Mouse leukemia	Cytoplasm	(?)
sis	Monkey sarcoma	Monkey sarcoma	Secreted	PDGF-like growth factor, β-chain
Ha-ras	Rat sarcoma	Human carcinoma, rat carcinoma	Plasma membrane	GTP-binding protein
Ki-ras	Rat sarcoma	Human carcinoma, leukemia, and sarcoma	Plasma membrane	GTP-binding protein
N-ras	—	Human leukemia and carcinoma	Plasma membrane	
myc	Chicken leukemia	Sarcoma, myelocytoma, and carcinoma	Nucleus	DNA-binding related to cell proliferation; transcriptional control
myb	Chicken leukemia	Human leukemia	Nucleus	(?)
B-lym	—	Chicken lymphoma, human lymphoma	Nucleus (?)	(?)
ski	Chicken sarcoma		Nucleus (?)	(?)
rel	Turkey leukemia	Reticuloendotheliosis	(?)	(?)
erb-A	Chicken leukemia		(?)	Thyroid hormone receptor (c-erb-Aα1); related to steroid hormone receptors, retinoic acid receptor, and vitamin D_3 receptor
ets	Chicken leukemia		(?)	DNA binding
elk (*ets*-like)				DNA-binding protein
jun		Osteosarcoma		Products associate to form AP1 gene transcription factor
fos		Fibrosarcoma		Products associate to form AP1 gene transcription factor

Source: Adapted from Hunter, T. The proteins of oncogenes. *Sci. Am.* 251:70, 1984.

[a] The second column gives the source from which each viral oncogene was first isolated and the cancer induced by the oncogene. Some names, such as *fps* and *fes*, may be equivalent genes in birds and mammals. The third column lists human and animal tumors caused by agents other than viruses in which the *ras* oncogene or an inappropriately expressed proto-oncogene has been identified.

Bibliography

Alberts, B., Bray, D., Lewis, J., Raff, R., Roberts, K., and Watson, J. D. *Molecular Biology of the Cell*, 3rd ed. New York: Garland Publishing, 1994.

Cuatrecasas, P. Hormone receptors, membrane phospholipids, and protein kinases. *The Harvey Lectures* Series 80:89, 1986.

DeGroot, L. J., (Ed.). *Endocrinology*. Philadelphia: Saunders, 1995.

Hunter, T. The proteins of oncogenes. *Sci. Am.* 251:70, 1984.

Krieger, D. T., and Hughes, J. C. (Eds.). *Neuroendocrinology*. Sunderland, MA: Sinauer Associates, 1980.

Litwack, G. (Ed.). *Biochemical Actions of Hormones*, Vols. 1–14. New York: Academic Press, 1973–1987.

Litwack, G. (Ed. in Chief) *Vitamins and Hormones*, Vol. 50. Orlando: Academic Press, 1995.

Norman, A. W., and Litwack, G. *Hormones*. Orlando: Academic Press, 1987.

Richter, D. Molecular events in expression of vasopressin and oxytocin and their cognate receptors. *Am. J. Physiol.* 255:F207, 1988.

Ryan, R. J., Charlesworth, M. C., McCormick, D. J., Milius, R. P., and Keutmann, H. T. *FASEB J.* 2:2661, 1988.

Saltiel, A. R. The paradoxical regulation of protein phosphorylation in insulin action. *FASEB J.* 8:1034, 1994.

Spiegel, A. M., Shenker, A., and Weinstein, L. S. Receptor–effector coupling by G proteins: implication for normal and abnormal signal-transduction. *Endocr. Rev.* 13:536, 1992.

Struthers, A. D. (Ed.) *Atrial Natriuretic Factor*. Boston: Blackwell Scientific Publications, 1990.

Weinberg, R. A. The action of oncogenes in the cytoplasm and nucleus. *Science* 230:770, 1985.

Questions

J. Baggott and C. N. Angstadt

1. In a cascade of hormones (e.g., hypothalamus, pituitary, and target tissue), at each successive level:

   A. the quantity of hormone released and its half-life can be expected to increase.

   B. the quantity of hormone released increases, but its half-life does not change.

   C. the quantity of hormone released and its half-life are approximately constant.

   D. the quantity of hormone released decreases, but its half-life does not change.

   E. the quantity of hormone released and its half-life can both be expected to decrease.

2. All of the following have an identical (or very similar) α sub-unit EXCEPT:

   A. growth hormone.

   B. thyroid-stimulating hormone.

   C. luteinizing hormone.

   D. follicle-stimulating hormone.

3. If a single gene contains information for the synthesis of more than one hormone molecule:

   A. all the hormones are produced by any tissue that expresses the gene.

   B. all of the hormone molecules are identical.

   C. cleavage sites in the gene product are typically pairs of basic amino acids.

   D. all of the peptides of the gene product have well-defined biological activity.

   E. the hormones all have similar function.

4. In the sequence of events associated with signal transduction, which one is out of place? Receptor binds hormone.

   A. Conformational change occurs in receptor.

   B. Receptor interacts with G-protein.

   C. α Subunit of G-protein hydrolyzes GTP.

   D. α Subunit of G-protein dissociates from β and γ subunits.

   E. α Subunit of G-protein binds to adenylate cyclase.

5. The direct effect of cAMP in the protein kinase A pathway is to:

   A. activate adenylate cyclase.

   B. dissociate regulatory subunits from protein kinase.

   C. phosphorylate certain cellular proteins.

   D. phosphorylate protein kinase A.

   E. release hormones from a target tissue.

6. Activation of phospholipase C initiates a sequence of events including all of the following EXCEPT:

   A. release of inositol 4,5-bisphosphate from a phospholipid.

   B. increase in intracellular Ca^{2+} concentration.

   C. release of diacylglycerol (DAG) from a phospholipid.

   D. activation of protein kinase C.

   E. phosphorylation of certain cytoplasmic proteins.

7. In the ovarian cycle:

   A. GnRH enters the vascular system via transport by a specific membrane carrier.

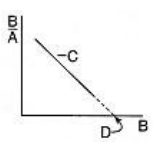
   B. the corpus luteum dies only if fertilization does not occur.

   C. inhibin works by inhibiting the synthesis of the α subunit of FSH.

   D. FSH activates a protein kinase A pathway.

   E. LH is taken up by the corpus luteum and binds to cytoplasmic receptors.

8. The Scatchard plot, shown in the accompanying figure, could be used to determine kinetic parameters of an enzyme. Which letter in the graph corresponds to total binding sites in a Scatchard plot or V_{max} in an enzyme kinetic plot?



9. With the anterior pituitary hormones, TSH, LH, and FSH:

   A. the α subunits are all different.

   B. the β subunits are specifically recognized by the receptor.

   C. the β subunit alone can bind to the receptor.

   D. hormonal activity is expressed through activation of protein kinase B.

   E. intracellular receptors bind these hormones.

10. In the interaction of a hormone with its receptor, all of the following are true EXCEPT:

A. more than one polypeptide chain of the hormone may be necessary.

B. more than one second messenger may be generated.

C. an array of transmembrane helices may form the binding site for the hormone.

D. receptors have a greater affinity for hormones than for synthetic agonists or antagonists.

E. hormones released from their receptor after endocytosis could interact with a nuclear receptor.

In the following questions, match the numbered hormone with the lettered kinase it stimulates.

A. protein kinase A

B. tyrosine kinase

C. protein kinase C

D. protein kinase G

11. Atrial natriuretic factor.

12. Gonadotropin-releasing hormone.

13. Insulin.

14. Vasopressin.

Answers

1. A Each successive step typically releases a larger amount of a longer lived hormone (p. 842).

2. A All of these are anterior pituitary hormones, but only the last three, the glycoprotein hormones, have an α subunit that is similar or identical from hormone to hormone (p. 846).

3. C One or more trypsin-like proteases catalyze the reaction (Figure 20.5). A: The POMC gene product is cleaved differently in different parts of the anterior pituitary (p. 849). B: Multiple copies of a single hormone may occur (p. 852), but not necessarily (Figure 20.5, p. 850). D: Some fragments have no known function. E: ACTH and β-endorphin, for example, hardly have similar functions (p. 847; Table 20.2).

4. C Hydrolysis of GTP returns the α subunit to its original conformation and allows it to associate with the β and γ subunits (p. 861).

5. B cAMP binding causes a conformational change in the regulatory subunits, resulting in the release of active protein kinase A (p. 862).

6. A Inositol 1,4,5-triphosphate (IP_3) is released from the phospholipid, phosphatidylinositol 4,5-bisphosphate (PIP_2) (p. 862).

7. D A: GnRH enters the vascular system through fenestrations (p. 868). B. The corpus luteum is replaced by the placenta if fertilization occurs (p. 870). C: The glycoprotein hormones share a common α subunit. Specific control of them would not involve a subunit they share. E: LH interacts with receptors on the cell membrane.

8. D A is free ligand concentration (analogous to substrate concentration), B is bound ligand concentration (analogous to), C is the equilibrium constant (analogous to K_m), and D is the extrapolated maximum number of binding sites (analogous to V_{max}) (p. 872).

9. B A: The α subunits are identical or nearly so (p. 872). B and C: Although specificity is conferred by the β subunits, which differ among the three hormones, binding to the receptor requires both subunits (p. 873). D: It is protein kinase A, and perhaps also protein kinase C in the case of TSH (p. 873). E: These large glycoprotein hormones do not penetrate the cell membrane; they bind to receptors on the cell surface (p. 874). See Figure 20.30.

10. D β Receptors bind isoproterenol more tightly than their hormones (p. 875). A and B: These are true of the glycoprotein hormones (p. 873). C: This appears to be true for the $_1$ receptor (Figure 20.31). E: This is possible, but entirely speculative; there are currently no known examples.

11. D See p. 885.

12. C See p. 883.

13. B See p. 879.

14. A See p. 880.