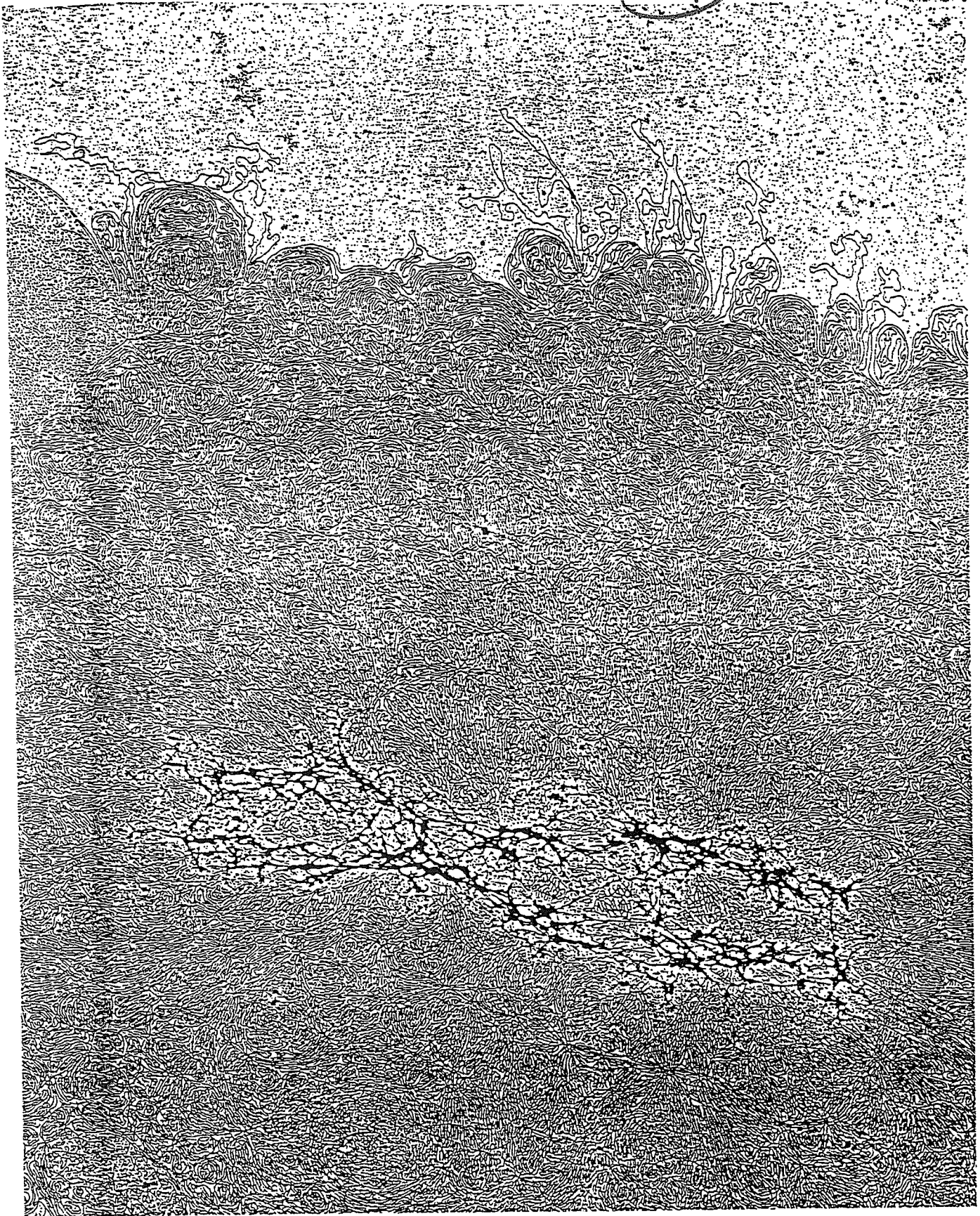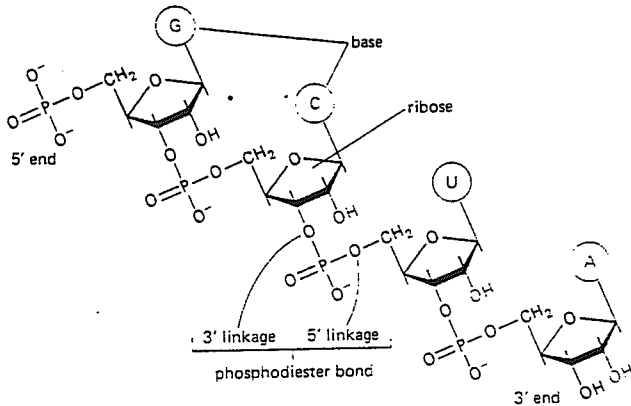*f... .R Alb..k of al M.l....b... B:.l... of the Cell*
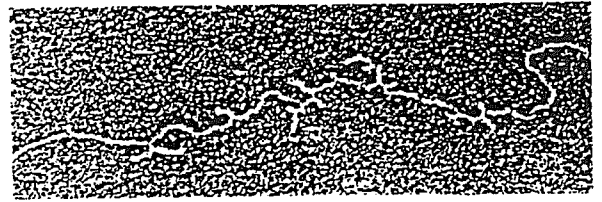
#13

Electron micrograph of DNA spilling out of a single disrupted human chromosome. Only about half of the DNA in one chromosome is shown. (Courtesy of James Paulson and Ulrich Laemmli.)
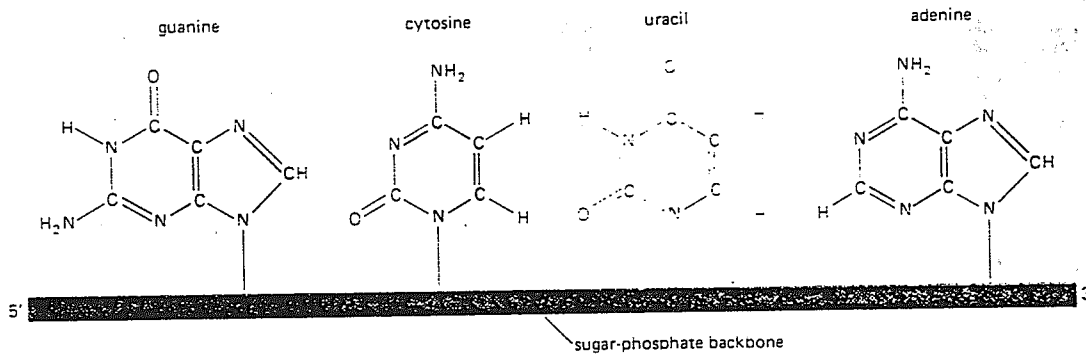
213

SUGAR-PHOSPHATE BACKBONE OF RNA

base

ribose

5' end

3' linkage   5' linkage

phosphodiester bond

3' end

ELECTRON MICROGRAPH OF RNA

FOUR BASES OF RNA

guanine   cytosine   uracil   adenine

5'   sugar-phosphate backbone   3'

RNA SINGLE STRAND
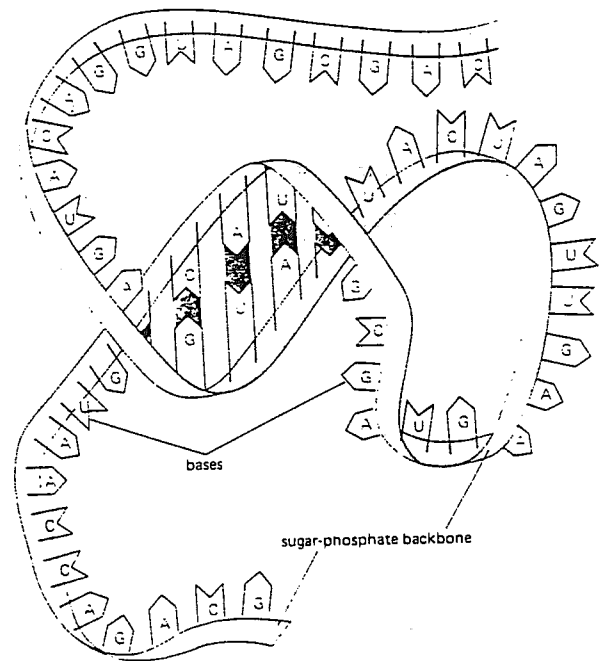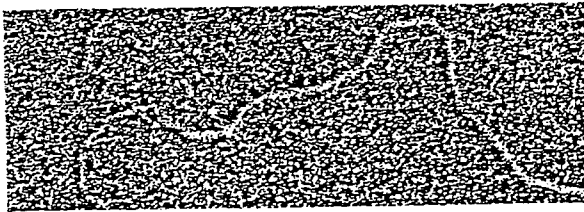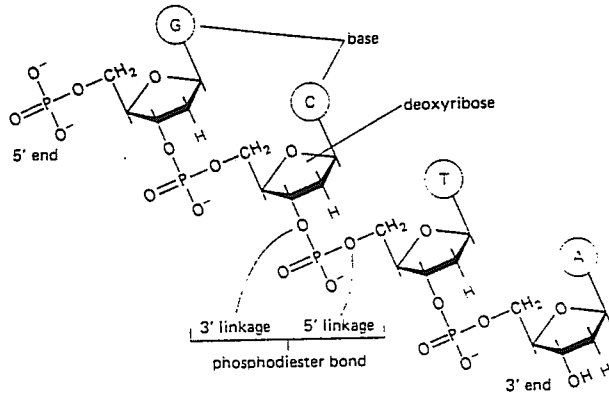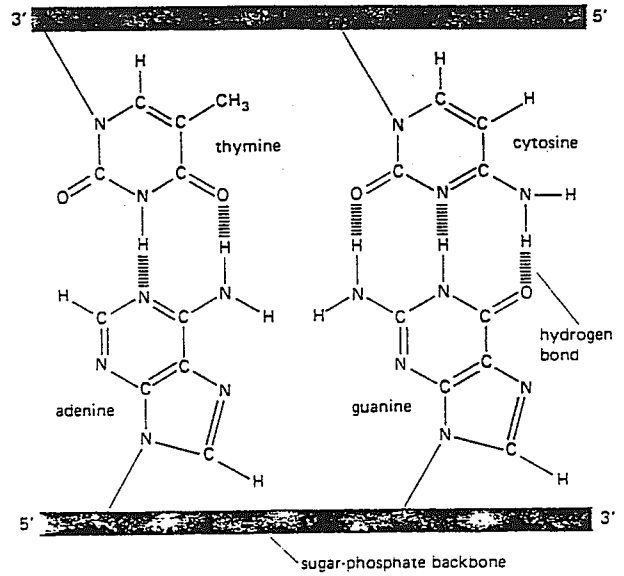
bases

sugar-phosphate backbone

Figure 3–6 The structure of DNA and RNA. Both DNA and RNA are linear polymers of nucleotides (see Panel H, pp. 60–61). DNA differs from RNA in two ways: (1) the sugar-phosphate backbone contains deoxyribose rather than ribose, and (2) it contains the base thymine (T) instead of uracil (U). Specific hydrogen bonding between G and C and between A and T (A and U in RNA) generates complementary base-pairing. In a DNA molecule two antiparallel strands that are complementary in their nucleotide sequence are paired in a right-handed double helix, forming about 10 base pairs per helical turn. RNA is single-stranded, but it contains local regions of short complementary base-pairing that can form from a random matching process. The conformation of these double-helical regions in RNA is shown in Figure 3–7. (Micrographs courtesy of Mei Lie Wong (DNA) and Peter Wellauer (RNA)).

SUGAR-PHOSPHATE BACKBONE OF DNA

base

deoxyribose

5' end

3' linkage   5' linkage

phosphodiester bond

3' end

OH H

FOUR BASES AS BASE PAIRS OF DNA

3'   5'

thymine

cytosine

adenine

guanine

hydrogen bond

5'   3'

sugar-phosphate backbone

ELECTRON MICROGRAPH OF DNA

DNA DOUBLE HELIX

sugar-phosphate backbone

5'

5'

3'

hydrogen bond

base

3'

— 1 helical turn = 3.4 nm —

minor groove   major groove
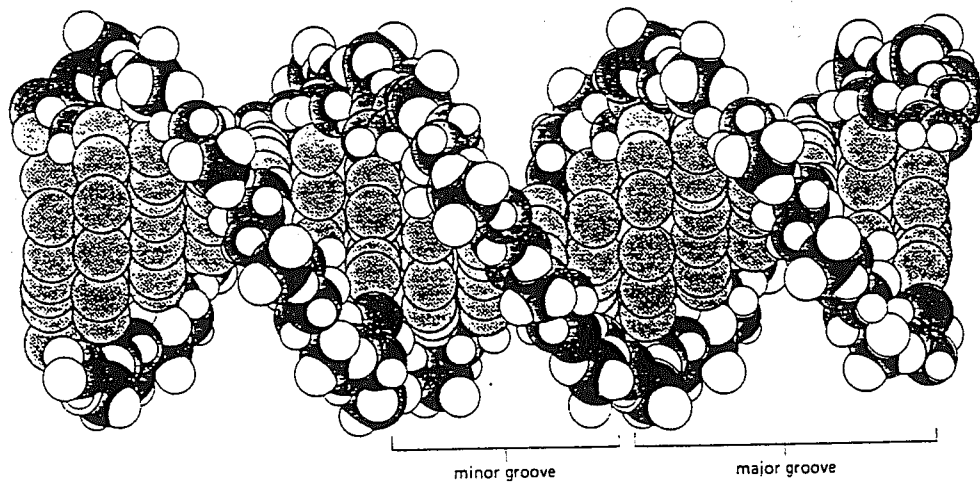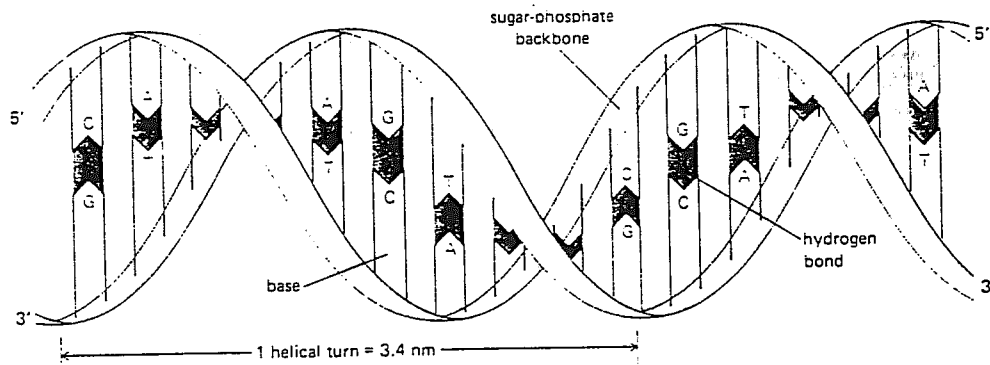
Basic Genetic Mechanisms

The capacity of cells to maintain a high degree of order in a chaotic universe stems from the genetic information that is expressed, maintained, replicated, and occasionally improved by four basic genetic processes—*protein synthesis, DNA repair, DNA replication,* and *genetic recombination.* These processes, which produce and maintain the proteins and the nucleic acids of a cell, are one-dimensional: in each of them, the information in a linear sequence of nucleotides is used to produce or alter either another linear chain of nucleotides (a DNA or RNA molecule) or a linear chain of amino acids (a protein molecule). Genetic events are therefore conceptually simple compared with most other cellular processes, which involve the expression of information contained in the three-dimensional surfaces of protein molecules. Perhaps that is why we understand genetic mechanisms in far greater detail than most other cellular events.

We shall begin with a detailed discussion of some of the mechanisms of protein synthesis that were introduced in Chapter 3. Then the mechanisms involved in the maintenance of DNA sequence integrity by DNA repair and its propagation by DNA replication are described. This leads to a discussion of viruses as self-replicating genetic entities that parasitize the genetic machinery of the cell, in which either DNA or RNA molecules serve as the primary repository of genetic information. The chapter concludes with a discussion of the mechanisms of genetic recombination in which DNA sequences are reassorted, a process of great importance to species adaptation in a changing environment.

Protein Synthesis

The Decoding of DNA into Protein[1]

Proteins generally constitute somewhat more than half of the total dry mass of a cell, and their synthesis is central to cell maintenance, growth, and development. Although protein synthesis is usually considered to start with the copying of DNA into messenger RNA (mRNA), a number of other preparatory steps are also required: a specific transfer RNA (tRNA) molecule must be at-

216

tached to each of the 20 common amino acids, and ribosomal subunits must be preloaded with auxiliary molecules. In the process of protein synthesis, all of these components are brought together in the cell cytoplasm in a ribosome complex. Here a single mRNA molecule is moved stepwise through a *ribosome* so that its sequence of nucleotides can be translated into a corresponding sequence of amino acids to create a distinctive protein chain.

RNA Polymerase Copies DNA into RNA[2]

The synthesis of an RNA copy of the nucleotide sequence in a limited region of DNA is catalyzed by the enzyme **RNA polymerase.** This process, called **DNA transcription,** is crucial for the transfer of information from DNA to protein. The RNA polymerase has been most thoroughly studied in procaryotes, where a single species of the enzyme mediates all RNA synthesis.

RNA polymerase initiates the transcription process after binding to a specific DNA sequence, called the **promoter,** that signals where RNA synthesis should begin. After binding to the promoter, the RNA polymerase unwinds about one turn of the DNA helix to expose a short stretch of single-stranded DNA that will act as a template for complementary base-pairing with incoming ribonucleotides. It then joins two of the incoming ribonucleoside triphosphate monomers together to begin an RNA chain. The RNA polymerase molecule then moves along the DNA template strand, extending the growing RNA chain in the 5'-to-3' direction by one nucleotide at a time (Figure 5–1). The enzyme

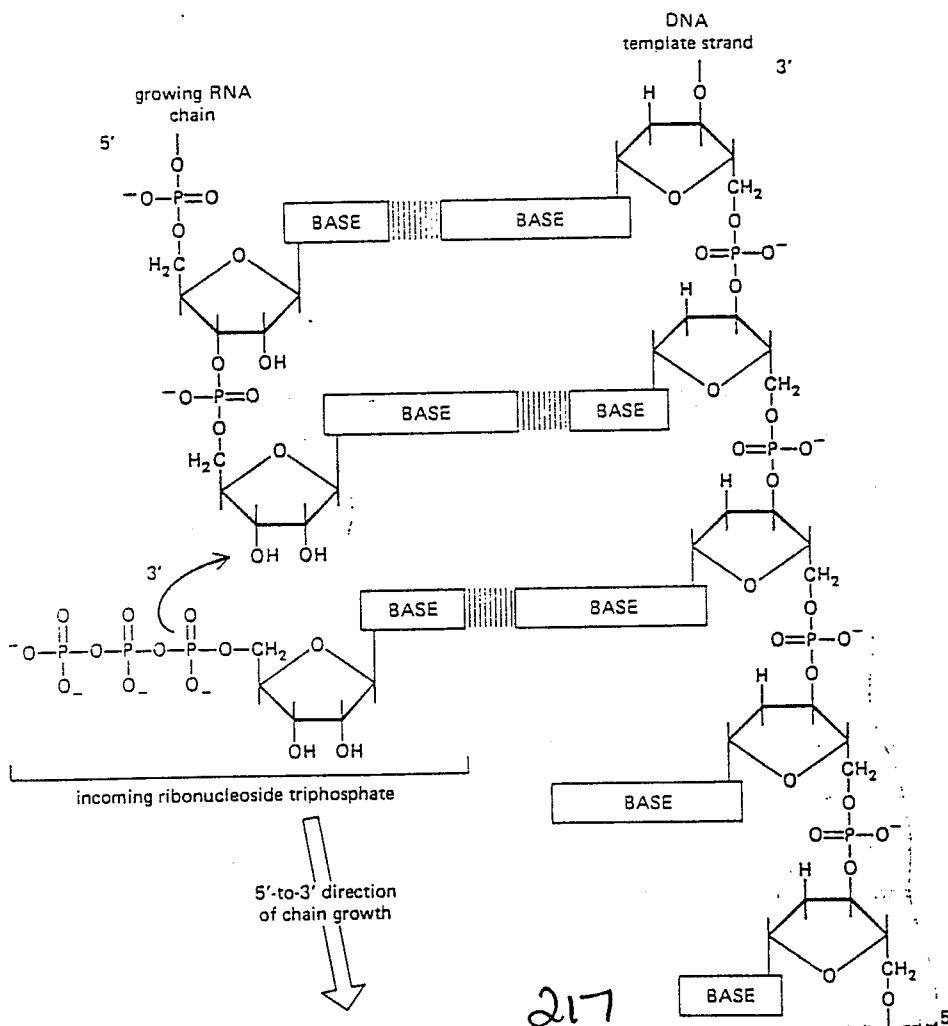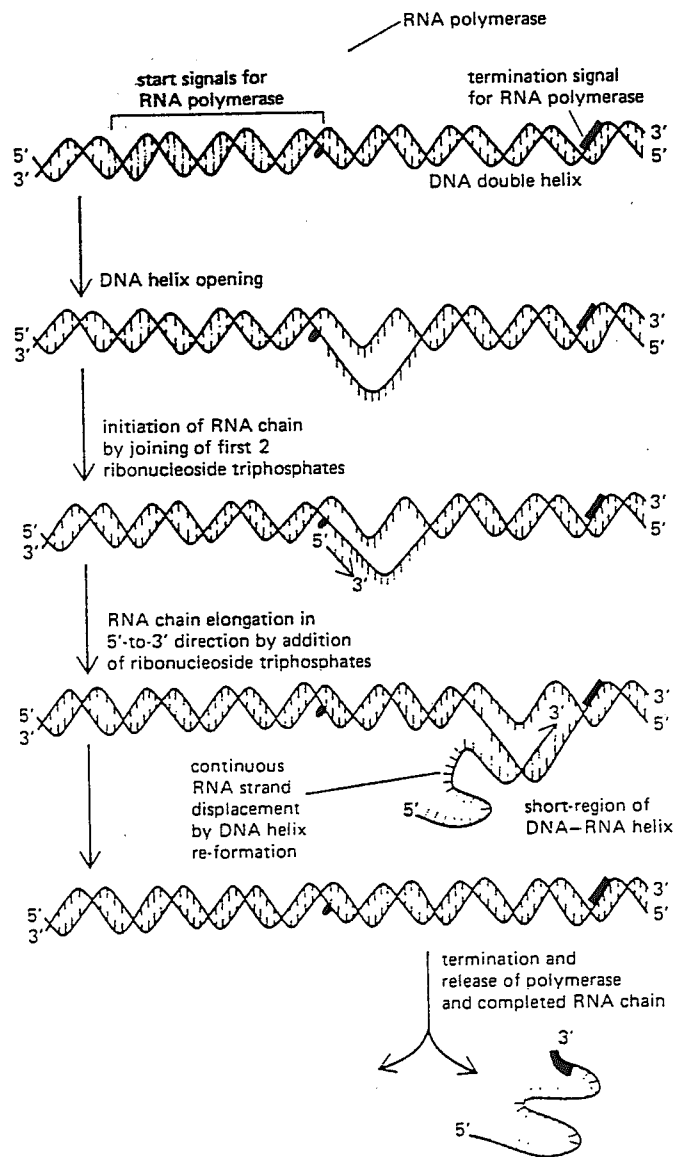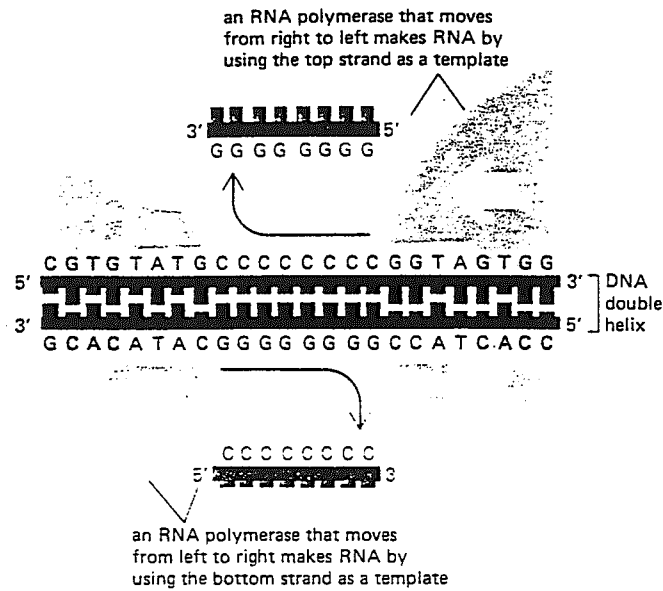**Figure 5–1** The reaction catalyzed by an RNA polymerase enzyme. In each step, an incoming ribonucleoside triphosphate is selected for its ability to base-pair with the exposed DNA template strand; a ribonucleoside monophosphate is then added to the growing 3'-OH end of the RNA chain *(colored arrow).* The new RNA chain therefore grows in the 5'-to-3' direction and is complementary in sequence to the DNA template strand.

RNA polymerase

start signals for
RNA polymerase

termination signal
for RNA polymerase

5'
3'

3'
5'

DNA double helix

DNA helix opening

5'
3'

3'
5'

initiation of RNA chain
by joining of first 2
ribonucleoside triphosphates

5'
3'

3'
5'

5'

3'

RNA chain elongation in
5'-to-3' direction by addition
of ribonucleoside triphosphates

5'
3'

3'
5'

continuous
RNA strand
displacement
by DNA helix
re-formation

3'

5'

short-region of
DNA–RNA helix

5'
3'

3'
5'

termination and
release of polymerase
and completed RNA chain

3'

5'

Figure 5–2 A diagram illustrating how RNA polymerase begins its synthesis at a special start signal on the DNA called a promoter and completes its synthesis at a special termination signal, whereupon both the polymerase and its completed RNA chain are released. Polymerization rates average about 30 nucleotides per second at 37°C. Therefore, an RNA chain of 5000 nucleotides takes about three minutes to complete. The RNA polymerase and the DNA helix are drawn approximately to scale.

continues to add nucleotides until it encounters a second special sequence in the DNA, the **termination signal**, at which point the polymerase releases both the DNA template and the newly made RNA chain. As the enzyme moves, a short RNA-DNA double helix is formed, but this is less stable than the DNA-DNA helix, which soon reestablishes itself, displacing the RNA. As a result, each completed RNA chain is released from the DNA template as a free, single-stranded RNA molecule (Figure 5–2).

In principle, any region of DNA could be copied into two different mRNA molecules—one from each of the two DNA strands. In fact, only one DNA strand is copied, although which it is can vary between neighboring genes. The "signpost" that indicates the strand to be copied is the promoter DNA sequence (the start signal). The promoter is oriented in such a way that it sets the RNA polymerase off in a particular direction across a given genetic region, and this automatically determines which of the two strands will be read (Figure 5–3).

an RNA polymerase that moves
from right to left makes RNA by
using the top strand as a template

3' ▊▊▊▊▊▊▊▊ 5'
G G G G G G G G

5' CGTGTATGCCCCCCCCGGTAGTGG 3' ⎤ DNA
                                      ⎥ double
3' GCACATACGGGGGGGGGCCATCACC 5' ⎦ helix

C C C C C C C
5' ▊▊▊▊▊▊▊ 3'

an RNA polymerase that moves
from left to right makes RNA by
using the bottom strand as a template

Figure 5–3 Because the DNA strand serving as template must be traversed from its 3' to its 5' end (see Figure 5–1), the direction of RNA polymerase movement determines which of the two DNA strands will serve as a template for the synthesis of RNA. The direction of polymerase movement is, in turn, determined by the orientation of the promoter sequence at which the RNA polymerase starts. In general, only one of the two DNA strands will be transcribed in any one region of the DNA double helix. Note that each RNA molecule made will be identical in its polarity and nucleotide sequence (except for the substitution of U for T) to the DNA strand that is paired to the strand used as template.

The DNA sequences that serve as promoters for the *E. coli* RNA polymerase have been well characterized. Their common features reveal that the polymerase recognizes two DNA sequences 6 nucleotides long that are separated from each other by about 25 nucleotides of unrecognized DNA (Figure 5–4). The DNA sequences that create a termination signal for this polymerase also share common features, as described in Figure 5–4. The eucaryotic cell contains three different RNA polymerases; they are basically similar to the bacterial enzyme, but their start and stop signals are different and they are less well characterized (see Chapter 8, p. 407).
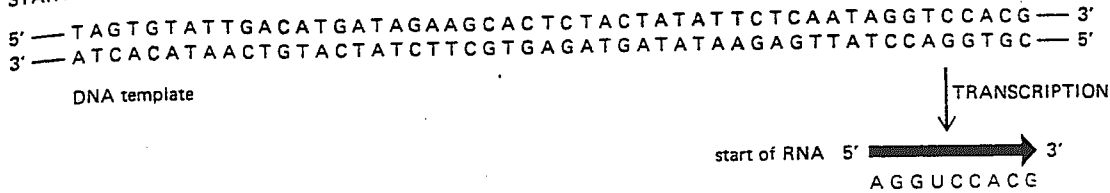
The above outline omits many details: in many cases, a number of other complex steps must occur before an mRNA molecule is produced. Thus, *gene-regulatory proteins* help to determine which regions of DNA are transcribed by the RNA polymerase and thereby play a major part in determining the proteins made by a cell. Moreover, although mRNA molecules are produced directly by DNA transcription in procaryotes, in higher eucaryotic cells most RNA transcripts are altered extensively—by a process called *RNA processing*—before they leave the cell nucleus and enter the cytoplasm to become mRNA. All of these specialized aspects of mRNA production will be discussed in detail in Chapter 8, when we consider the cell nucleus. For the present, let us assume that functional mRNA molecules have been produced by a cell and proceed to examine how they direct protein synthesis.

Protein Synthesis Is Inherently Very Complex[3]

The molecular processes that underlie protein synthesis are very complex. Although we can describe many of them, they do not make conceptual sense in the way that DNA transcription, DNA repair, and DNA replication do. For example, we now know that not one but three main classes of RNA molecules (mRNA, tRNA, and rRNA) are involved in protein synthesis, but we do not fully understand why this must be so. Thus, the details of protein synthesis must largely be learned as fact without an obvious conceptual framework.

A true understanding of the mechanism of protein synthesis should help to illuminate the early events by which life itself came into existence. Here

) START SIGNAL

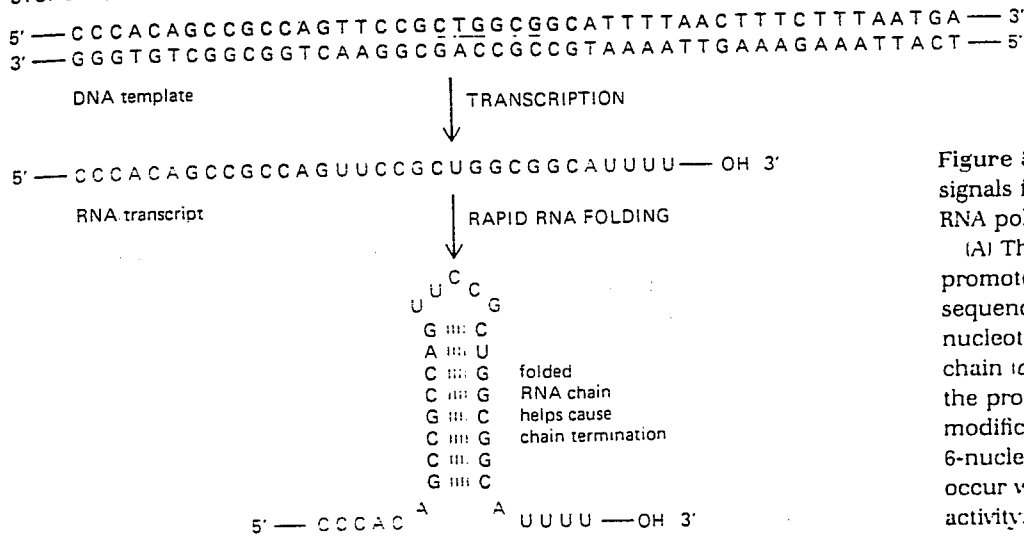5' —— T A G T G T A T T G A C A T G A T A G A A G C A C T C T A C T A T A T T C T C A A T A G G T C C A C G —— 3'
3' —— A T C A C A T A A C T G T A C T A T C T T C G T G A G A T G A T A T A A G A G T T A T C C A G G T G C —— 5'

DNA template

TRANSCRIPTION

start of RNA   5' ▬▬▬▬▶ 3'

A G G U C C A C G

3) STOP SIGNAL

5' —— C C C A C A G C C G C C A G T T C C G C T G G C G G C A T T T T A A C T T T C T T T A A T G A —— 3'
3' —— G G G T G T C G G C G G T C A A G G C G A C C G C C G T A A A A T T G A A A G A A A T T A C T —— 5'

DNA template

TRANSCRIPTION

5' —— C C C A C A G C C G C C A G U U C C G C U G G C G G C A U U U U —— OH 3'

RNA transcript

RAPID RNA FOLDING

```
              C
          U U   C
        U         G
        G ⅲⅲ C
        A ⅲⅈ U
        C ⅲⅲ G    folded
        C ⅲⅲ G    RNA chain
        G ⅲⅲ C    helps cause
        C ⅲⅲ G    chain termination
        C ⅲⅰ G
        G ⅲⅲ C
          A       A
5' —— C C C A C       U U U U —— OH 3'
```

Figure 5–4   The start and stop signals for RNA synthesis by bacterial RNA polymerase of *E. coli.*

(A) The polymerase starts at a promoter sequence. The two short sequences about −35 and −10 nucleotides from the start of the RNA chain (color) are thought to define the promoter. Only limited modification of each of the two 6-nucleotide sequences shown can occur without eliminating promoter activity. The rest of the DNA sequence shown is not thought to be important.
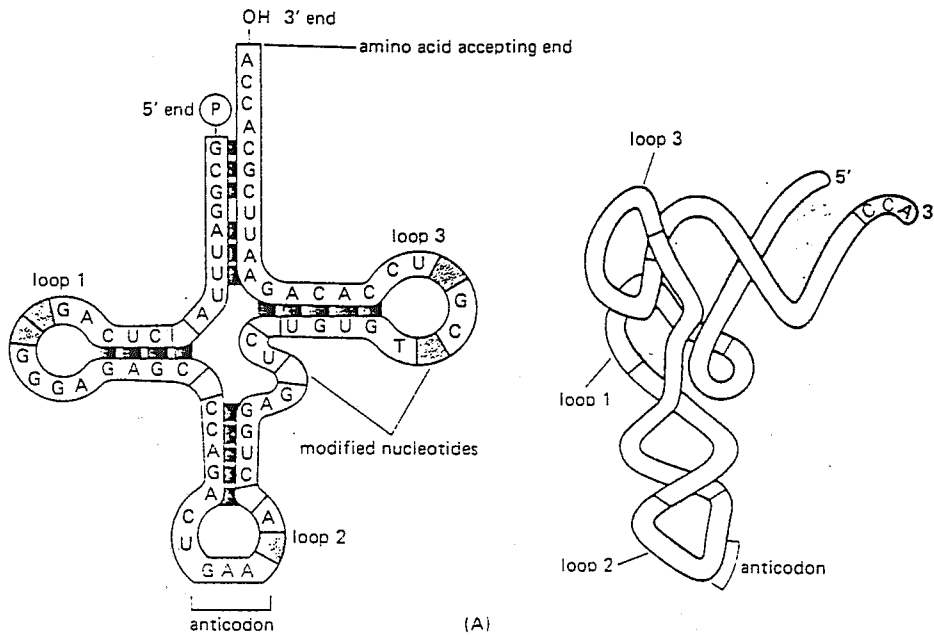
(B) The polymerase stops when it synthesizes a run of U residues on the template that is located immediately following a self-complementary nucleotide sequence. The appropriate combination of DNA sequences therefore constitutes a termination signal. As indicated, the self-complementary sequence (gray) will rapidly form a hairpin helix in the newly synthesized RNA chain. Many different self-complementary sequences have been shown to function in this way; therefore, the hairpin helix rather than the nucleotide sequence is what seems to be required to create this part of the termination signal.

the key relationship must be that between RNA chain chemistry and polypeptide synthesis. How did it come about that the ordering of amino acids in polypeptides became determined by the sequences of nucleotides in RNA chains? So far we know of no chemical features of RNA that could lead to its preferential association with amino acids, not only providing the energy to make peptide bonds, but also dictating which amino acids are linked together. Until we do, we must attempt to clarify still further the details of protein synthesis as it occurs today and hope that the primordial moments of life may thus somehow be revealed.
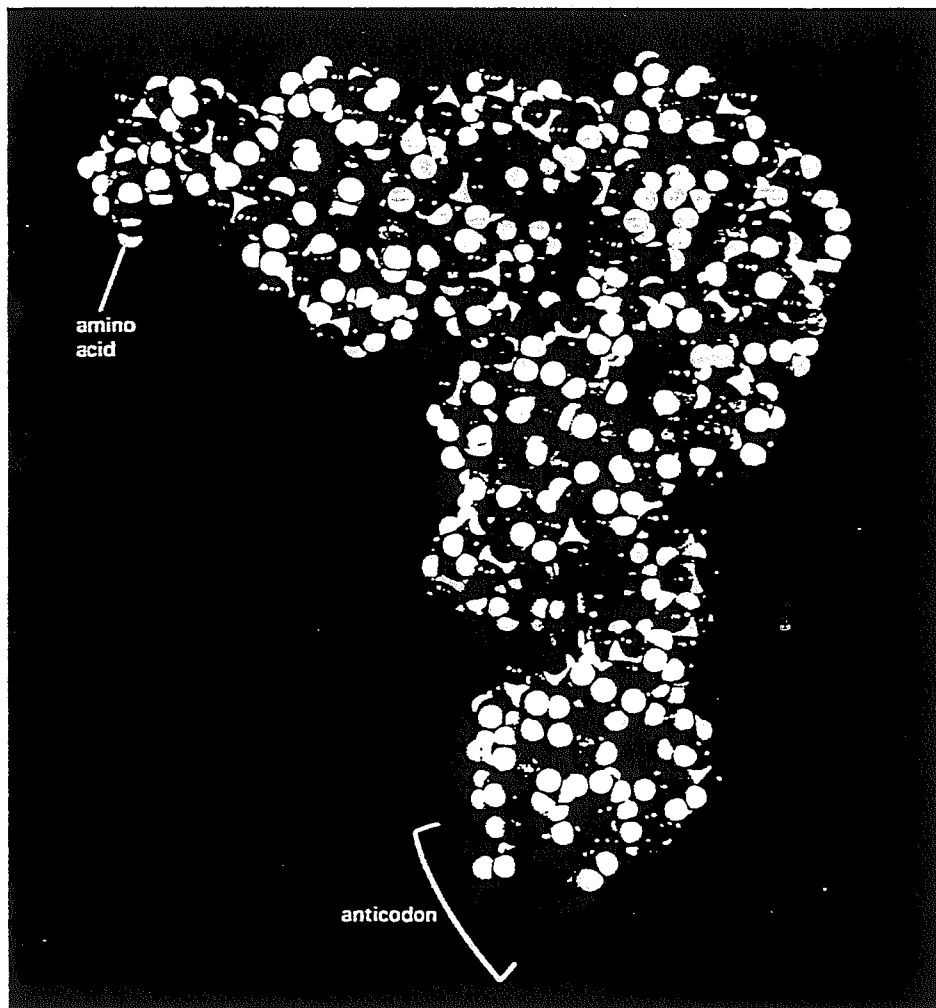
The central agents in protein synthesis are the tRNA (transfer RNA) molecules to which amino acids are attached prior to their polymerization into polypeptides (Figure 5–5). By becoming attached at their carboxyl ends, amino acids are activated to high-energy forms from which peptide bonds form spontaneously to yield polypeptides. Such activation processes are obligatory for protein synthesis since free amino acids cannot be added directly to a growing polypeptide chain. (Only the reverse process in which a peptide bond is hydrolyzed by the addition of water occurs spontaneously.)

Only the tRNA molecule, and not its attached amino acid, determines where the amino acid is added during protein synthesis. This was established by an ingenious experiment in which an amino acid attached to a specific tRNA was chemically converted into a different amino acid (cysteine → alanine). If such hybrid tRNA molecules are allowed to instruct protein synthesis in a cell-free system, the wrong amino acid is inserted at every point in the protein chain where that tRNA is used (Figure 5–6). Thus the success of de-
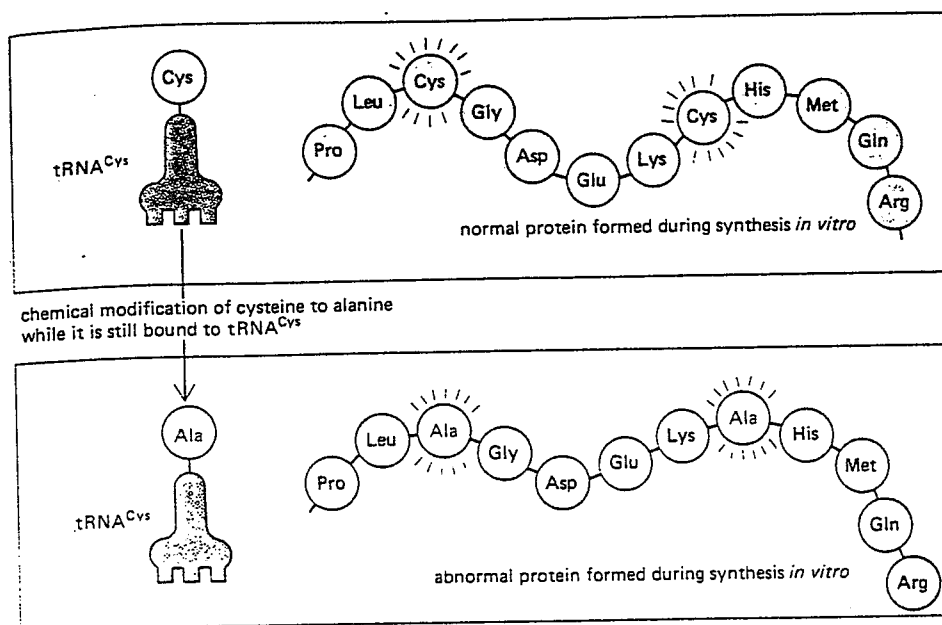
(B)

221

Figure 5-6   An experiment demonstrating that only tRNA, and not its attached amino acid, "recognizes" the site at which each amino acid is to be added during protein synthesis.
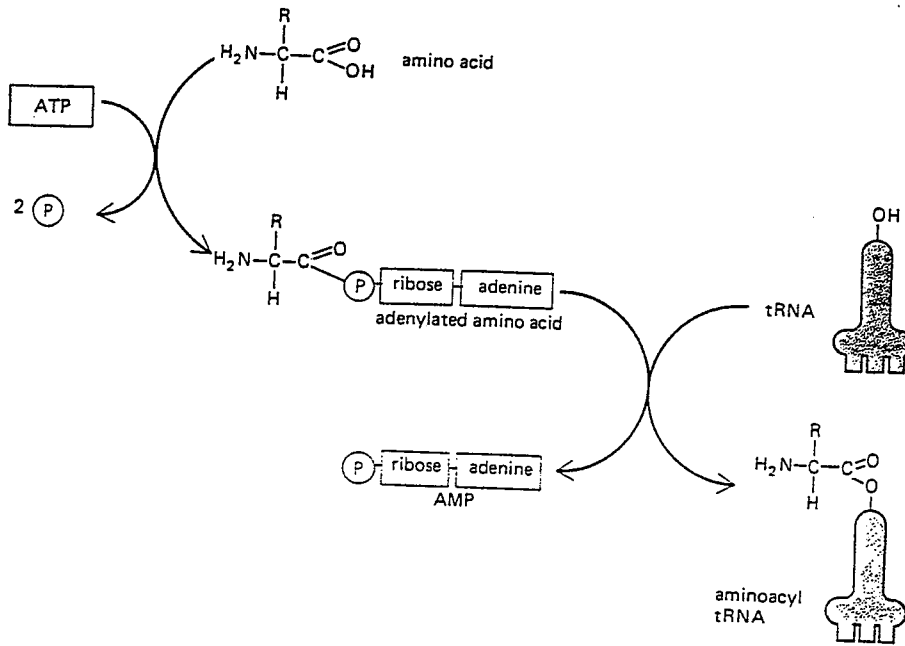
coding depends in part on the accuracy of the mechanism that normally links each activated amino acid to its corresponding tRNA molecule.

Equally important to successful mRNA decoding is the accuracy of the base-pairing between the *codons* in mRNA and the *anticodons* in each tRNA molecule (see Figure 5-9). These base pairs form effectively only when the mRNA and tRNA molecules are attached to appropriate binding slots on ribosomes. Unfortunately, we may wait many years before we deeply understand these ribosome-mediated binding interactions. As shown on page 209, ribosomes are very complex molecular aggregates, and the exact determination of their structure at the molecular level is not likely in the near future.

Specific Enzymes Couple Each Amino Acid to Its Appropriate tRNA Molecule[4]

How does each tRNA molecule recognize its corresponding amino acid? The answer is that a special set of enzymes, called **aminoacyl-tRNA synthetases,** couple each amino acid to its appropriate tRNA molecule. There is a different synthetase enzyme for every amino acid: one attaches glycine to tRNA^{Gly}, another attaches alanine to tRNA^{Ala}, and so on. The coupling reaction occurs in two steps, as illustrated in Figure 5-7, to create an **aminoacyl-tRNA** molecule (Figure 5-8).
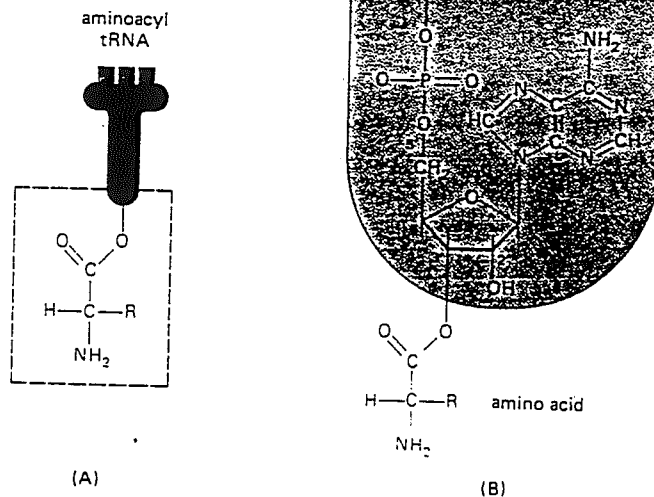
The tRNA molecules serve as the final "adaptors" that convert nucleic-acid-sequence information into protein-sequence information. However, the aminoacyl-tRNA synthetase enzymes comprise a second set of specific adaptors of equal importance to the decoding process. Thus, the genetic code is translated by means of two linked sets of adaptors, each matching one molecular surface to another with great specificity; their combined action serves to identify each amino acid with a particular sequence of three nucleotides in the mRNA molecule—a codon for that amino acid (Figure 5-9).

Amino Acids Are Added to the Carboxyl-Terminal End of a Growing Polypeptide Chain

The fundamental reaction of protein synthesis is the formation of a peptide bond between the carboxyl group at the end of a growing polypeptide chain and a free amino group on an amino acid. A protein chain is consequently synthesized stepwise from its amino-terminal end to its carboxyl-terminal end. Throughout the entire process the growing carboxyl end of the polypeptide chain remains activated by the covalent attachment of a tRNA molecule (a *peptidyl-tRNA* molecule). The covalent linkage is disrupted in each cycle but is immediately replaced by an identical linkage carried in by the new amino acid (Figure 5-10). Thus, in protein synthesis each amino acid added carries with it the activation energy for the addition of the *next* amino acid rather than for its own addition. (This is an example of "head growth," described in Chapter 2 [Figure 2-34, p. 82].)

(A)

(B)

amino acid
(tryptophan)

tRNA specific for
tryptophan (tRNA^{Trp})

aminoacyl
tRNA synthetase
specific for
tryptophan

linkage of tryptophan
to tRNA^{Trp}

tRNA^{Trp} binds to
the UGG codon

NET RESULT: TRYPTOPHAN IS RECOGNIZED BY ITS CODON

5′   3′
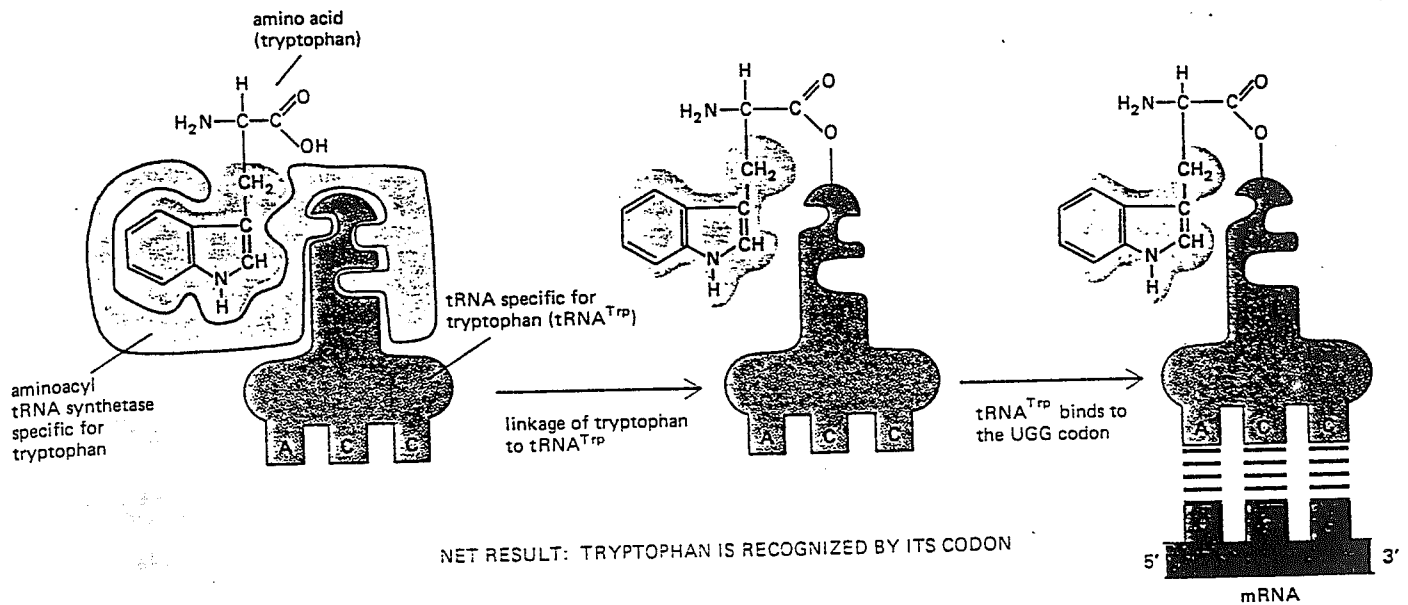
mRNA

Each Amino Acid Added Is Selected by a Complementary Base-pairing Interaction Between Its Linked tRNA Molecule and an mRNA Chain[5]

In the course of translation, the mRNA sequence is read three nucleotides at a time as the translation machinery moves in the 5′-to-3′ direction along the mRNA molecule. Each amino acid is specified by the triplet of nucleotides (**codon**) in the mRNA molecule that pairs with a sequence of three complementary nucleotides at the anticodon tip of a tRNA molecule. Because only one of the many different types of tRNA molecules in a cell can base-pair with each codon, the codon determines the specific amino acid residue to be added to the growing polypeptide chain end (Figure 5–11).

Figure 5–9 Schematic diagram illustrating how the genetic code is translated by means of two linked "adaptors": the aminoacyl-tRNA synthetase enzyme, which couples a particular amino acid to its corresponding tRNA, and the tRNA molecule, which then binds to the appropriate nucleotide sequence on the mRNA.
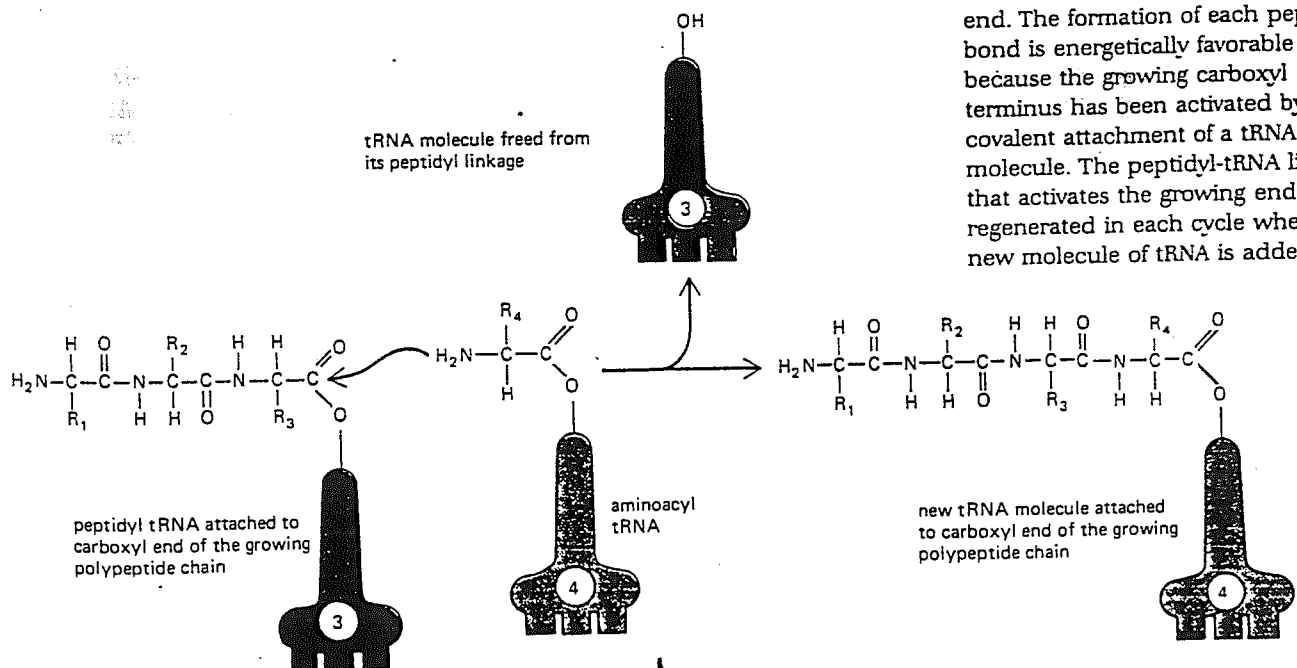
Figure 5–10 A polypeptide chain grows by the stepwise addition of amino acids to its carboxyl-terminal end. The formation of each peptide bond is energetically favorable because the growing carboxyl terminus has been activated by the covalent attachment of a tRNA molecule. The peptidyl-tRNA linkage that activates the growing end is regenerated in each cycle when a new molecule of tRNA is added.

tRNA molecule freed from
its peptidyl linkage

peptidyl tRNA attached to
carboxyl end of the growing
polypeptide chain

aminoacyl
tRNA

new tRNA molecule attached
to carboxyl end of the growing
polypeptide chain

Figure 5-11 Each amino acid added to the growing end of a polypeptide chain is selected by complementary base-pairing between the anticodon on its attached tRNA molecule and the next codon on the mRNA chain.

There are 64 different possible sequences composed of three nucleotides (4 × 4 × 4), and most of them occur somewhere in most mRNA molecules. Three of these 64 codons do not code for amino acids but instead specify the termination of a polypeptide chain; they are known as *stop codons*. That leaves 61 codons to specify only 20 different amino acids: most amino acids are therefore represented by more than 1 codon. For this reason, the genetic code is said to be *degenerate* (Figure 5-12).

The degeneracy of the genetic code means that either (1) a single tRNA molecule can base-pair with more than one codon or (2) there is more than one tRNA for each amino acid. In fact, both statements are true. For some amino acids there is more than one tRNA molecule. In addition, some tRNA molecules are constructed so that they require accurate base-pairing only at the first two positions of the codon and can tolerate a mismatch (or *wobble*) at the third.

The nucleotides in a completed nucleic acid chain (like the amino acids in proteins, p. 333) can be covalently modified in order to modulate the biological activity of the nucleic acid. In particular, many different modified nucleotides exist in a population of tRNA molecules, each produced by covalent modification after the tRNA has been synthesized: a few examples are illustrated in Figure 5-13. Some of the unusual nucleotides affect the conformation and base-pairing of the anticodon and thereby facilitate the recognition of the appropriate mRNA codon (or codons) by the tRNA molecule. Another type of posttranscriptional modification is also common to tRNA molecules: tRNAs are initially made as part of a somewhat longer RNA transcript, from which they are first cut out and then completed by an enzyme that adds the three nucleotides CCA to their 3' end. Both types of tRNA processing reactions occur in both procaryotes and eucaryotes; in contrast, an extensive processing of mRNA precursors occurs only in eucaryotes (see Chapter 8).

The Events in Protein Synthesis Are Catalyzed on the Ribosome[6]

The protein synthesis reactions that we have just described require a complex catalytic machinery to guide them. For example, the growing end of the polypeptide chain must be kept aligned with the mRNA molecule in such a way as to ensure that each successive codon in the mRNA engages with a tRNA molecule. This means that the growing end of the polypeptide must be moved exactly three nucleotides along the mRNA after the addition of each amino acid. This and the other events in protein synthesis are made possible by a large multienzyme complex composed of protein and RNA molecules—the **ribosome.**

Eucaryotic and procaryotic ribosomes are quite similar in design and in function. Each is composed of one large and one small subunit. Eucaryotic
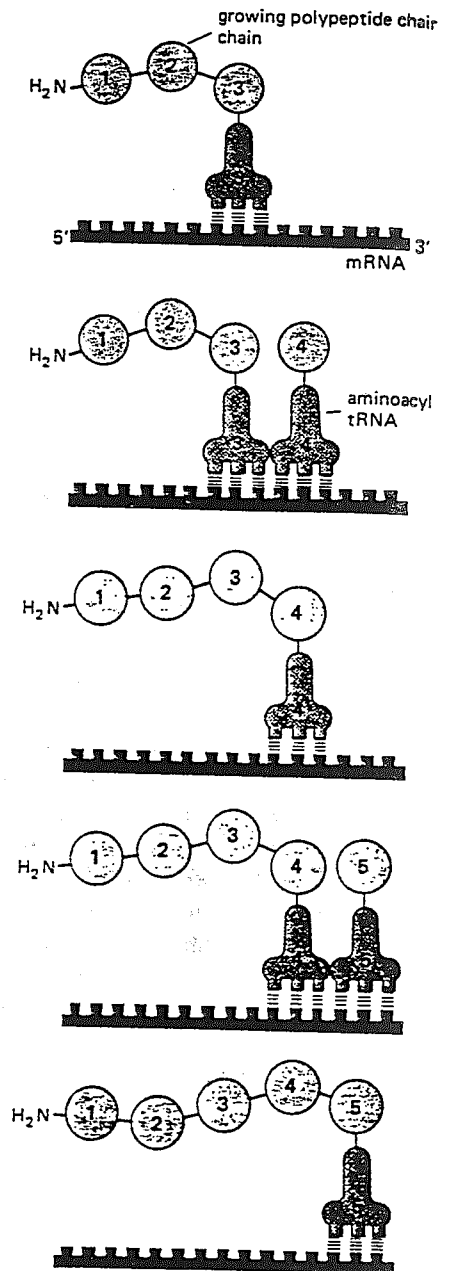
Figure 5-12 The genetic code. Codons are written with the 5'-terminal nucleotide on the left. Note that most amino acids are represented by more than one codon and that variation at the third nucleotide in a codon is common.

GCA GCC GCG GCU	AGA AGG CGA CGC CGG CGU	GAC GAU	AAC AAU	UGC UGU	GAA GAG	CAA CAG	GGA GGC GGG GGU	CAC CAU	AUA AUC AUU	UUA UUG CUA CUC CUG CUU	AAA AAG	AUG	UUC UUU	CCA CCC CCG CCU	AGC AGU UCA UCC UCG UCU	ACA ACC ACG ACU	UGG	UAC UAU	GUA GUC GUG GUU	UAA UAG UGA
Ala	Arg	Asp	Asn	Cys	Glu	Gln	Gly	His	Ileu	Leu	Lys	Met	Phe	Pro	Ser	Tnr	Trp	Tyr	Val	stop
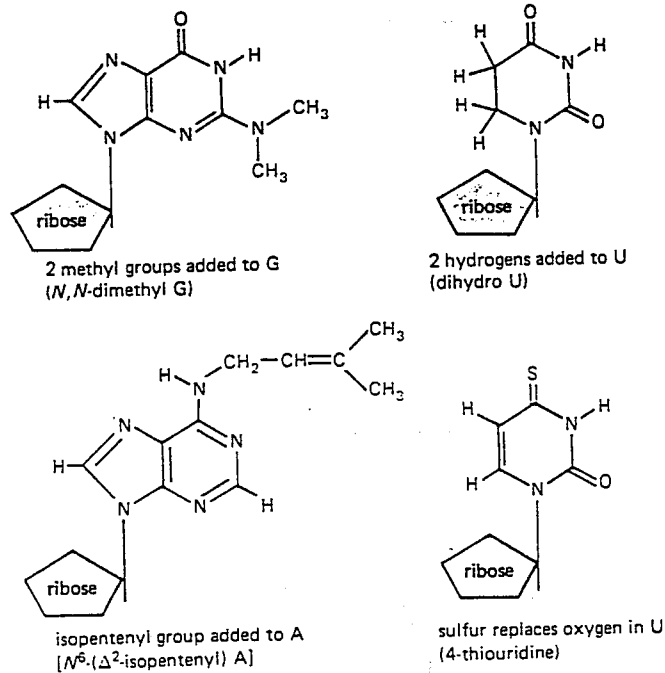
2 methyl groups added to G
(*N,N*-dimethyl G)

2 hydrogens added to U
(dihydro U)

isopentenyl group added to A
[*N*⁶-(Δ²-isopentenyl) A]

sulfur replaces oxygen in U
(4-thiouridine)

ribosomes are roughly half RNA by weight: the small subunit consists of 1 ribosomal RNA (rRNA) molecule bound to about 33 different ribosomal proteins, while the large subunit consists of 3 different rRNA molecules bound to more than 40 different ribosomal proteins. Procaryotic ribosomes are slightly smaller and contain fewer components (Figure 5–14). Both types of ribosomes have a groove that accommodates a growing polypeptide chain and a groove that accommodates an mRNA molecule. These grooves cover a stretch of about 30 amino acids and about 35 RNA nucleotides, respectively.

Figure 5–14 A comparison of the structures of procaryotic and eucaryotic ribosomes. Despite the differences in structure, they function in very similar ways.

29 nm
21 nm
70S
MW 2,800,000

50S
MW 1,800,000

30S
MW 1,000,000

5S RNA
120 nucleotides

23S RNA
3000 nucleotides

16S RNA
1500 nucleotides

~34 proteins

~21 proteins

PROCARYOTIC RIBOSOME

~32 nm
~22 nm
80S
MW 4,500,000

40S
MW 1,500,000

60S
MW 3,000,000

18S RNA
2000 nucleotides

5.8S RNA
160 nucleotides

28S RNA
5000 nucleotides

5S RNA
120 nucleotides

~33 proteins
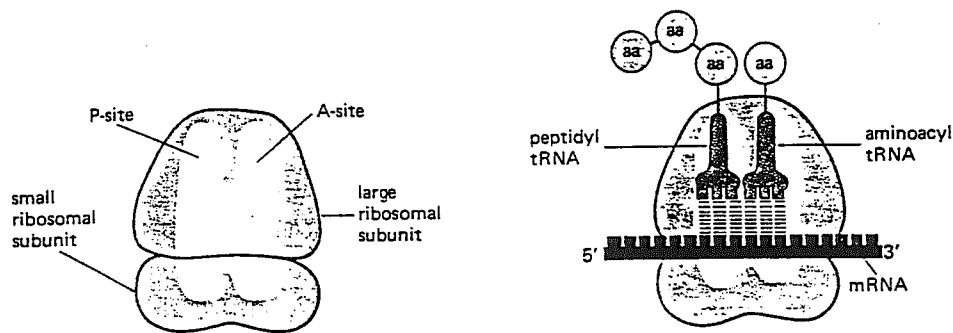
~45 proteins

EUCARYOTIC RIBOSOME

Figure 5—15 Schematic diagram of the three major RNA binding sites on a ribosome. An empty ribosome is shown at the left and a loaded ribosome at the right.

A Ribosome Moves Stepwise Along the mRNA Chain[6,7]

A ribosome contains two different binding sites for tRNA molecules: one site holds the tRNA molecule that is normally linked to the growing end of the polypeptide chain and is called the **peptidyl-tRNA binding site, or P-site;** the other holds the incoming tRNA molecule charged with an amino acid and is called the **aminoacyl-tRNA binding site, or A-site.** A tRNA molecule at either site is held in such a way that the anticodon nucleotides must form base pairs with a complementary mRNA codon in order to fit into the site. The A- and P-sites are so close together that the tRNA molecules held in each site form base pairs with adjacent codons in the mRNA molecule (Figure 5—15).

The process of polypeptide chain elongation on a ribosome can be considered as a cycle with three discrete steps (Figure 5—16). In step 1, an aminoacyl-tRNA molecule becomes bound to a vacant ribosomal A-site adjacent to an occupied P-site by forming base pairs with the three mRNA nucleotides exposed at the A-site. In step 2, the carboxyl end of the polypeptide chain is uncoupled from the tRNA molecule in the P-site and joined by a peptide bond to the amino acid linked to the tRNA molecule in the A-site. This reaction is catalyzed by **peptidyl transferase,** an enzyme that is tightly bound to the ribosome. In step 3, the new peptidyl-tRNA in the A-site is translocated to the P-site as the ribosome moves exactly three nucleotides along the mRNA molecule. This step requires energy and is driven by a series of conformational changes induced in one of the ribosomal proteins by the hydrolysis of a bound GTP molecule (see p. 137). As part of the translocation process of step 3, the free tRNA molecule that was generated in the P-site during step 2 is ejected from the ribosome to reenter the cytoplasmic tRNA pool. Therefore, upon completion of step 3, the unoccupied A-site is free to accept a new tRNA molecule linked to the next amino acid, which starts the cycle again. Since each cycle requires only about one-twentieth of a second in a bacterium under optimal conditions, the complete synthesis of an average-size protein of 400 amino acids is accomplished in about 20 seconds.

In most cells, protein synthesis consumes more energy than any other biosynthetic process. All told, four high-energy phosphate bonds are split to

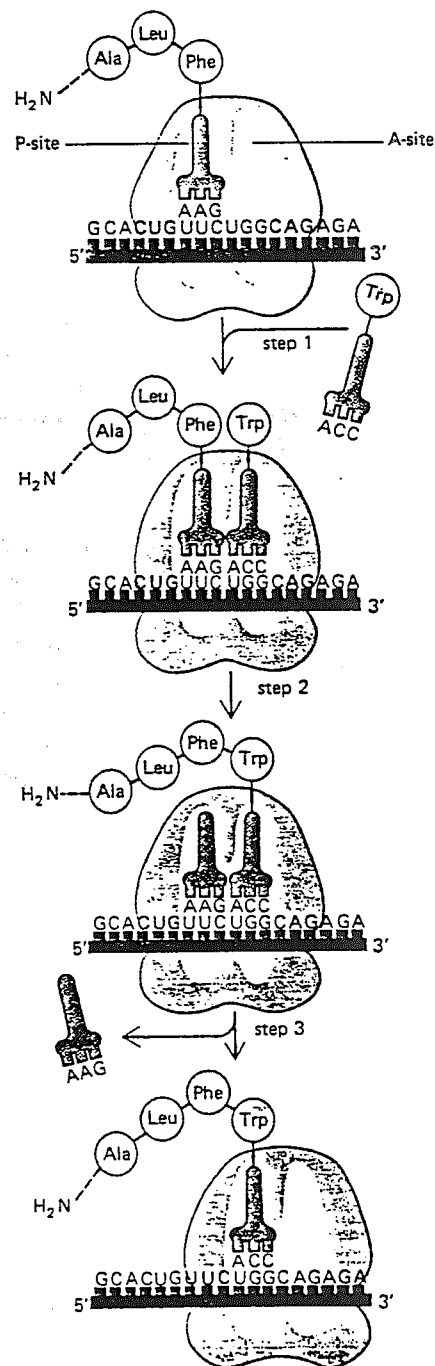**Figure 5—16** The elongation phase of protein synthesis on a ribosome. The three-step cycle shown is repeated over and over during the synthesis of a protein chain. In step 1, an aminoacyl-tRNA molecule binds to the A-site on the ribosome; in step 2, a new peptide bond is formed; and in step 3, the ribosome moves a distance of three nucleotides along the mRNA chain, "resetting" the ribosome so that the cycle can be repeated.

make each new peptide bond. Two of these are required to charge each tRNA molecule with an amino acid (Figure 5-7). And two more drive two of the cyclic reactions occurring on the ribosome during synthesis itself: one for the aminoacyl-tRNA binding in step 1, and one for the ribosome translocation in step 3.

A Protein Chain Is Released from the Ribosome Whenever One of Three Different Termination Codons Is Reached[6,8]

As already noted, three of the codons in an mRNA molecule are **stop codons,** which terminate the translation process. A protein called *release factor* binds directly to any stop codon (UAA, UAG, or UGA) that reaches the A-site on the ribosome. This binding disturbs the activity of the nearby peptidyl transferase enzyme, causing it to catalyze the addition of a water molecule instead of the free amino group of an amino acid to the peptidyl-tRNA. As a result, the carboxyl end of the growing polypeptide chain is freed from its attachment to a tRNA molecule. Since it is only this attachment that normally holds the growing polypeptide to the ribosome, the completed protein chain is released into the cell cytoplasm (Figure 5-17).

The Initiation Process Sets the Reading Frame for Protein Synthesis[6,9]

In principle an RNA sequence can be decoded in any one of three different *reading frames*, each of which will specify a completely different polypeptide chain (see Figure 3-13, p. 108). Which of the three frames is actually read is determined when a ribosome engages with an mRNA molecule to form an *initiation complex*. This complex is assembled at the exact spot on the mRNA where the polypeptide chain is to begin.

The initiation process is complicated, involving a number of steps catalyzed by proteins called **initiation factors,** many of which are themselves composed of several polypeptide chains. Because of their complexity, many of the details of initiation are still uncertain. However, it is clear that each ribosome is assembled onto an mRNA chain as two separate subunits, the small ribosomal subunit being added first. Before the mRNA is bound, a special **initiator tRNA** molecule, recognizing the codon AUG and carrying methionine, is loaded onto each small subunit. This loading reaction is catalyzed by one of the initiation factors, called *initiation factor 2*, or *IF-2*. In some eucaryotic cells the overall rate of protein synthesis is controlled by this factor (see the following page).

The small ribosomal subunit binds to the region of the mRNA molecule where protein synthesis is to begin by pairing its bound initiator tRNA molecule with a particular AUG start codon (Figure 5-18). An mRNA molecule usually contains many AUG sequences, each of which codes for methionine. But the vast majority of these will not serve as start codons. As explained elsewhere, which AUG is recognized as a start codon depends on other parts of the mRNA nucleotide sequence (p. 332).

At the completion of the initiation process, all of the initiation factors associated with the small ribosomal subunit up to this point are discharged to make way for the binding of a large ribosomal subunit to the small one. In this way, a complete functional ribosome is formed. The initiator tRNA molecule ends up bound to the P-site of the ribosome, so that the synthesis of a protein chain can begin directly with the binding of a second aminoacyl-tRNA molecule to the A-site of the ribosome (Figure 5-18). Further steps in the elongation phase of protein synthesis then proceed as described previously (see step 2 of Figure 5-16).
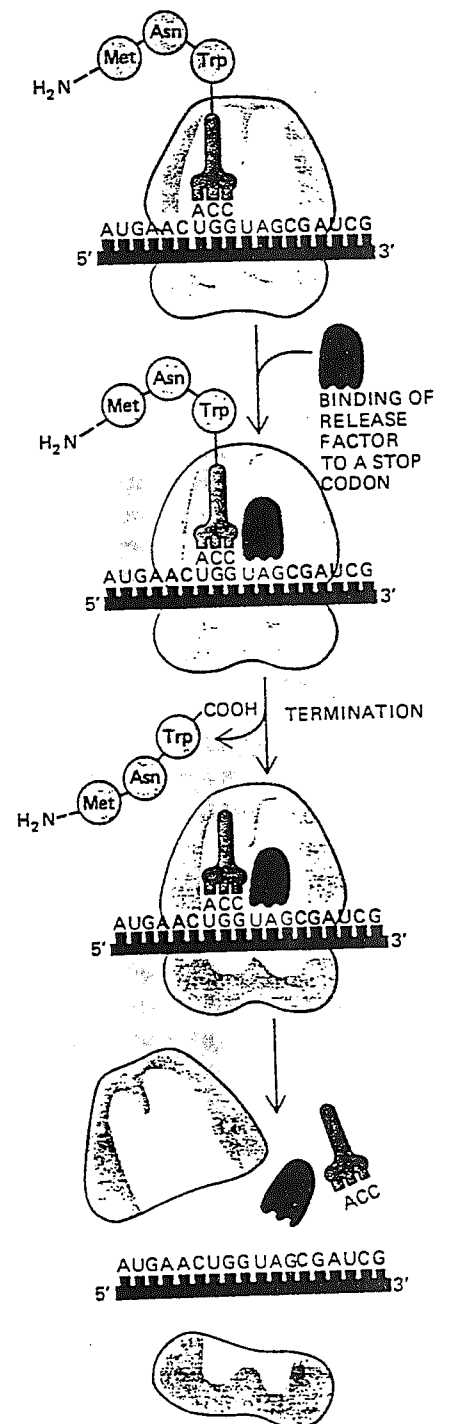
Figure 5-17   The final phase of protein synthesis. The binding of release factor to a stop codon terminates translation: the completed polypeptide is released, and the ribosome dissociates into its two separate subunits.

Figure 5–18   The initiation phase of protein synthesis. Events are illustrated as they occur in eucaryotes, but a very similar process occurs in bacteria. Step 1 and step 2 refer to steps in the elongation reaction shown in Figure 5–16.

Because an initiator tRNA molecule always carries the amino acid methionine (or the amino-formyl derivative of methionine in procaryotes), all newly made proteins have a methionine as their amino-terminal residue. This methionine is usually not important for protein function and is often removed shortly after its incorporation by the limited proteolysis of a few residues that occurs at the amino terminus of many proteins.

The Overall Rate of Protein Synthesis in Eucaryotes Is Controlled by Initiation Factors[10]

As will be discussed in Chapter 11, the cells in a multicellular organism multiply only when they are in an appropriate environment. The mechanism by which extracellular signals stimulate cells to grow and divide is not known, but one of their major effects must be to increase the overall rate of protein synthesis (p. 617). What determines this rate? Direct studies in tissues are very difficult, but when cells in tissue culture are starved of essential amino acids, glucose, or serum, there is a marked inhibition of their rate of polypeptide-chain *initiation*, which can be shown to result from inactivation of the protein synthesis initiation factor IF-2. Moreover, in at least one type of cell (immature red blood cells), the activity of IF-2 is known to be reduced in a controlled way by the phosphorylation of one of its three protein subunits. This suggests that eucaryotic protein synthesis rates are controlled in part by specific protein kinases, which in their active form inhibit the initiation of protein synthesis. One can speculate that unknown growth signals in tissues might cause cells to multiply by inactivating such protein kinases when cells are in the proper environment.

The initiation factors required for protein synthesis are more numerous and more complex in eucaryotes than in procaryotes, even though they perform the same basic functions. Many of the extra components could be regulatory proteins that coordinate cell growth in multicellular eucaryotes by controlling protein synthesis.

Many Inhibitors of Procaryotic Protein Synthesis Are Useful as Antibiotics[11]

Many of the most effective antibiotics used in modern medicine act by inhibiting bacterial protein synthesis. A number of these drugs exploit the structural and functional differences between procaryotic and eucaryotic ribosomes. Their selectivity often enables such compounds to be used at relatively high concentrations in the human body without undue toxicity. Because different antibiotics bind to different protein subcomplexes on bacterial ribosomes, they often inhibit different steps in the synthetic process. Some of the more common compounds in this group are listed in Table 5–1, along with their specific effects. Also listed in this table are several other commonly used inhibitors of protein synthesis, some of which act on eucaryotic cells. Because they can be used to block specific steps in the processes that lead from DNA to protein, the compounds listed are widely used for a variety of biochemical and cell biological studies.
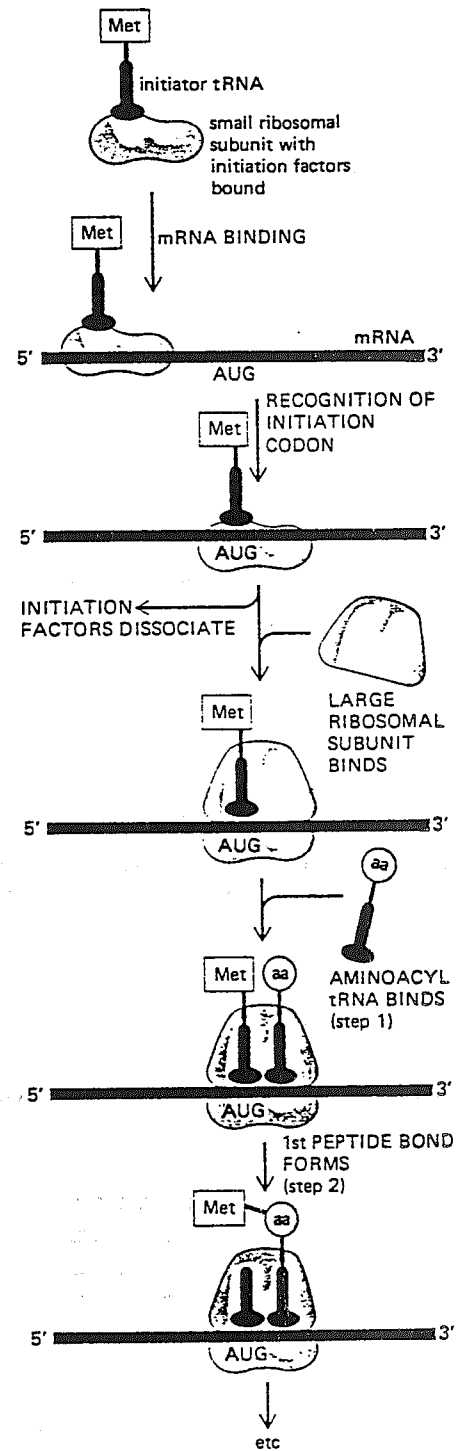
229

Table 5–1   Inhibitors of Protein or RNA Synthesis

Inhibitor	Specific Effect
*Acting Only on Procaryotes**	
Tetracycline	blocks binding of aminoacyl-tRNA to A-site of ribosome
Streptomycin	prevents the transition from initiation complex to chain-elongating ribosome
Chloramphenicol	blocks the peptidyl transferase reaction on ribosomes (step 2 in Figure 5–16)
Erythromycin	blocks the translocation reaction on ribosomes (step 3 in Figure 5–16)
Rifamycin	blocks initiation of RNA chains by binding to RNA polymerase (prevents RNA synthesis)
Acting on Procaryotes and Eucaryotes	
Puromycin	causes the premature release of nascent polypeptide chains by its addition to growing chain end
Actinomycin D	binds to DNA and blocks the movement of RNA polymerase (prevents RNA synthesis)
Acting Only on Eucaryotes	
Cycloheximide	blocks the peptidyl transferase reaction on ribosomes (step 2 in Figure 5–16)
α-Amanitin	blocks mRNA synthesis by binding preferentially to RNA polymerase II

*The ribosomes of eucaryotic mitochondria (and chloroplasts) often resemble those of procaryotes in their sensitivity to inhibitors.

Summary

Before the synthesis of a particular protein can begin, the corresponding mRNA molecule must be produced by DNA transcription processes and exported to the cell cytoplasm. An extensive machinery is then called into play. The process begins with the binding of a small ribosomal subunit to an mRNA molecule. A unique initiator tRNA molecule positions the small ribosomal subunit over a special start codon on the mRNA. A large ribosomal subunit is added to complete the ribosome, and the elongation phase of protein synthesis ensues. Each amino acid is added to the carboxyl-terminal end of the growing polypeptide by means of a cycle of three sequential steps: aminoacyl-tRNA binding, followed by peptide bond formation, followed by ribosome translocation. The ribosome progresses from codon to codon in the 5′-to-3′ direction along the mRNA molecule until one of three stop codons is reached. A release factor then binds to the stop codon, terminating translation and releasing the completed polypeptide from the ribosome.

An aminoacyl-tRNA molecule serves as a decoding device that allows a particular sequence of three ribonucleotides in the mRNA to be translated as a unique amino acid in the newly synthesized protein. Each of the 20 amino acids is fitted to a particular codon by means of a two-step recognition process: in the first step the amino acid is recognized by a unique aminoacyl-tRNA synthetase enzyme that links it to a specific tRNA molecule; and in the second step a particular sequence of three nucleotides in the mRNA chain is recognized by the anticodon of the tRNA molecule. There is at least one specific aminoacyl-tRNA synthetase and at least one specific tRNA for each amino acid.

230