Translation of RNA: The Genetic Code and Protein Metabolism

12.1 The Process of Protein Synthesis
- Characteristics of Translation

12.2 The Three Stages of Protein Synthesis
- Initiation, Elongation, and Termination
- Polyribosomes
- Protein Synthesis and Energy
- Inhibition of Protein Synthesis by Antibiotics

12.3 Post-translational Processing of Proteins
- Protein Folding
- Biochemical Modifications
- Protein Targeting
- The Proteasome and Protein Degradation

12.4 Regulation of Protein Synthesis and Gene Expression
- Principles of Regulating Gene Expression
- Three Classes of Regulatory Proteins
- Examples of Gene Regulation
- Regulation of Gene Expression by RNA

Ricin, a protein found in the maturing seeds of the castor bean plant (Ricinus communis), is one of the most toxic naturally occurring chemicals known. If injected, the fatal dose in an adult is about 1 µg per kilogram of body weight. When inhaled or ingested, the fatal dose is slightly higher. Initial symptoms are flulike and death usually occurs within 36–72 hours. Ricin inactivates the 28S ribosomal RNA, thus it is a potent inhibitor of eukaryotic protein synthesis. The toxin has received special attention because of its use in terrorists’ acts and potential use as a chemical or biological agent in war.

Castor bean seeds are processed throughout the world by pressing to extract the natural laxative, castor oil. Ricin, which is very water soluble, does not extract with the oil but remains in the waste mash after pressing. The toxin may be isolated from the mash in just a few chemical steps.

An antidote for ricin poisoning has not yet been found. Dr. Ellen Vitetta and colleagues at the University of Texas Southwestern Medical Center, Dallas, are attempting to develop a recombinant DNA vaccine for use in humans. They have prepared a vaccine that is effective in mice and will soon be tested in humans. (See Biochemistry in the Clinic 12 for biochemical and medical information about ricin.)

(© Kjell B. Sandved/Photo Researchers.)
We turn now to the final step in the flow of genetic information, the synthesis of proteins. We have previewed the transfer pathway, DNA → RNA → proteins; we began our detailed journey with DNA, the storehouse of all genetic information. The term “replication” describes the duplication of DNA for future generations of cells. The message in DNA is in the form of a linear sequence of four nucleotide bases (ATGC). Within a cell the 4-letter language of DNA is transcribed into a similar language for mRNA (AUGC). In the transcription process we have noted a one-to-one correspondence between the nucleotide bases. This correspondence has its origin in the formation of hydrogen bonded complementary base pairs (A-U, A-T, G-C). The final step, transfer of sequential information in mRNA to the amino acid sequence of proteins, is much more complicated. A translation process is required to convert the 4-letter language of RNA into the 20-letter language of protein molecules (amino acids). Amino acids and nucleotide bases are quite different chemically, and no direct biochemical correlations or correspondence have been observed. We describe in this chapter how Francis Crick in 1958 suggested that some kind of “adaptor” molecules carried amino acids to the mRNA template for protein synthesis. Those molecules that act as translators are now identified as tRNAs. These small nucleic acids are linked together with the appropriate amino acid.

Other important concepts of protein synthesis introduced in this chapter include ribosomes, the ribozyme-containing molecular machines that make proteins, the workings of the genetic code, the chemistry and mechanics of protein synthesis, and post-translational modification of protein products.

The chapter concludes with a description of the regulation of protein synthesis and gene expression. It would be wasteful for a cell to synthesize all proteins at all times. Only a small percentage of the cell’s proteins are required continuously. The cell uses intricate mechanisms to regulate and coordinate the presence of appropriate proteins at the proper concentrations.
Eukaryotes have larger, more complex ribosomes than prokaryotes (Figure 12.1b). The intact mammalian ribosome is an 80S particle with 60S and 40S subunits. At least 80 distinct proteins are present in the 80S particle. Functionally, ribosomes are molecular structures that facilitate all the steps of protein synthesis: (1) They move along mRNA templates deciphering the code for conversion from nucleotide to amino acid sequence; (2) they bring to the template the proper adaptor molecule “charged” with the proper amino acid; and (3) they catalyze the formation of peptide bonds between amino acids using energy from ATP or GTP. During protein synthesis the two ribosomal subunits combine together so as to form a channel through which the mRNA moves. If the purified RNA and protein components of ribosomes are mixed under proper conditions, they self-assemble into intact functional ribosomes, capable of directing protein synthesis.

The sizes of intact ribosomes, their individual subunits, and their component biomolecules are correlated with their sedimentation coefficients ($S$). This is a measure of how rapidly a molecule or particle sediments during gradient centrifugation (see Window on Biochemistry 1-2; Chapter 10, Sections 10.1 and 10.3). The higher the $S$ value, the larger the particle and the more rapidly it settles. The intact *E. coli* ribosome has a sedimentation coefficient of 70S. The two subunits are defined as 50S and 30S particles. The *E. coli* ribosome consists of approximately 66% RNA and 34% protein. Its structural organization is described in Table 12.1 and Figure 12.1a. New X-ray crystallographic, high-resolution structures of the subunits of prokaryotic ribosomes were published in 2000. The new pictures confirmed the long-accepted structure as described here and also provided updated information about the catalytic mechanism of protein synthesis (Section 12.2).

### Table 12.1
**Structural organization of the *E. coli* ribosome**

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Intact Ribosome</th>
<th>Large Subunit</th>
<th>Small Subunit</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sedimentation coefficient</td>
<td>70S</td>
<td>50S</td>
<td>30S</td>
</tr>
<tr>
<td>Molecular mass (kilodaltons)</td>
<td>2520</td>
<td>1590</td>
<td>930</td>
</tr>
<tr>
<td>RNA content</td>
<td>66%</td>
<td>23S (2904 bases)</td>
<td>16S (1542 bases)</td>
</tr>
<tr>
<td></td>
<td>5S (120 bases)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Protein content</td>
<td>34%</td>
<td>34 different proteins</td>
<td>21 different proteins</td>
</tr>
</tbody>
</table>

**Figure 12.1** (a) Model for the *E. coli* ribosome showing the individual and combined subunits. The individual 30S and 50S subunits reversibly associate into the 70S complex, which functions in protein synthesis. The 70S complex moves along the mRNA to be translated. (b) Model for the eukaryotic ribosome made up of an 80S particle.
The concept that translation occurs only in the cellular cytoplasm has become one of the central tenets of modern cell biology. The idea is so accepted as biology dogma that it is often the basis of exam questions of the type: “Where does translation occur?” The question, if answered by today’s biology student (and biology professor), would be a resounding “cytoplasm only!” But this dogma, like many other ideas in science, is being challenged by research from several laboratories around the world showing evidence for nuclear translation. Before we describe the new results, let’s review what we have read in this text about protein synthesis from DNA, with a special focus on the preliminary processes and the differences between prokaryotic and eukaryotic cells.

The steps of protein production prior to translation, which have been described in Chapter 11, differ considerably in prokaryotic and eukaryotic cells. In prokaryotic cells synthesis of mRNA from DNA (transcription) occurs in the uncompartmentalized cytoplasm. Most prokaryotic mRNA requires little or no post-transcriptional alteration before it is translated into protein structures by the ribosomal particles. In fact, it is sometimes translated while it is being transcribed. Biologists generally accept the fact that the transcription and translation processes are coupled in bacteria. In compartmentalized eukaryotic cells, however, transcription and translation processes are separated by time and space. Transcription occurs in the nucleus and the primary transcripts require extensive post-transcriptional modification before they can be exported to the cytoplasm for translation. Important changes include removal of noncoding introns, 5’-capping, and poly A addition (see Section 11.5). Preparing mRNA for translation is considered to be an important function of the nucleus. Processed mRNA transcripts are then sent out to the cytoplasm where they attach to ribosomes for translation. These basic concepts about protein synthesis have developed from many observations in research laboratories during the last 25–30 years and have become fundamental tenets in biology.

In 2001, Peter Cook at the University of Oxford, United Kingdom, published research results suggesting that translation occurs not only in the cytoplasm, but also in the eukaryotic cell nucleus. From his studies, he estimated that as much as 15% of cellular protein synthesis may occur in the nucleus. His results also led him to the conclusion that transcription and translation may be coupled in eukaryotic cells just like in bacteria. These surprising results, which have been supported in several other labs, are contrary to biological dogma, very controversial, and not yet accepted by mainstream biology. Some biologists, who hold their beliefs about cell biology with all the zeal of a religion, label the new postulates as “heresy.” Most arguments against nuclear translation emphasize that no one has yet shown that the nucleus contains the complex molecular machinery necessary for protein synthesis, and that many of Cook’s experiments lacked the proper controls. If Cook’s results are shown to be correct, very important questions must be answered. What is the nature of the nuclear template, mRNA? Does the process occur with the same mechanism as in the cytoplasm? What happens to the nuclear proteins? Biologists around the world are waiting for more supporting evidence before they believe this major change in a fundamental tenet of modern biology.

Protein Synthesis Begins at the Amino Terminus  An early view of protein synthesis portrays a polymerization process whereby amino acids are added one at a time to a growing end. But what is the direction of elongation? Does it proceed from the amino terminus to the carboxyl terminus or vice versa (Figure 12.2)? Howard Dintzis and colleagues in 1961 answered this question by incubating tritium-labeled leucine with rabbit reticulocytes (immature red blood cells) that were synthesizing hemoglobin. The location of radioactive leucine in product hemoglobin allowed Dintzis to conclude that amino acids are added to the growing carboxyl terminus of the polypeptide. Thus, protein synthesis proceeds from the amino terminus to the carboxyl terminus (mode 1 in Figure 12.2).

The experiments performed by Dintzis had an additional benefit in that they provided useful information about the time required for synthesis of a polypeptide. A single ribosome in the rabbit reticulocytes completed the α-chain of hemoglobin (146 amino acids) in about 3 min at 37°C, a rate of slightly less than one residue added per second. An E. coli ribosome can construct a 100-amino acid polypeptide at 37°C in about 5 s (20 residues per second).
Amino Acids Are Activated and Combined with Specific tRNAs  

We come now to what was perhaps the most difficult concept to elucidate in protein synthesis: How is the 4-letter language of mRNA sequence translated into the 20-letter language of proteins? Francis Crick suggested that an “adaptor” molecule spanned the information gap between the two languages. In Crick’s words:

One would expect, therefore, that whatever went on to the template (mRNA) in a specific way did so by forming hydrogen bonds. It is therefore a natural hypothesis that the amino acid is carried to the template by an “adaptor” molecule, and that the adaptor is the part which actually fits on to the RNA. In its simplest form one would require twenty adaptors, one for each amino acid.

What sort of molecules such adaptors might be is anybody’s guess. . . . But there is one possibility which seems inherently more likely than any other—that they might contain nucleotides. This would enable them to join on to the RNA template by the same “pairing” of bases as is found in DNA, or in polynucleotides. (Crick, F., 1958. Symp. Soc. Exp. Biol. 12: 138–163.)

We now know that the tRNA molecules serve as the adaptor molecules. Crick’s suggestion that nucleotides might be involved was indeed insightful. As shown in Figure 12.3, the amino acid is covalently linked by an ester bond to the 2’- or 3’-hydroxyl end of a specific tRNA. At the other end of the tRNA, three adjacent nucleotide bases (the anticodon) bind to the mRNA template via hydrogen bonding between complementary base pairs.

Amino acids are linked to tRNAs by enzymes called **aminoacyl-tRNA synthetases**. It is critical that these enzymes display two kinds of specificity: They must recognize both the proper tRNA and the proper amino acid for linkage. This is a difficult task for the enzymes because all tRNA molecules have very similar primary, secondary, and tertiary structures except at the anticodon, where each is unique. In addition, the 20 amino acids, which have similar structural characteristics, must be distinguished...
from each other. With some exceptions cells are unable to correct the wrong pairing of an amino acid and tRNA. Most organisms have 20 aminoacyl-tRNA synthetases, one for each amino acid. As many as 60 different tRNA molecules have been isolated from organisms, so some amino acids have more than one tRNA. (We explain the significance of this later in this section.)

**Aminoacyl-tRNA Synthetases** Aminoacyl-tRNA synthetases catalyze a reaction sequence that requires the energy from cleavage of two phosphoanhydride bonds in ATP:

\[
\text{amino acid} + \text{ATP} \rightleftharpoons \text{aminoacyl-adenylate} + \text{PP}_i
\]

\[
\text{aminoacyl-adenylate} + \text{tRNA} \rightleftharpoons \text{aminoacyl-tRNA} + \text{AMP}
\]

\[
\text{PP}_i + \text{H}_2\text{O} \rightleftharpoons 2 \text{P}_i
\]

The aminoacyl-adenylate formed by the reaction of an amino acid with ATP is an unstable, enzyme-bound intermediate that contains a reactive anhydride linkage (Figure 12.4). The overall equilibrium constant for the first two reactions is approximately 1; therefore, a thermodynamic driving force is necessary to pull the reaction to completion. That extra energy comes from the hydrolysis of the pyrophosphate product (PP\(_i\)). Although we have emphasized the importance of this reaction process in bringing together an amino acid with its correct tRNA, the reaction has another important function. By linkage of the amino acid via an ester bond, the amino acid becomes activated for later peptidyl bond formation in protein synthesis.
Figure 12.4  Mechanism of aminoacyl-tRNA synthetase action with formation of an aminoacyl-adenylate intermediate. Energy from the cleavage of two phosphoanhydride bonds (ATP and PPi) drives the reaction to completion.
Relationships Between mRNA Base Sequence and Protein Amino Acid Sequence  Several terms have been used to describe the functional characteristics of the genetic code (Section 1.5):

- **Triplet.** A set of three nucleotide bases on mRNA code for one amino acid.
- **Nonoverlapping.** A set of three adjacent bases is treated as a complete group. The set of bases, called the **codon**, is used once for each translation step.
- **No punctuation.** There are no punctuation marks (intervening bases) between triplets. Therefore, an mRNA is read from start to finish without commas or other interruptions. (Perhaps the argument could be made that termination signals are periods.)
- **Degenerate.** A single amino acid may have more than one triplet code. There is usually a sequential relationship between degenerate codes.
- **Universal.** The same genetic code is used in all organisms except for a few exceptions, including mitochondria and some algae.

The first experiments designed to “break” the genetic code were reported by Marshall Nirenberg and Heinrich Matthaei in the early 1960s. They incubated cell-free extracts containing ribosomes, tRNA, amino acids, and aminoacyl-tRNA synthetases with synthetic mRNA of known sequences. For example, when poly U [U-(U)n-U] was used as template, the polypeptide product was polyphenylalanine, indicating that the base codon U-U-U was translated as the amino acid phenylalanine (Table 12.2). Similar experiments with other synthetic polyribonucleotides allowed researchers to elucidate completely the genetic code.

Table 12.2
The genetic code

<table>
<thead>
<tr>
<th>Second Base of Codon</th>
<th>A</th>
<th>G</th>
</tr>
</thead>
<tbody>
<tr>
<td>UUU</td>
<td>Phe</td>
<td>UAU</td>
</tr>
<tr>
<td>UUC</td>
<td>Leu</td>
<td>UAC</td>
</tr>
<tr>
<td>UUA</td>
<td>Leu</td>
<td>UAA</td>
</tr>
<tr>
<td>UUG</td>
<td>Leu</td>
<td>UAG</td>
</tr>
<tr>
<td>CUC</td>
<td>Leu</td>
<td>CAU</td>
</tr>
<tr>
<td>CUC</td>
<td>Leu</td>
<td>UAA</td>
</tr>
<tr>
<td>CUG</td>
<td>Leu</td>
<td>UAG</td>
</tr>
<tr>
<td>ACC</td>
<td>Pro</td>
<td>CAU</td>
</tr>
<tr>
<td>CCC</td>
<td>Pro</td>
<td>UAA</td>
</tr>
<tr>
<td>CGG</td>
<td>Pro</td>
<td>UAG</td>
</tr>
<tr>
<td>AUC</td>
<td>Ile</td>
<td>AAC</td>
</tr>
<tr>
<td>ACA</td>
<td>Ile</td>
<td>AAA</td>
</tr>
<tr>
<td>ACU</td>
<td>Ile</td>
<td>AAG</td>
</tr>
<tr>
<td>GUC</td>
<td>Val</td>
<td>GAC</td>
</tr>
<tr>
<td>GCC</td>
<td>Val</td>
<td>GAA</td>
</tr>
<tr>
<td>GCA</td>
<td>Val</td>
<td>GAG</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>First Base of Codon</th>
<th>Thr</th>
<th>Lys</th>
<th>Ser</th>
<th>Arg</th>
</tr>
</thead>
<tbody>
<tr>
<td>AUG</td>
<td>Thr</td>
<td>Lys</td>
<td>Ser</td>
<td>Arg</td>
</tr>
<tr>
<td>AUG</td>
<td>Thr</td>
<td>Lys</td>
<td>Ser</td>
<td>Arg</td>
</tr>
<tr>
<td>AUG</td>
<td>Thr</td>
<td>Lys</td>
<td>Ser</td>
<td>Arg</td>
</tr>
<tr>
<td>AUG</td>
<td>Thr</td>
<td>Lys</td>
<td>Ser</td>
<td>Arg</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Third Base of Codon</th>
<th>G</th>
<th>C</th>
</tr>
</thead>
<tbody>
<tr>
<td>GUU</td>
<td>Ala</td>
<td>Glu</td>
</tr>
<tr>
<td>GUC</td>
<td>Ala</td>
<td>Glu</td>
</tr>
<tr>
<td>GUA</td>
<td>Ala</td>
<td>Glu</td>
</tr>
<tr>
<td>GAG</td>
<td>Ala</td>
<td>Glu</td>
</tr>
</tbody>
</table>

Note: AUG is the start codon and UAA, UAG, and UGA are stop codons as highlighted in table.

Synthetic polynucleotides were also used to determine the direction mRNA is read, 5’ → 3’ or 3’ → 5’. The template 5’ A-A-A-(A-A-A)n-A-A-C was synthesized and used for protein synthesis. The triplet AAA codes for lysine and AAC for asparagine. The experimental polypeptide product from this template...
was $H_2N$-Lys-(Lys)$_n$-Asn-COO$^\cdot$. indicating that the direction of translation was 5' $\rightarrow$ 3'. It is significant that this is the same direction that the mRNA is formed. Hence, the ribosomes may begin translating the mRNA immediately after or perhaps even during its transcription.

As indicated earlier, some amino acids have degenerate codons; that is, they have more than one codon. Phenylalanine uses the two triplets UUU and UUC; threonine has four, ACU, ACC, ACA, and ACG. From a study of Table 12.2, it is apparent that when degeneracy exists, the first two bases are usually the same; only the third differs. It has been concluded that the first two bases of each codon are primary determinants for specificity, and one would expect strong interactions between these bases and two complementary bases on the tRNA called the **anticodon**. The third base interaction between the codon and anticodon is probably weaker and is sometimes called the **wobble base**. The wobble factor provides three benefits: (1) It adds flexibility to the genetic code; (2) the weaker interaction causes faster dissociation of the tRNA from the mRNA, thereby speeding protein synthesis; and (3) it minimizes the effects of mutations.

### Before You Go On...

1. What is the amino acid sequence of a polypeptide translated from the following mRNA?

   5' AUGUUUAACAAA
12.2 The Three Stages of Protein Synthesis

Figure 12.5 Steps of protein synthesis: (a) initiation, (b) formation of 70S complex, (c) elongation by peptide bond synthesis.
Figure 12.5—Continued Steps of protein synthesis: (d) translocation, (e) continued elongation to termination codon, and (f) termination. IF = initiation factors; RF = release factor.

Steps of protein synthesis:

(d) Translocation

(e) Continued elongation to termination codon

(f) Termination

mRNA 5' → 3'

Translocation

OH
GTP
Met
Phe
Asn
AAA
UAA
AAU

GDP + P_i + GTP

Asn + fMet

Many elongation, translocation steps

(fMet)
Phe
Asn
UUA

Release factor

(f)

Peptidyl transferase

H_2O

COO^-

OH

Asn-Phe-fMet + 30S + 50S

Dissociation of subunits

mRNA 5' → 3'

mRNA 5' → 3'

mRNA 5' → 3'
12.2 The Three Stages of Protein Synthesis

The 30S subunit and mRNA combine to place the two aminoacyl-tRNA binding sites (P for peptidyl; A for aminoacyl) directly over the codons for the first two amino acids. The ribosomal subunit is directed to this placement by the Shine–Dalgarno sequence of bases on the mRNA. (Eukaryotic mRNAs have no Shine–Dalgarno sequence, but translation probably begins at the first AUG codon of the mRNA.) This initiation region of approximately five to ten bases is centered at about ten bases upstream from the start codon. This purine-rich region pairs with a complementary pyrimidine-rich region of the 16S RNA in the 30S subunit. This correct alignment facilitates binding of the first activated amino acid, fMet-tRNA as shown in Figure 12.6, which will become the N-terminus. The binding of tRNA-activated amino acid is specified by codon–anticodon base pairing. All proteins in bacterial cells begin with the amino acid formylmethionine at the amino terminus. The start codon on the mRNA is AUG (see Table 12.2). The complete package of participants at this point is called the 30S initiation complex (see Figure 12.5a). Initiation proceeds with joining of the 50S subunit to the 30S complex. Formation of the new 70S initiation complex is accompanied by the hydrolysis of GTP (GTP $\rightarrow$ GDP + $\text{Pi}$ + energy) (Figure 12.5b). The 70S complex is now ready to begin the elongation stage of protein synthesis. (Eukaryotic initiation requires at least nine initiation factors.)

Stage 2 The participants in the elongation stage include the 70S initiation complex, the next aminoacyl-tRNA, and proteins called elongation factors (EF). This stage begins with entry of the second amino acid that is guided into the A ribosomal binding site by elongation factors. As with fMet-tRNA, the new aminoacyl-tRNA is selected by codon–anticodon interactions that specify binding. The complex is set for formation of the first peptide bond that links the carboxyl group of fMet to the amino group of the second amino acid (Figure 12.5c). This reaction produces a dipeptidyl-tRNA that is now positioned in the A ribosomal binding site.

An expanded view of peptide bond formation is shown in Figure 12.7. The formation of the peptide bond is catalyzed by peptidyl transferase, an RNA enzyme or ribozyme associated with the 50S ribosome. Highly resolved structures of the
Figure 12.7  Expanded view of peptide bond formation. The amino acid (Gly) to be added to the starting polypeptide chain is in the A site of the ribosome. The growing chain is in the P site. Peptidyl transferase, an RNA enzyme, catalyzes formation of the new peptide bond. Peptide bond formation is initiated by the removal of a proton from the amino group of bound glycine. The nitrogen atom then directs a nucleophilic attack on the ester bond holding fMet.
prokaryotic 70S ribosomal subunits have been produced by X-ray crystallography (Figure 12.8). These updated pictures show that there are no proteins in the vicinity of the peptidyl transferase active site where catalysis occurs and new covalent bond formation takes place. Only ribosomal RNA molecules are present at the active site. This new evidence confirms that the protein synthesis activity of the ribosome has its basis in RNA. RNA molecules in the subunits function as catalysts by orienting the two substrates, mRNA and aminoacyl-tRNA, close together so that peptide bond formation can occur. Apparently, proteins play predominately a structural role in ribosomes.

Formation of the peptide bond is a thermodynamically favorable reaction because of the reactive ester bond between fMet and its tRNA. The stage is now set for one of the mechanical motions of the ribosomal particles. The 70S ribosomal unit moves the distance of one codon toward the 3' end of the mRNA (Figure 12.5d). This action, called translocation, is accompanied by displacement of deacylated tRNA from the codon region and P site and eventually releases the free tRNA into the cytoplasm. At the same time the dipeptidyl-tRNA is shifted from the A site to the P site, leaving the A site vacant. Each translocation step is coupled to the hydrolysis of GTP. The third aminoacyl-tRNA of the peptide that is bound at the vacant A site is selected by codon–anticodon interactions and facilitated by elongation factors. The binding of each additional aminoacyl-tRNA is coupled to GTP hydrolysis. In fact, two GTPs are necessary for each amino acid, one for delivery and one for each translocation step. Protein synthesis proceeds with formation of the new peptide bond between the second and third amino acids, resulting in a tripeptide still linked to the tRNA of the third amino acid (Figure 12.5e).

**Stage 3** Termination is signaled by special codons on the mRNA. The elongation process continues as described above until the ribosomal A site moves to one of the codons for termination, UAG, UGA, or UAA. No aminoacyl-tRNAs exist with complementary anticodons for binding to these codons. Instead, protein release factors (RF) bind to the A site, thus activating peptidyl transferase. In place of peptide bond formation, the transferase catalyzes the hydrolysis of the ester bond linking the carboxyl group of the newly synthesized protein to the tRNA in the P site (Figure 12.5f). The protein product is released from the complex, tRNA and other factors diffuse into the cytoplasm while the 70S ribosome dissociates into its subunits to initiate translation of another mRNA molecule.
Before You Go On...

1. The enzyme peptidyl transferase, which catalyzes amide bond formation in the translation process, is a ribozyme. Describe chemical differences between the active site environment of a ribozyme and that of a regular protein enzyme. What kinds of functional groups are present at each? What kinds of catalysts, ribozymes or protein enzymes, have the potential to display greater variety in types of reactions catalyzed? Why?

Polyribosomes

We have demonstrated the process of protein synthesis using only one ribosome on a mRNA molecule. In reality, clusters of ribosomes called **polyribosomes** can simultaneously translate a mRNA molecule, thereby making many identical protein molecules from a single copy of mRNA (Figure 12.9). Each ribosomal particle along the mRNA is at a different stage of translation. Those ribosomes closer to the 3’ end of the mRNA are nearer to completion of the polypeptide product than those closer to the 5’ end. The actual number of ribosomes on a single mRNA depends on the size of the mRNA. The maximum number is probably about 1 ribosome per 80 nucleotides. The presence of polyribosomes, of course, adds immensely to the efficiency of protein synthesis. mRNA molecules have a relatively short lifetime, especially in prokaryotes, so it is important to use each one as efficiently as possible.

Protein Synthesis and Energy

The total energy requirements for protein synthesis are quite high (Table 12.3). To summarize energy needs:

1. Two phosphoanhydride bonds in ATP are cleaved for the activation of each amino acid and the synthesis of a specific aminoacyl-tRNA by aminoacyl-tRNA synthetase.
2. One GTP is required for entry of each amino acid into the ribosomal A site.
3. One GTP is required during each translocation step.
Table 12.3
Energy requirements for protein synthesis for each amino acid added

<table>
<thead>
<tr>
<th>Process</th>
<th>Reaction</th>
<th>Number of Phosphoanhydride Bonds Broken</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Selection and activation of amino acid</td>
<td>Amino acid + tRNA → aminoacetyl-tRNA</td>
<td></td>
</tr>
<tr>
<td></td>
<td>ATP → AMP + P_i</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>PP_i+H_2O → 2 P_i</td>
<td>1</td>
</tr>
<tr>
<td>2. Entry of aminoacyl-tRNA into ribosomal site A</td>
<td>GTP+H_2O → GDP + P_i</td>
<td>1</td>
</tr>
<tr>
<td>3. Translocation</td>
<td>GTP + H_2O → GDP + P_i</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Total:</td>
<td>4</td>
</tr>
</tbody>
</table>

Therefore, the total requirement is four phosphoanhydride bonds cleaved for each amino acid incorporated into a protein. As we shall see in Chapter 14, cleavage of each phosphoanhydride bond releases approximately 30 kJ/mol. This makes protein synthesis a very expensive, energy-demanding process. However, the high level of accuracy needed to synthesize functional proteins justifies this high cost.

Before You Go On...

1. How many phosphoanhydride bonds must be expended to make a dipeptide?
   Assume you are starting with a functional mRNA, ribosomal subunits, tRNA, and free amino acids.

Inhibition of Protein Synthesis by Antibiotics

Because protein synthesis is an essential biological process, it is an important target for drug design, especially for agents that kill infectious microorganisms. Prokaryotic protein synthesis is different enough from the more complex eukaryotic process that it is possible to find chemical substances that inhibit bacterial protein synthesis but have little effect on the process in eukaryotes. We have learned much about drug design by studying the natural antibiotics synthesized by some microorganisms as toxins to prevent the growth of competing microorganisms.

Puromycin  Extensive studies with the antibiotic puromycin, produced by a Streptomyces bacterium, have provided us with important knowledge about antibiotic action as well as detailed information about the mechanism of protein synthesis. Puromycin inhibits prokaryotic protein synthesis by mimicking aminoacyl-tRNA molecules (Figure 12.10). Its structure is similar enough to the amino acid tRNA ester that it binds to the ribosomal A site. The amino group on the antibiotic participates normally in peptide bond formation, producing a protein molecule with puromycin at the carboxyl end. However, the peptidyl–puromycin cannot be translocated to the P site for elongation because the very strong amide bond cannot be broken. Protein synthesis is terminated because the weakly bound peptidyl–puromycin dissociates from the ribosome. The polypeptide chain being synthesized (tagged with puromycin) is prematurely terminated, and thus incomplete and most likely biologically inactive.

Other Antibiotics  The structure and function of several other important antibiotic inhibitors of protein synthesis are given in Figure 12.11 and Table 12.4. Many of the antibacterial agents shown here have been used clinically for years to fight
bacterial infections, but the problem of antibiotic resistance is becoming increasingly serious. Bacteria can become resistant to antibiotics through genetic mutations, by transfer of resistance genes, by inactivation of the toxin by chemical modification, or by simply making membrane proteins that pump out unnatural chemicals (see Biochemistry in the Clinic 9).

Table 12.4
Antibiotic inhibitors of protein synthesis

<table>
<thead>
<tr>
<th>Antibiotic</th>
<th>Mode of Action</th>
</tr>
</thead>
<tbody>
<tr>
<td>Puromycin</td>
<td>Causes early termination by mimicking the action of an aminoacyl-tRNA; acts on prokaryotes and eukaryotes</td>
</tr>
<tr>
<td>Streptomycin</td>
<td>Causes misreading of mRNA and inhibits initiation; acts on prokaryotes</td>
</tr>
<tr>
<td>Tetracycline</td>
<td>Binds to the A site of ribosomes and blocks entry of aminoacyl-tRNAs; acts on prokaryotes</td>
</tr>
<tr>
<td>Erythromycin</td>
<td>Binds to ribosome and inhibits translocation; acts on prokaryotes</td>
</tr>
<tr>
<td>Chloramphenicol</td>
<td>Binds to 50S subunit and inhibits peptidyl transferase; acts on prokaryotes</td>
</tr>
<tr>
<td>Cycloheximide</td>
<td>Inhibits translocation of eukaryotic peptidyl-tRNA</td>
</tr>
<tr>
<td>Linezolid</td>
<td>Blocks formation of 70S initiation complex in prokaryotes</td>
</tr>
</tbody>
</table>
The synthetic antibiotic **linezolid (Zyvox)** is the first of a newly discovered class of protein synthesis inhibitors called the oxazolidinones. This class of compounds has a five-membered heterocyclic ring containing both nitrogen and oxygen atoms (Figure 12.11). Linezolid inhibits protein synthesis by blocking the interaction between mRNA and the two ribosomal subunits, thus preventing the formation of the 70S initiation complex (Table 12.4). The drug is active against essentially all Gram-positive bacteria including those that are vancomycin resistant. It is used especially for treatment of skin and soft tissue infections, pneumonia, and tuberculosis.

Figure 12.11 Structures of antibiotic inhibitors of prokaryotic protein synthesis.
When medical doctors encounter patients suffering from diseases caused by bacterial infections, the treatment often involves a 1- to 2-week course of a strong antibiotic such as tetracycline, erythromycin, vancomycin, or the new synthetic antibiotic linezolid. As you just read in the section on antibiotics, these compounds kill bacteria by inhibiting specific steps in the process of prokaryotic protein synthesis. But what should we do, if anything, about the many germs (some harmless and even beneficial) that are always present on our bodies and in our homes, but seldom cause serious problems?

Manufacturers of personal care products and household cleaning agents are adding antibacterial agents to limit the spread of these microorganisms. These germ-killing agents are now being incorporated into dishwashing detergents, hand creams, sponges, hand soaps, and cosmetics, and even toothpaste, toys, mattress pads, and food cutting boards.

The most widely used biocide is the chlorinated chemical triclosan (see structure). The action of triclosan and other biocides like alcohol and detergents has always been thought to be nonspecific—they simply kill all microorganisms by an extensive disruption of the cell membranes. If the biocides have no specific biological targets, then it could be assumed that bacteria could not develop a resistance to the chemical. This is different from the antibiotics that target specific biochemical events and thus allow bacteria to develop resistant strains by mutations. New studies on triclosan now show that its biocide action does involve a specific biological target. The chemical inhibits enoyl-ACP reductase, an important enzyme in bacterial fatty acid synthesis (see Section 18.3) in several microorganisms. In addition, the bacteria were able to acquire genetic resistance to triclosan by mutating the gene that encodes the reductase. Manufacturers of products containing triclosan and other antibacterial additives state that the U.S. government has approved the use of the chemicals and that they are safe.

Nonantibiotic Inhibitors Two interesting nonantibiotic inhibitors of eukaryotic protein synthesis are (1) ricin, a very toxic protein, present in castor bean seeds, that has been used as a biological warfare agent and by terrorists (see p. 352 and Biochemistry in the Clinic 12), and (2) diphtheria toxin, an enzyme secreted by Corynebacterium diphtheriae bacteria carrying the phage corynephage β.

For a nonclinical and more general use of antibacterial agents in household and personal care products, see Window on Biochemistry 12-2.

Post-translational Processing of Proteins

Learning Objective
Understand post-translational processing of proteins including folding, chemical modifications, and ubiquitin-initiated turnover.

Most polypeptide translation products are not yet in their biologically active form (native conformation) immediately after synthesis. Several steps of folding and biochemical modification are often required before a protein can perform its essential and specific function. In addition, most proteins (except those made in mitochondria and chloroplasts) are synthesized in the cytoplasm and they must be transported to other cellular regions where they serve their specific biological purpose(s).

Protein Folding

The first action for most proteins after synthesis is to fold into their native conformation. In some cases the protein molecules begin the folding process before synthesis is complete. As discussed in Section 4.3, the final tertiary structure of a protein is
determined by the primary structure (amino acid sequence). The protein folding process probably begins with the formation of local secondary structure (α-helix, β-conformation) to provide a nucleus or seed. The remainder of the protein chain then continues to fold around the initiation nucleus. The process often has the characteristics of cooperativity, indicating that each folding step facilitates the formation of other favorable interactions. The goal of protein folding is to form the maximum number of strong interactions (hydrophobic, hydrogen bonding, van der Waals, ionic) so a stable, three-dimensional arrangement can result. Many polypeptides receive assistance in the folding process. Proteins called chaperones act as catalysts to guide and facilitate folding (see Section 4.3). Some chaperones are enzymes that couple ATP hydrolysis to the protein folding process. The chaperones also bind to folding proteins in order to hide exposed hydrophobic amino acid residues so interactions do not occur out of order.

Most favorable interactions holding a protein in its native conformation are non-covalent in nature; however, covalent disulfide bonds (S—S) often cross-link cysteine residues that may be far apart in the sequence of the polypeptide chain. Experimental studies on the importance of disulfide bonds in tertiary structure show that they do not directly influence the folding of most proteins into their native conformations. Instead, disulfide bonds lock the protein into its final form after most other stabilizing, non-covalent interactions are in place. Some proteins are unable to fold completely into their native conformations until further biochemical modifications occur.

Biochemical Modifications

Many proteins must undergo covalent changes of their primary structures before they are biologically functional. Most of these biochemical changes involve chain cleavage processes, amino acid residue alterations, and addition of other factors such as carbohydrates and prosthetic groups.

Proteolytic Cleavage All prokaryotic proteins begin with the amino acid N-formyl-methionine; however, for about 50% of the proteins synthesized in E. coli, the N-terminal amino acid is removed by hydrolysis. In a similar fashion, the methionine residue at the N-terminus of eukaryotic proteins is removed. In addition, approximately one-half of eukaryotic proteins are modified by addition of an acetyl group to their amino end:

\[
\text{CH}_3\text{C—NH—}\]

In Chapter 6 the covalent modification of enzymes was described as a means to regulate activity. Some enzymes are initially constructed as inactive precursors, called zymogens, that are cleaved at one or more specific peptide bonds to produce the active form of the enzyme. Chymotrypsin is a protease that is regulated by such covalent modification (see Figure 6.8).

Amino Acid Modification The activity of enzymes and other proteins can be altered by biochemical modification on amino acid residues (see Chapters 3 and 6). The chemical processes of phosphorylation and hydroxylation are the most common. The hydroxyl groups of serine, threonine, and tyrosine residues are modified by transfer of a phosphoryl group, —PO_3^—, from ATP. The activity of the enzyme glycogen phosphorylase is regulated by phosphorylation of specific serine residues. Side chain hydroxylation of proline and lysine residues are common alterations in the structural protein collagen. Chemical modifications occur not on the free amino acid molecules but after they are incorporated into proteins.

Attachment of Carbohydrates Glycoproteins, those proteins that have covalently bonded carbohydrates, are involved in many biological functions including immunological protection, cell–cell recognition, and blood clotting. Carbohydrate
chains up to 15 residues long are covalently bonded to proteins at the hydroxyl groups of serine and threonine or on the side chain amide nitrogen of asparagine (see Section 7.5).

**Addition of Prosthetic Groups**  Many proteins depend on the presence of a covalently bound cofactor or prosthetic group for biological activity. During our studies in metabolism we will encounter proteins with prosthetic groups such as heme, FAD, biotin, and pantothenic acid. Recall that most cofactors and prosthetic groups are derived from vitamins.

**Protein Targeting**

Except for a few proteins made in mitochondria and chloroplasts, essentially all proteins are synthesized by ribosomes in the cytoplasm of cells. (Perhaps some are synthesized in the nucleus; see Window on Biochemistry 12-1.) However, proteins are needed not only in the cytoplasm but also in other cellular regions and organelles. Bacteria typically need proteins in four basic compartments: outer membrane, plasma (inner) membrane, periplasmic space (between membranes), and cytoplasm. Eukaryotic cells have several compartments needing proteins: membranes, mitochondria, chloroplasts, nucleus, lysosomes, and others. Each region, whether it is membrane or the nucleus, has the requirement for a unique set of proteins. How are proteins sorted and transported to their final destinations? These are the subjects of protein targeting.

In general, proteins that must be transported from the cytoplasm across plasma or vesicle membranes are synthesized with a short sequence of extra amino acid residues called the signal sequence. Except for those transported to the nucleus, all proteins that are to cross membranes are labeled with a **signal sequence** of 14 to 26 amino acids at the amino terminus. The signal sequence is usually removed by hydrolysis when the protein reaches its final destination. (For nuclear proteins such as DNA polymerase I, the signal sequence is internal and is not cleaved.) The general features of a prokaryotic (E. coli) signal sequence are shown in Figure 12.12. The label consists of (1) a short sequence of positively charged amino acids followed by (2) a hydrophobic region of 7 to 13 amino acids, (3) a nonhelical region containing proline or glycine, and (4) a cleavage site for peptidase action to remove the signal sequence.

Signal sequences for eukaryotic proteins are very similar to prokaryotic proteins; however, more complicated processes are necessary for transport to cellular organelles and other regions or for export to other cells. The best understood targeting system in eukaryotic cells involves the **endoplasmic reticulum** (ER). Proteins bound for membranes, lysosomes, or export are synthesized on ribosomes attached to the ER. Following synthesis, their signal sequences mark them for transport.

![Figure 12.12](image)

**Figure 12.12** General features of a prokaryotic signal sequence. The sequence consists of a basic region of positively charged amino acids, a hydrophobic region, and a nonhelical region. The cleavage site is for removal of the signal sequence by the action of a peptidase.
12.3 Post-translational Processing of Proteins

Post-translational Processing of Proteins

through the ER membrane. Inside the ER, the signal sequence is cut off and chemical modification on the proteins occurs, usually by addition of carbohydrates. Proteins are then moved from the ER to the Golgi apparatus, where they are sorted on the basis of carbohydrate markers, packaged in vesicles, and transported to their final destinations.

The Proteasome and Protein Degradation

Cellular and extracellular proteins are continuously being degraded and replaced by newly synthesized protein molecules. Although this process appears wasteful, it serves many purposes including removal of damaged or misfolded proteins and destruction of regulatory proteins that are not needed at the time. It also gives an organism the chance to adapt to changing conditions. The rate of turnover varies greatly from protein to protein. For example, two rat liver enzymes, RNA polymerase I and cytochrome c, have half-lives of 1.3 and 150 minutes, respectively. Human hemoglobin molecules may exist in erythrocytes for as long as 100 days.

Extracellular proteins taken up by cells are usually destroyed by the actions of lysosomal proteases. Degradation of unwanted intracellular proteins in prokaryotes and eukaryotes is carried out by ATP-dependent proteases associated with large protein complexes called the proteasomes. Proteins destined for destruction are marked in various ways. In eukaryotic cells, the ubiquitin pathway is important for protein labeling and degradation. Ubiquitin is a small protein of 76 amino acid residues found in all eukaryotic cells (Figure 12.13). Its amino acid sequence is highly conserved. For example, the sequence is identical in the ubiquitin of humans and fruit flies. By an ATP-dependent process, ubiquitin becomes covalently attached to unneeded or damaged proteins via an unusual peptide bond between the carboxyl terminus of ubiquitin and the ε-amino group of lysine residues in the targeted protein (Figure 12.14).

---

**Figure 12.13** Structure of the protein ubiquitin that marks defective proteins for degradation by proteasomes. (Structure (pdb1UBQ) determined by S. Vijay-Kumar, C. E. Bugg, and W. J. Cook.)

**Figure 12.14** Linkage of ubiquitin to a defective protein. Note the unusual amide bond (in red) to the ε-amino group of a lysine residue in the defective protein. Several ubiquitin molecules may be attached to a protein in order to mark it for destruction. Ubiquitin is a small polypeptide of 76 amino acid residues.
The labeled protein is then degraded by the proteolytic actions of the proteasome (Figure 12.15). The 2004 Nobel Prize in chemistry was awarded to two Israelis—Aaron Ciechanover and Avram Hershko—and one American—Irwin Rose—for their elucidation of ubiquitin-mediated protein degradation.

The intact proteasome complex in eukaryotic cells is the 26S proteasome. The barrel-shaped, 1700-kD complex is about 45 nm long and is composed of a 20S core plus two additional structures known as 19S caps (Figure 12.15). A cavity that has an opening diameter of about 1.3 nm runs down the inside center of the barrel. Entry to the central chamber, where proteolytic destruction occurs, is regulated by the 19S caps. Intracellular proteins that are labeled with ubiquitin are sought by the 19S caps, unfolded, and directed into the core. Once inside, the marked proteins are cleaved by proteolytic action. Several types of catalytic activity appear to take place including trypsin-like cleavage at basic amino acid residues, chymotrypsin-like action at aromatic acid residues, and cleavage at acidic amino acid residues. The initial peptide products range in size from 3 to 23 amino acid residues. These are further degraded to free amino acids after release from the proteasome.

Figure 12.15 Functioning of the 26S proteasome in protein degradation.

The labeled protein is then degraded by the proteolytic actions of the proteasome (Figure 12.15). The 2004 Nobel Prize in chemistry was awarded to two Israelis—Aaron Ciechanover and Avram Hershko—and one American—Irwin Rose—for their elucidation of ubiquitin-mediated protein degradation.

The intact proteasome complex in eukaryotic cells is the 26S proteasome. The barrel-shaped, 1700-kD complex is about 45 nm long and is composed of a 20S core plus two additional structures known as 19S caps (Figure 12.15). A cavity that has an opening diameter of about 1.3 nm runs down the inside center of the barrel. Entry to the central chamber, where proteolytic destruction occurs, is regulated by the 19S caps. Intracellular proteins that are labeled with ubiquitin are sought by the 19S caps, unfolded, and directed into the core. Once inside, the marked proteins are cleaved by proteolytic action. Several types of catalytic activity appear to take place including trypsin-like cleavage at basic amino acid residues, chymotrypsin-like action at aromatic acid residues, and cleavage at acidic amino acid residues. The initial peptide products range in size from 3 to 23 amino acid residues. These are further degraded to free amino acids after release from the proteasome.

Figure 12.15 Functioning of the 26S proteasome in protein degradation.

The labeled protein is then degraded by the proteolytic actions of the proteasome (Figure 12.15). The 2004 Nobel Prize in chemistry was awarded to two Israelis—Aaron Ciechanover and Avram Hershko—and one American—Irwin Rose—for their elucidation of ubiquitin-mediated protein degradation.

The intact proteasome complex in eukaryotic cells is the 26S proteasome. The barrel-shaped, 1700-kD complex is about 45 nm long and is composed of a 20S core plus two additional structures known as 19S caps (Figure 12.15). A cavity that has an opening diameter of about 1.3 nm runs down the inside center of the barrel. Entry to the central chamber, where proteolytic destruction occurs, is regulated by the 19S caps. Intracellular proteins that are labeled with ubiquitin are sought by the 19S caps, unfolded, and directed into the core. Once inside, the marked proteins are cleaved by proteolytic action. Several types of catalytic activity appear to take place including trypsin-like cleavage at basic amino acid residues, chymotrypsin-like action at aromatic acid residues, and cleavage at acidic amino acid residues. The initial peptide products range in size from 3 to 23 amino acid residues. These are further degraded to free amino acids after release from the proteasome.

12.4 Regulation of Protein Synthesis and Gene Expression

Learning Objective
Understand the molecular processes that regulate gene expression.

The typical bacterial cell (E. coli, for example) has about 4000 genes in its DNA genome that upon transcription and translation could potentially lead to a like number of polypeptide products. The situation in eukaryotic cells is even more overwhelming. With an estimated 20,000 to 25,000 genes in the human genome, there is the potential for hundreds of thousands of protein products in the human organism. But all proteins in prokaryotic or eukaryotic cells are not required at all times, so it would be a waste of energy and supplies to synthesize all proteins continuously. In fact, only a fraction of genes is expressed at any given time. Hence, very careful regulation of gene expression is required.

Some proteins and enzymes must be present in a cell at all times. Others are needed in only small amounts and at specific times. Some types of cells require specific proteins that are not present in all cells. The environments of some cells, especially prokaryotic, are constantly changing. Genes must be turned on and off during different development stages. The mechanisms by which organisms respond to these situations and balance protein synthesis and degradation are extremely complicated,
and we have much less than a complete understanding of regulatory processes. As with many other biological processes, our understanding of the regulation of gene expression is at a higher level for prokaryotes than for eukaryotes.

There are many steps from DNA to protein that can potentially be regulated (Figure 12.16). A cell can control the amount of protein by regulating (a) the rate of transcription, (b) the rate of post-transcriptional processing, (c) the rate of mRNA degradation, (d) the rate of protein synthesis (translation), (e) the rate of post-translational processing, and (f) the rate of protein degradation. Although examples of regulation at all of these stages are known, most gene expression is controlled at the level of transcription initiation. Control at this stage limits the number of mRNA molecules synthesized since the rate of synthesis of any protein is most closely related to the quantity of its corresponding mRNA. The quantity of mRNA is determined as the rate of mRNA synthesis minus the rate of mRNA breakdown. Here we will concentrate on the regulation of transcription initiation.

There are two fundamental types of gene expression:

1. Expression of constitutive genes refers to continuous transcription, resulting in a constant level of certain protein products. These housekeeping genes maintain a constant supply of their products for general cell maintenance and central metabolism.

2. Expression of inducible genes or repressible genes refers to those that can be activated (induced) to increase the level of mRNA and protein product or can be deactivated (repressed) to decrease the level of mRNA and protein. These genes are regulated by the action of RNA polymerase and molecular signals such as regulatory proteins, hormones, and metabolites.
Principles of Regulating Gene Expression

Several fundamental principles of regulating gene expression are presented here.

**Operons** In many prokaryotic cells, the genes for proteins that are related in function are clustered into units on the chromosome called **operons** (Figure 12.17). For example, the genes for enzymes for a single metabolic pathway would be expected in such a cluster. The components within these units are:

1. The genes to be transcribed and translated (called the structural genes).
2. The promoter region, which is responsible for RNA polymerase binding to the initiation site (see Section 11.4).
3. A binding site for activators.
4. A binding site for repressors (called an **operator**).

All structural genes are regulated by nucleotide sequences upstream from the start site (+1). The operon model, which was postulated by Jacques Monod and François Jacob in 1960, serves as a paradigm for prokaryotic gene regulation; however, other regulatory processes also are important.

**RNA Polymerase** The key participant in transcription is **RNA polymerase**. Recall that transcription is initiated when RNA polymerase binds to a promoter region on the DNA. The nucleotide sequences in promoter regions vary considerably, which changes the affinity of RNA polymerase binding. The number of mRNA molecules made is influenced by the affinity of RNA polymerase binding. For constitutive genes, RNA polymerase binding to promoter leads to transcription without further regulation. For inducible or repressible genes, other levels of control are superimposed onto the transcription process. Expression of these genes is controlled by regulatory proteins and other molecular signals. Regulation of eukaryotic gene expression is a more complicated process in several ways:

1. Complex sets of regulatory elements are present in promoter regions.
2. Three classes of RNA polymerases with different modes of regulation are present.
3. The DNA is much more complex in size and structure.

**Regulatory Proteins** RNA polymerase activity is mediated by **regulatory proteins**. Two major types of proteins that influence the action of RNA polymerase have been identified. Both types recognize and bind to specific DNA sequences on the chromosome that turns the polymerase “on” or “off”:

1. **Activators** are regulatory proteins that bind next to the promoter regions and assist the binding of RNA polymerase to the adjacent promoter (Figure 12.17).
   This increases the rate of gene transcription. Some eukaryotic activators bind to DNA regions called **enhancers** that may be some distance from the promoter.
2. **Repressors** are proteins that bind to specific base sequences in the promoter regions (called operators in prokaryotic cells). When bound, these proteins prevent the RNA polymerase from gaining access to the promoter; thus, transcription is blocked.
The primary mode of action of the regulatory proteins is to recognize and bind to specific base sequences on the chromosome. The proteins have discrete DNA binding domains within their polypeptide chains. The regulatory proteins can “read” DNA sequences and bind to specific regions. Additionally, some regulatory proteins also have protein binding domains that allow them to bind to RNA polymerase, other regulatory proteins, and with themselves to form dimers, trimers, and so on. Binding between protein molecules most likely involves hydrophobic interactions and hydrogen bonding. The action of a repressor or activator protein is influenced by other molecular signals such as small metabolites or hormones. The binding of these small molecules to regulatory proteins may increase or decrease the affinity of the protein–DNA interaction. Before we discuss the details of regulatory protein structure and function we will pause to consider the common types of gene regulation.

**Types of Regulatory Processes** Most regulatory processes can be classified into one of four mechanistic types (Figure 12.18). Positive regulation refers to the action of an activator on transcription. Two modes are possible. In the first type, transcription proceeds until a specific molecular signal binds to the activator, causing it to dissociate from the DNA (Figure 12.18a). In the second positive mode, only the activator–molecular signal complex binds to DNA. If the molecular signal dissociates, the activator falls off the DNA stopping transcription (Figure 12.18b). In negative regulation (Figure 12.18c), the repressor binds to DNA in the absence of molecular signal and mRNA synthesis is halted; binding of molecular signal by repressor causes dissociation from DNA and synthesis of mRNA proceeds. (d) Negative regulation: Repressor with bound molecular signal binds to DNA and mRNA synthesis is inhibited; dissociation of the molecular signal causes dissociation of the repressor from DNA and mRNA synthesis proceeds.
regulation, binding of repressor causes inhibition of transcription. Some repressors
dissociate from the DNA when a specific signal molecule is present (Figure 12.18c),
whereas other repressors require the binding of a signal molecule, which is necessary
for repressor binding and transcription inhibition (Figure 12.18d). All four types of
regulation are found in prokaryotic cells; positive regulation is an especially common
feature of eukaryotic gene expression. At the conclusion of this section several exam-
ples of some of these common modes of gene regulation are provided.

Binding of Regulatory Proteins Regulatory proteins have common structural features
and binding characteristics. Now that many regulatory proteins have been discovered
and studied, it is becoming clear that they have much in common. Regulatory proteins
have discrete binding domains that allow them to recognize and bind to specific DNA
sequences. The DNA binding domains on the proteins are relatively small, consisting
of 20 to 100 amino acid residues. How do the proteins interact with DNA? Molecular
recognition is the result of an exact fit between the surfaces of two molecules. Exact
fit implies favorable interactions that hold the molecules together. At the outside edge
of the DNA double helix is the major groove where the nucleotide bases are suffi-
ciently exposed for possible hydrogen bonding to proteins. Hydrogen bonding between
amino acid residues and nucleotide bases is possible without disrupting the base pair-
ing and without unwinding the double helix (Figure 12.19). The amino acid residues of
the regulatory proteins participate in hydrogen bonding to DNA through the side
chains of lysine, arginine, glutamate, asparagine, and glutamine. In each DNA–protein
complex many possible contacts lead to specific and relatively tight binding.

Three Classes of Regulatory Proteins The structures of many regulatory proteins have been elucidated, making it possible to search for similar characteristics. It has been discovered that about 80% of the currently known regulatory proteins can be classified into one of three classes based on the presence of common structural motifs (see Sections 4.2 and 4.3):

1. The helix–turn–helix motif
2. The zinc finger motif
3. The leucine zipper motif
The helix–turn–helix motif, which has a length of about 20 amino acid residues, consists of two short \( \alpha \)-helical regions of 7 to 9 amino acids connected by a \( \beta \) turn, often caused by the amino acid glycine. It is the most common DNA binding domain in prokaryotic regulatory proteins. One of the \( \alpha \)-helixes, the recognition helix, can bind to DNA by fitting snugly into the major groove (Figure 12.20). Two of the best known regulatory proteins with this structural motif are lac repressor and trp repressor.

A second structural motif common in regulatory proteins is the zinc finger motif. To date, this DNA binding domain has been found only in eukaryotic regulatory proteins. This interesting motif is a region composed of about 30 amino acids. The key features are four amino acid residues whose combined side chains form a binding site for a single \( \text{Zn}^{2+} \). Three families of zinc finger proteins have been discovered. They are distinguished by the zinc binding amino acid residues: (a) Cys-Cys-His-His, (b) Cys-Cys-Cys-Cys, and (c) Cys-Cys-His-Cys (Figure 12.21). The sulfur and/or nitrogen atoms in each of the four amino acid side chains coordinate to the zinc ion. Two well-studied zinc finger regulatory proteins are the transcription factor TFIIIA from *Xenopus laevis* (the African clawed toad) and glucocorticoid receptor protein. Proteins with zinc fingers probably bind to DNA in the major groove and wrap around the double helix axis (Figure 12.22).
The third structural motif found in regulatory proteins is the leucine zipper motif, which allows the proteins to interact with each other (Figure 12.23). The characteristic feature is an α-helix region of approximately 30 amino acids, with the amino acid leucine occurring about every seventh residue. The leucine residues protrude from one side of the protein. This structural motif allows two molecules of the regulatory protein to combine by forming hydrophobic interactions between the leucine-rich areas. The combined proteins with alternating leucine residues take on a “zipperlike” appearance. Hence, this structural motif serves to provide protein–protein interactions and functions in protein dimerization. Regulatory proteins with leucine zipper motifs often have a DNA binding domain, composed of an α-helix with a high content of lysine and arginine, that is adjacent to the protein binding region. Thus, leucine zipper proteins have the distinction of two possible types of interactions, one with the DNA and one with another regulatory protein molecule.

Examples of Gene Regulation

With some of the fundamental principles of gene regulation described, we can apply them to actual regulation systems. Several examples of regulatory mechanisms are listed in Table 12.5 and described next.

The lac Operon The lac operon regulatory system, discovered in 1960 by Monod and Jacob, was the first regulatory mechanism understood in detail. The structural genes in this operon code for three enzymes that are required for lactose metabolism. In the absence of substrate lactose in the growth medium for E. coli cells, the lac operon is repressed. The specific lac repressor molecule binds to the operator, thereby inhibiting transcription of the structural genes by RNA polymerase and slowing the synthesis of enzymes for lactose metabolism (see Figure 12.18c). When lactose is present, it binds to the lac repressor, thus causing its dissociation from the operator site. Under these conditions the lactose-metabolizing enzymes are synthesized. The lac operon is also regulated by glucose concentrations in the growth medium. Glucose is the preferred nutrient for E. coli growth and its presence represses the metabolism of other carbohydrates, including lactose. Glucose acts through the regulatory protein, catabolite activator protein (CAP). When lactose is present and
Regulation of Protein Synthesis and Gene Expression

When glucose is absent, CAP binds near the lac promoter and enhances transcription of lactose metabolism enzymes. When lactose and glucose both are present, CAP dissociates from the DNA and decreases the production of lactose metabolism enzymes. The action of glucose on CAP is mediated by the second messenger cyclic AMP (cAMP) (see Special Topics in Modern Biochemistry I).

The trp Operon

The trp operon regulatory system controls the synthesis of tryptophan in E. coli. The trp operon is composed of a regulatory region (promoter, operator, etc.) and five structural genes that code for enzymes needed to synthesize tryptophan from chorismate (see Section 19.2, Figure 19.10). If the amino acid tryptophan is present in the growth medium, the biosynthetic enzymes are not needed and transcription of the structural genes is repressed. Their regulation is mediated by Trp repressor. Tryptophan present in the growth medium binds to protein repressor forming a complex that associates with the operator region. The Trp–repressor complex binds to the operator region thus blocking the transcription of the five structural genes by RNA polymerase.

Steroid Hormone Response Elements

Steroid hormone control of metabolism in humans is mediated through a complex regulatory system. These hydrophobic hormones, such as estrogen, progesterone, and glucocorticoids, present in the bloodstream, are able to diffuse through the plasma membranes of target cells (see Section 8.4). In the cell nucleus, the hormones form complexes with regulatory proteins and bind to specific promoter regions in the DNA called hormone response elements (HRE). Binding of the hormone–regulatory protein complex to DNA may enhance or repress transcription depending on the hormone. The glucocorticoid receptor system is described in Table 12.5.

Toxic Metals and Metallothionein Gene Regulation

Organisms must be protected against the ubiquitous presence in the environment of toxic heavy metals such as mercury, lead, cadmium, zinc, and copper. Most organisms require for normal growth very low concentrations of some of the metals (see Section 1.2). However, they are very toxic in the concentrations often found in rivers, streams, and watershed areas. In eukaryotes, ingested metals are detoxified by the protein metallothionein (MT). MT is a low-molecular-weight, cysteine-rich protein that has a high affinity for binding metal ions. It is present in kidney cells, liver cells, and many other cell types. The water-soluble, complexed metal ions are readily removed from cells. MT is always present.

---

Table 12.5
Examples of gene regulation

<table>
<thead>
<tr>
<th>Regulatory System</th>
<th>Regulatory Protein and Structural Motif</th>
<th>Molecular Signal</th>
<th>Mechanism of Action</th>
</tr>
</thead>
<tbody>
<tr>
<td>lac operon (E. coli)</td>
<td>Tetramer of identical subunits, helix–turn–helix</td>
<td>Lactose or allolactose</td>
<td>Negative; signal causes dissociation of regulatory protein</td>
</tr>
<tr>
<td>lac operon (E. coli)</td>
<td>Catabolite activator protein, dimer of identical subunits, helix–turn–helix</td>
<td>Glucose and cAMP</td>
<td>Glucose presence represses lac genes; glucose absence stimulates lac genes</td>
</tr>
<tr>
<td>trp operon (E. coli)</td>
<td>Dimer of identical subunits, helix–turn–helix</td>
<td>Tryptophan</td>
<td>Negative; signal binds to repressor causing binding to operator; inhibits synthesis of tryptophan</td>
</tr>
<tr>
<td>Steroid hormone response elements (eukaryotes)</td>
<td>Glucocorticoid receptor protein, dimer, two zinc fingers</td>
<td>Glucocorticoid</td>
<td>Regulation of carbohydrate metabolism</td>
</tr>
<tr>
<td>Metal response elements (eukaryotes)</td>
<td>MTF-1 (fish), six zinc fingers</td>
<td>Heavy metal ions, Zn$^{2+}$, Hg$^{2+}$, etc.</td>
<td>Synthesis of metallothionein, a protein to complex toxic metal ions</td>
</tr>
</tbody>
</table>
in cells in relatively low concentrations, but the concentration of MT increases in response to an increase in metal ion concentration. The levels of MT in an organism are often used as a biomarker to evaluate the extent of heavy metal contamination in the local environment.

The MT gene promoter region is structurally complex. It consists of several elements including a TATA box, at least four regions called metal response elements (MRE), and one glucocorticoid response element. The MREs are designed to accept specific transcription factor proteins. These transcription factors must be complexed with metal ions in order to bind to the specific regions in DNA. This binding induces RNA polymerase to transcribe the structural gene for MT, thus producing increased MT.

The complex regulation of MT synthesis is used to an advantage in the treatment of Wilson’s disease. This is a rare genetic disorder characterized by excessive storage of copper in the body especially in the liver, kidney, and brain. One experimental drug is zinc acetate. The presence of zinc ions stimulates the synthesis of MT which can then bind and remove the excess copper ions. Copper ions bind to MT with 100,000 times the affinity of zinc so copper can readily displace bound zinc.

**Regulation of Gene Expression by RNA**

Up to this point in Section 12.4, the discussion has focused on the role of regulatory proteins in controlling gene expression and, hence the synthesis of proteins. In fact, until a few years ago, it was assumed by most biologists that only proteins act to regulate gene expression in response to cellular conditions. The standard method of gene suppression in prokaryotic cells involves regulatory proteins that sense and bind a small molecule such as a hormone or metabolite, and then bind to DNA or RNA to slow the synthesis of certain proteins. It has now been discovered that some forms of RNA are able to monitor the cellular concentrations of certain biomolecules and act to initiate necessary changes in gene expression (protein or enzyme concentrations). Two important types of regulation by RNA will be described.

**Regulation of Gene Expression by RNA Interference**

The discovery that RNA can silence or suppress the expression of certain genes began with a study of petunia color. In an attempt to deepen the purple color of the flower, Richard Jorgensen and colleagues introduced into petunia plants extra copies of the gene that is responsible for making the purple pigment. Instead of the expected enhanced color, many flowers appeared variegated and even white. The expression of the introduced gene as well as the similar endogenous gene for the pigment was somehow suppressed rather than enhanced. Similar experiments with other species, including the nematode *C. elegans*, showed the gene-silencing phenomenon to be general. Most surprising, however, was that even greater gene suppression could be observed by introduction of double-stranded, sense RNA (dsRNA, RNA with the same nucleotide sequence as the target mRNA). This phenomenon is now called RNA interference (RNAi) or post-transcriptional gene silencing (PTGS) in plants.

The biochemical mechanism of RNAi has been studied in *Drosophila* and is thought to be similar in most organisms (including humans; see Figure 12.24). The incorporated dsRNA (called the RNA trigger; same nucleotide sequence as the targeted mRNA) is first attacked by an enzyme called Dicer. This enzyme, an ATP-dependent endonuclease, catalyzes the cleavage of the long dsRNA (usually must be >300 nucleotide pairs) into smaller nucleotide fragments of 21–23 base pairs. These fragments, called small interfering RNAs (siRNAs), then assemble into endonuclease-containing complexes called RNA-induced silencing complexes (RISCs). (The Dicer step may be bypassed by incorporating synthetic siRNAs into a cell.) The duplex siRNAs are then unwound into single strands by helicase-type action. The complexes with single-strand RNA then bind to complementary regions of the target mRNA and cleave it to prevent further translation.
Regulation of Protein Synthesis and Gene Expression

The search is now on for practical applications of the RNAi effect. Researchers are finding this type of gene silencing to be especially useful to make functional gene knockouts in order to assess gene function. There is also hope that the RNAi phenomenon can be directed toward the design of sense RNA drugs to treat diseases caused by overexpressed or malfunctioning proteins.

**Regulation of Gene Expression by Ribozymes (Riboswitches)** Another example of RNA as a messenger to control gene expression comes from the discovery of ribozymes that act as metabolite-responsive genetic switches (called riboswitches). Recent studies in prokaryotic, archaea, and some eukaryotic cells have uncovered a new type of ribozyme (Section 6.6) that senses the presence of a small molecule such as a hormone or metabolite in the cell and acts to regulate the concentration of proteins that act on the small molecule. These unique regulatory ribozymes were first discovered from a study of vitamin synthesis in bacteria. When vitamins like B1 (thiamine) and B12 (cyanocobalamin) reach certain concentrations in the cell, they interact with a special RNA that then relays a message to slow down the translation machinery. In general, the section of the RNA that makes up the riboswitch precedes the coding parts of larger mRNA molecules. When a metabolite reaches a sufficient concentration in a cell, it binds to the ribozyme which is connected to the specific mRNA and slows production of proteins that synthesize the metabolite. In some cases, the ribozyme switches off protein synthesis by cleaving itself from the mRNA. There is some promise for therapeutic application if riboswitches can be designed to bind specific drugs that could then regulate translation of overexpressed proteins.
Biochemistry in the Clinic

Ricin—The Killer Protein

Ricin, a protein found in the seeds of the castor bean plant (Ricinus communis), is one of the most toxic natural substances known. The protein is easily isolated from the waste mash remaining after extraction of castor oil from the seeds. Ricin has received special attention recently because of its potential use as a bioweapon (see p. 352).

Ricin has several properties that make it a favored agent for bioterrorism: it is highly toxic; it is easily obtained; it is available in powder, mist, and pellet form so poisoning may be by injection, inhalation, ingestion, or skin or eye exposure; it is water soluble and very stable; and there is no known antidote. As little as 75 μg if injected, will kill an adult. Accidental poisoning by ingestion of the seeds is also possible. Chewing on six to eight seeds could be fatal to some adults, whereas a child may die from chewing on just one or two seeds.

Symptoms of Ricin Exposure

Symptoms of ricin poisoning depend on the route of exposure and on the amount.

- **Ingestion:** Flu-like symptoms develop within 6–8 hours including abdominal pain, vomiting, and bloody diarrhea. In a few days, severe dehydration, lowered blood pressure, seizures, and bloody urine are common.

- **Inhalation:** Initial symptoms within a few hours are difficulty breathing; fever, cough, nausea, and tightness in the chest; heavy sweating; and fluid buildup in the lungs (pulmonary edema).

- **Skin or eye exposure:** Redness and pain of the skin and eyes result.

If death occurs, it is usually within 36–72 hours. If the patient is still alive after 3–5 days, she or he will likely recover.

**Treatment**

No antidote currently is known for ricin poisoning. Medical care depends on the type of exposure and involves treating the symptoms: helping the patient breathe (if inhaled); providing intravenous fluids; supplying medications for seizure and low blood pressure; flushing the stomach with activated charcoal to increase removal of ricin (if ingested); thoroughly washing skin or eyes if exposed.

Ricin Structure and Mode of Action

Ricin is a heterodimeric protein consisting of an A chain of about 30,000 molecular weight (267 amino acid residues) and a B chain of about the same size (262 amino acid residues). The two chains are connected by a covalent disulfide bond (see figure). The A chain has 8 α-helices and 8 β-sheets. The A chain is the toxic subunit and has RNA

Before You Go On...

1. Predict how small metabolite molecules may be bound to sensor ribozymes in riboswitches. Do you expect such bonding to be covalent or noncovalent?
N-glycosidase activity that catalyzes the hydrolysis of a specific adenine base from the 28S ribosomal RNA. (The reaction for depurination is shown in Figure 11.14a.) The target adenine is in the GAGA loop. This causes the inactivation of the ribosome and inhibition of protein synthesis. The ricin B chain is a lectin (see Window on Biochemistry 7-3) that recognizes and binds to galactose or N-acetylgalactosamine on the outside membrane of the target cells. The interaction facilitates the transport of the protein through the membrane and into the cell where the A chain acts on 28S RNA. It has been estimated that just a single ricin molecule that enters the cell can inactivate over 1500 of the 28S ribosomes per minute and kill the cell. Ricin is also currently being studied clinically as an anticancer agent and for use in bone marrow transplants.

Study Questions
1. Write the reaction catalyzed by the A chain of ricin. Use the general structure shown below to begin, but fill in all atoms for the substrate and products.

\[
\begin{align*}
\text{To 5'} & \\
\text{Ribose—adenine + H}_2\text{O} & \xrightarrow{\text{ricin A}} & \text{To 3'}
\end{align*}
\]

2. What properties of ricin would lead researchers to test it as an anticancer drug?

References
http://www.ansci.cornell.edu/plants/toxicagents/ricin/ricin.html
An informative website on ricin developed by the Animal Science Department at Cornell University.
Website on ricin from the U.S. Centers for Disease Control and Prevention.

■ Regulation of gene expression is controlled at the level of transcription initiation. Regulatory proteins recognize and bind to specific base sequences on the chromosomes, which turns RNA polymerases “on or off.”
■ Regulatory proteins have common structural features and are classified accordingly: helix–turn–helix proteins, zinc finger proteins, and leucine zipper proteins.

STUDY EXERCISES

Understanding Terms
12.1 Define the following terms in 25 words or less.

a. Translation
b. Codon
c. Aminoacyl-adenylate intermediates
d. 70S initiation complex
e. Polyrribosome
f. Protein targeting
g. Ubiquitin
h. Regulatory proteins
i. Operon
j. Zinc fingers
k. Repressors
l. Anticodon
m. Antibiotics
n. Peptidyl transferase
o. lac operon
p. Post-translational processing

Reviewing Concepts
12.2 Write the sequence of amino acids in a polypeptide translated from the following mRNAs.

a. 5' UUUCUAGAUAGAGUU
b. 5' GGAGGAGUAAGUUGU

12.3 Write out the sequence of bases in mRNA that encodes the message for the peptide His-Asn-Pro. Is there more than one possible mRNA sequence that answers this question? Explain.
12.4 Compare the following characteristics for the DNA binding domain in helix–turn–helix regulatory proteins and zinc fingers.
   a. Amino acid composition
   b. Number of amino acid residues
   c. Sequence of amino acids
   d. Presence of α-helices, β-sheets, turns
   e. Presence of metal ions

12.5 Differentiate between the A and P binding sites on the ribosome–mRNA complex.

12.6 Assume that a mRNA molecule has the following sequence:

   5' AUGCUCAUCUCGGAGAAGC

   a. What polypeptide sequence would be translated from this mRNA?
   b. Assume that nucleotide 4 (C) is deleted from the mRNA. What polypeptide sequence would now be translated from this modified mRNA?

12.7 The word “translation” describes which of the following steps in the flow of genetic information?
   a. DNA → RNA
   b. DNA → DNA
   c. RNA → RNA
   d. RNA → protein

12.8 Briefly describe the function of each of the following in protein synthesis.
   a. Ribosomes
   b. Codon
   c. Anticodon
   d. Aminoacyl-tRNA synthetases
   e. Peptidyl transferase

12.9 Which of the following statements about protein synthesis are true?
   a. Protein synthesis occurs by adding amino acids to the amino terminus of the growing polypeptide chain.
   b. The first amino acid to be incorporated into a protein is usually N-formylmethionine.
   c. Amino acids are selected and activated by a special form of tRNA.
   d. Peptidyl transferase is a ribozyme.
   e. Peptidyl transferase catalyzes the formation of the peptide bond.

12.10 Complete the following questions by identifying characteristics of the aminoacyl-tRNA synthetase–catalyzed reaction.
   a. Source of energy (ATP or GTP)?
   b. Fate of ATP molecule?
   c. Enzyme-bound intermediate?

12.11 Differentiate between constitutive and inducible genes.

12.12 Which of the following features are characteristic of the translation process?
   a. Occurs at the surfaces of mitochondria.
   b. The first amino acid incorporated is the amino terminus.
   c. Amino acids are activated by linkage to cyclic AMP.
   d. Aminoacyl-tRNA synthetases catalyze activation of amino acids.
   e. The process occurs on ribosomal particles.

12.13 Which of the following biochemical processes are part of post-translational modification events?
   a. Modification of amino acid residues
   b. Addition of FAD and other prosthetic groups to proteins
   c. Formation of disulfide bonds
   d. Proteolytic cleavage

12.14 Many types of molecules and molecular associations are required for the translation process. Circle those listed below that are not directly involved in translation.
   a. mRNA
   b. Aminoacyl-tRNA
   c. DNA
   d. Peptidyl transferase
   e. DNA polymerase III
   f. Ribosomes

12.15 Compare and contrast the structures and functional characteristics of the two types of regulatory proteins, activators and repressors.

**Working Problems**

12.16 Approximately how many phosphoanhydride bonds must be hydrolyzed during translation to synthesize a 300-amino acid protein? Assume that you are starting with a functional mRNA, ribosomal subunits, tRNA, and free amino acids.

12.17 Recall from Chapter 11 that mutations are caused by base changes in DNA. Which of the following amino acid changes are caused by a single base mutation?
   a. Ser → Pro
   b. Val → Ser
   c. Leu → Phe
   d. Ser → Phe
   e. Lys → Glu
   f. Gly → Ala

12.18 Show how the amino acid residues asparagine and glutamine in regulatory proteins could hydrogen bond to the thymine:adenine pair in DNA without disrupting complementary base pairing.

12.19 Write the reaction catalyzed by each of the following enzymes and describe physical properties of the enzyme.
   a. Aminoacyl-tRNA synthetases
   b. Peptidyl transferase
   c. RNA polymerase

12.20 Could the 20 amino acids be incorporated into proteins using a genetic code consisting of one nucleotide base = one amino acid? How many different types of amino acids could be coded with this ratio?

12.21 Explain how metal ions are bound in the protein metallothionein. What atoms in cysteine residues complex with metal ions?

12.22 Name the type of bond that links the following molecules.
   a. Amino acid to amino acid in proteins
   b. Nucleotide to nucleotide in DNA
   c. Amino acid to tRNA
   d. Nucleotide to nucleotide in RNA
   e. Codon in mRNA to anticodon in aminoacyl-tRNA

12.23 What is the advantage to the cell of having polyribosomes working on a single mRNA molecule rather than just one ribosome?
12.24 The amino acids hydroxylysine and hydroxyproline are not incorporated into a protein using the genetic code and therefore there is no triplet code listed for them in Table 12.2. Explain how these amino acids become present in proteins.

12.25 Could the 20 amino acids be incorporated into proteins using a genetic code consisting of two nucleotide bases = one amino acid? Explain.

12.26 Complete the following biochemical reactions of amino acid residue modification that are described in this chapter.

\[ \text{O} \quad \text{NH} \quad \text{CH} \quad \text{C} \quad \text{ATP} \]
\[ \overset{k\text{inase}}{\text{CH}_2} \quad \text{OH} \]

12.27 Describe the kind of chemical bonding that holds together a codon with its anticodon. **Hint:** Is it covalent, noncovalent, or both?

12.28 Circle those amino acids in the list below that are altered by biochemical reactions when they are present as residues in proteins.

Ala     Pro
Tyr     Ser
Lys     Thr
Val

12.29 Describe the kind of chemical bond that binds ubiquitin to proteins targeted for destruction.

12.30 Describe the action of the trp operon in 25 words or less.

12.31 The list below contains several terms that describe the characteristics of the genetic code. Select all of the terms that are correct.

a. Different for each species
b. Universal
c. Doublet (two nucleotide bases per amino acid)
d. Triplet (three nucleotide bases per amino acid)
e. Overlapping
f. Nonoverlapping
g. Punctuated
h. No punctuation
i. Degenerate

**Writing Biochemistry**

12.32 Your biochemistry instructor has asked each student in the class to prepare and hand in two essay questions on the material covered in Chapter 12. Some of the questions submitted will be selected by your instructor for use on the next hour exam. The only rule is that the questions should require a 25- to 50-word answer. Prepare two questions for submission and also write answers to the questions for your instructor to use as key.

12.33 Select one of the antibiotics in Figure 12.11 for study. Use an advanced biochemistry textbook or a literature reference to learn all you can about the compound. Then, write a 200-word review of the antibiotic. Your writing should include a discussion of the source of the antibiotic and how it functions to control the growth of microorganisms.

12.34 Protein synthesis is controlled by the interactions of regulatory proteins with specific regions on DNA. The most important types of bonding between the proteins and DNA include hydrogen bonding, hydrophobic interactions, and ionic interactions.

a. Describe these three types of bonding and show how amino acid residues in proteins can interact with nucleotides in DNA.
b. Interactions between regulatory proteins and DNA are said to be reversible and noncovalent. Explain why both of these characteristics are essential in regulatory protein–DNA interactions.

12.35 Large amounts of energy are required for protein synthesis. The energy from the cleavage of four phosphoanhydride bonds in ATP and GTP are necessary. Outline the steps of protein synthesis to show how the energy from ATP and GTP is used.

**FURTHER READING**


Website for Ambion, The RNA Company. Read technical bulletin, “RNA Interference and Gene Silencing—History and Overview.”
Website from the University of Leeds, UK. See which antibacterial agents are inhibitors of protein synthesis?
Website for the Pfizer drug, linezolid (Zyvox).