

## CONTENTS

- General Considerations
- Major Breakthroughs in Protein Synthesis
- Central Dogma of Molecular Genetics
- Phase of Protein Synthesis
- The Two Key Components in Protein Synthesis
- Activation of Amino Acids
- Initiation of Protein Synthesis
- Elongation of the Polypeptide Chain
- Termination of Polypeptide Synthesis and Release of Polypeptide Chain
- Folding and Processing of Polypeptide Chain
- Energy Requirements for Peptide Bond Formation
- Inhibitors of Protein Synthesis
- Eukaryotic Protein Synthesis
- Protein Synthesis in Mitochondria and Chloroplasts
- Salient Features of Ribosomal Protein Synthesis
- Biosynthesis of Short Peptides
- Evolution of Protein Synthesis



E site P site A site

**Transfer RNA-binding sites**

Three tRNA-binding sites are present on the 70S ribosome. They are called the A (for aminoacyl), P (for peptidyl), and E (for exit) sites. Each tRNA molecule contacts both the 30S and the 50S subunit.

## CHAPTER

## 28

# Biosynthesis of Proteins

## GENERAL CONSIDERATIONS

Proteins are the end products of most metabolic pathways. A typical cell requires thousands of different protein molecules at any given moment. These must be synthesized in response to the cell's current requirements, transported (or targeted) to the appropriate cellular location and ultimately degraded when the cell no longer requires them. In fact, *the protein synthesis is well understood than protein targeting or degradation.*

*Protein synthesis is the most complex of all the biosynthetic mechanisms.* In eukaryotic cells, protein synthesis requires the participation of over 70 different ribosomal proteins ; 20 or more enzymes to activate the amino acid precursors; about 15 auxiliary enzymes and other specific protein factors for initiation, elongation and termination of polypeptides; perhaps 100 additional enzymes for the final processing of different kinds of proteins and 40 or more types of transfer and ribosomal RNAs (Lehninger, Nelson and Cox, 1993). Thus, about 300 different macromolecules are required to synthesize polypeptides or proteins. Many of these macromolecules are organized into the complex 3-D structure of the ribosome to carry out stepwise translocation of the mRNA as the polypeptide is assembled. Despite this great complexity, proteins are made at exceedingly high rates. For example, in *E. coli* cell at 37°C, a complete polypeptide chain of 100 amino acid residues is synthesized in about 5 seconds. The synthesis of different proteins in each cell is

strictly regulated so that only the required number of molecules of each protein is made under certain metabolic conditions. The targetting and degradative processes keep pace with synthetic process so that an appropriate concentration of proteins is maintained in the cell.

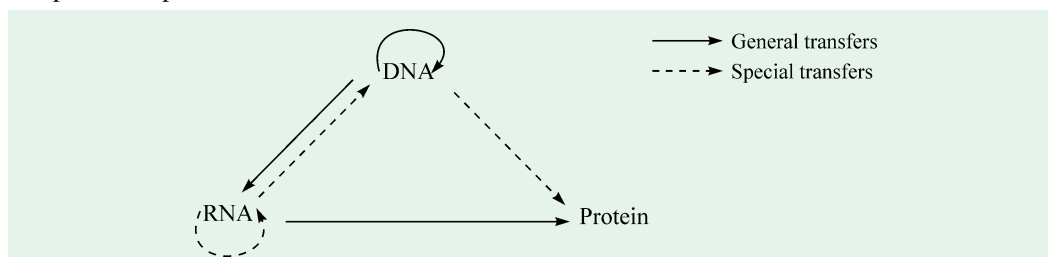
Of the total chemical energy used for all the biosynthetic processes, a cell utilizes as much as 90% for protein synthesis only. In *E. coli*, the numbers of different types of proteins and RNA molecules involved in protein synthesis are similar to those in eukaryotic cells. When totalled, the 20,000 ribosomes, 1,00,000 related protein factors and enzymes and 2,00,000 tRNAs present in a typical bacterial cell (with a volume of  $100 \text{ nm}^3$ ) can account for more than 35% of the cell's dry weight.

The transformation of hereditary information in biological systems falls into 3 categories (Fig. 28-1) :

**1. General (or Information) transfers.** These are those that can occur in all cells and are of 3 types: DNA, RNA and protein syntheses.

**2. Special transfers.** These are those that occur in cells only under special circumstances and are also of 3 types: RNA replication, reverse transcription and DNA translation.

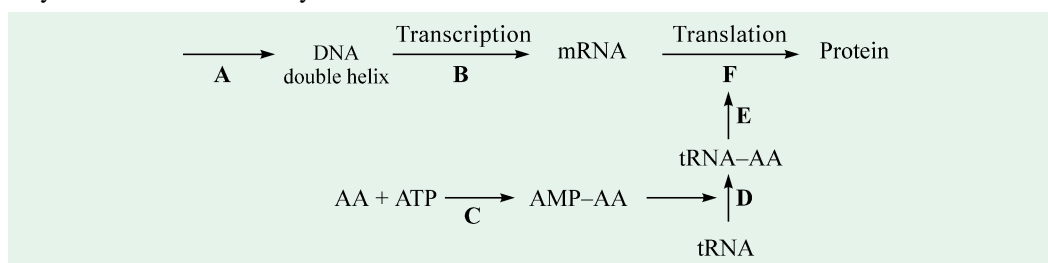
**3. Forbidden (or Unknown) transfers.** These are those that have never been detected experimentally or predicted theoretically. They too are of 3 types: protein to DNA, protein to RNA and protein to protein.



**Fig. 28-1. Transfer of genetic information in biological systems**

The central dogma of molecular biology has been depicted here. Forbidden transfers are not diagrammed.

In fact, information in the gene is contained in the base sequence of one of the strands of the double helix of DNA. Messages are transcribed onto a strand of RNA which serves as a template for the specific transfer RNA units carrying the amino acids for protein synthesis. The sequence of step may be shown schematically as :



In the above scheme, some possible points of control have been indicated by the inclusion of letters A through F (depicted in boldface).

Let us take an overview of protein synthesis. A protein is synthesized in the amino-to-carboxyl direction by the sequential addition of amino acids to the carboxyl end of the growing peptide chain. The activated precursors are aminoacyl-tRNAs, in which the carboxyl group of an amino acid is joined to the  $3'$ -OH of a transfer RNA (tRNA). The linking of an amino acid to its corresponding tRNA is catalyzed by an *aminoacyl-tRNA synthetase*. This activation reaction, which is analogous to

**720 FUNDAMENTALS OF BIOCHEMISTRY**

the activation of fatty acids, is driven by ATP. For each amino acid, there is at least one kind of tRNA and activating enzyme.

*Protein synthesis is a continuous anabolic process and takes place intracellularly. There is no evidence to suggest that proteins can replicate. Self replication seems to be confined to nucleic acids exclusively. The protein synthesis, in its simplest form, takes place as depicted in Fig. 28–2.*

**MAJOR BREAKTHROUGHS IN PROTEIN SYNTHESIS**

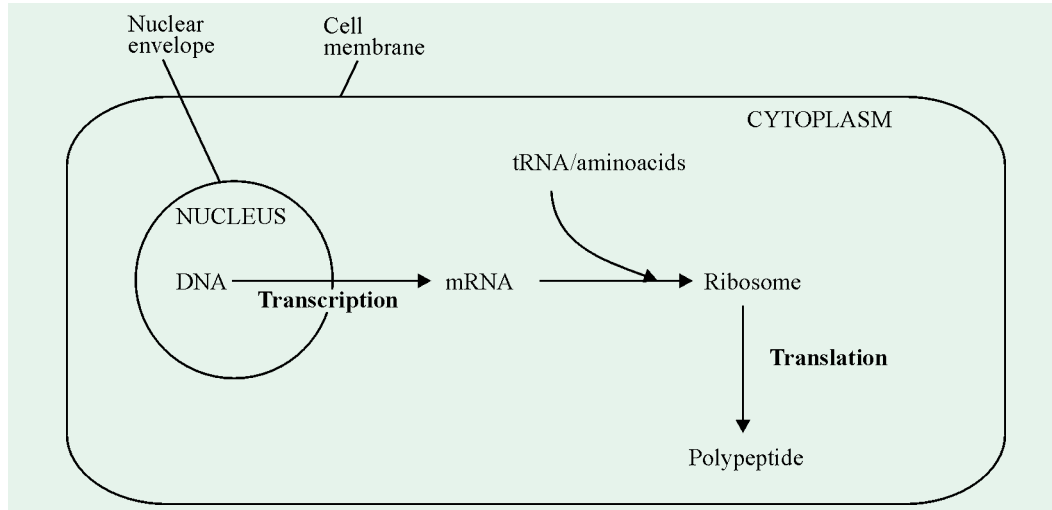
Three major events in 1950s set the stage for protein synthesis :

**(a) Ribosomes as the Site of Protein Synthesis**

In the early 1950s, **Paul Zamecnik** *et al* conducted an experiment to find out the site of protein synthesis. They injected radioactive amino acids into rats and then after different time intervals, the liver was removed, homogenized and fractionated by centrifugation. The subcellular fractions were then examined for the presence of radioactive protein. After few hours or days of the labelled amino acids injected, all the subcellular fractions contained labelled proteins. However, when the liver was removed and fractionated only minutes after injection of the labelled amino acids, labelled protein was found only in a fraction containing ribonucleoprotein particles, which were later named ribosomes.



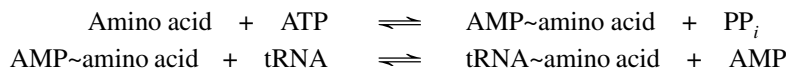
**PAUL ZAMECNIK**



**Fig. 28–2. Summary diagram of the main steps involved in protein synthesis**  
(After Green NPO, Stout GW and Taylor DJ, 1984)

**(b) RNA as a Receptor Molecule**

From studies on protein synthesis in rat liver, **Mahon Hoagland**, in 1955, identified the activated form of amino acids. He proposed that the cytoplasmic enzymic reaction requires ATP and an amino acid and produces an aminoacyl-adenylate (activated amino acid) and inorganic pyrophosphate. Later, in 1957, Hoagland, Zamecnik and **Mary Stephenson** found that amino acid activation also utilizes a small RNA (tRNA), found in the soluble fraction of the cell, to form aminoacyl-tRNAs; the reaction being catalyzed by aminoacyl-tRNA synthetases. In other words, the role for a nucleic acid as a receptor molecule in amino acid activation was, hence, identified.



### (c) Adaptor Hypothesis

The third major advance in protein synthesis took place when **Francis H.C. Crick** (1953) found out how the genetic information that is coded in the 4-letter language of nucleic acids is translated into the 20-letter language of proteins. He opined that tRNA must serve the role of an adaptor, one part of the tRNA molecule binding a specific amino acid and some other part of the tRNA recognizing a short nucleotide sequence (a trinucleotide) in mRNA coding for that amino acid. This idea, termed as adaptor hypothesis, was soon verified. The tRNA adaptor ‘translates’ the nucleotide sequence of an mRNA into the amino acid sequence of a polypeptide. This overall mRNA-guided process of protein synthesis is often referred to as **translation**. As might be expected, translation is a more complex process than either the replication or the transcription of DNA, which takes place within the framework of a common base-pairing language. In fact, translation necessitates the coordinated interplay of more than a hundred kinds of macromolecules.

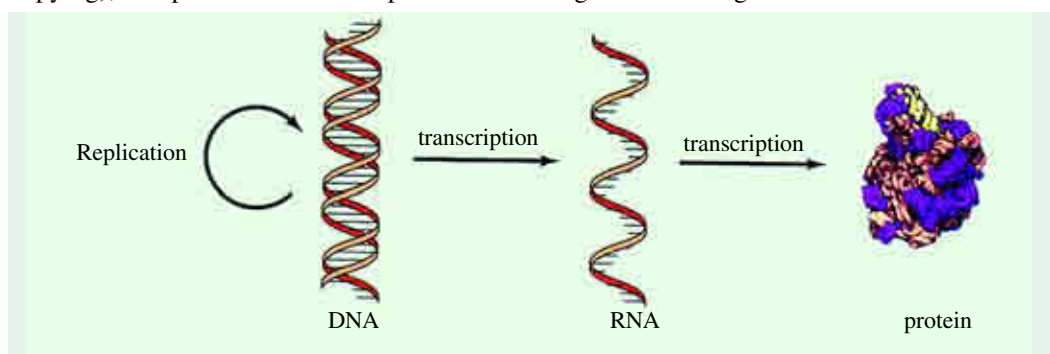
Adaptor is also spelt as **adapter**.

## CENTRAL DOGMA OF MOLECULAR GENETICS

The 3 different disciplines—genetics, physics and biochemistry — have contributed a lot to the advancement of knowledge of genetical information pathways. The discovery of the double-helical structure of DNA by **Watson and Crick** (1953) epitomized the contributions of these 3 fields. *Genetics* contributed the concept of coding by genes. *Physics* made possible the determination of molecular structure of DNA by x-ray diffraction analysis. *Biochemistry* revealed the chemical composition of DNA.

The structure of DNA itself suggested how DNA might be copied so that the information contained therein could be transmitted from one generation to the next. Understanding of how the information contained in DNA was converted into functional proteins became possible through the discovery of messenger RNA, transfer RNA and the genetic code. All these advancements led to the understanding of the process of protein synthesis through the **central dogma of molecular genetics**, which postulates that genetic information flows from nucleic acids to protein. This concept, first forwarded by Francis H.C. Crick in 1958, consists of 3 steps or processes (Fig. 28–3). The *first step* is **replication**, *i.e.*, the copying of parental DNA to form daughter DNA molecules having identical nucleotide sequences. The *second step* is **transcription** (= copying), the process in which parts of coded genetic message

The term ‘**dogma**’ (meaning literally ‘a body of beliefs’) is a misnomer here. It was introduced by Francis H.C. Crick (1958) at a time when little evidence supported these ideas. The ‘dogma’ is now a well-established principle.



**Fig. 28–3. Central dogma of molecular genetics**

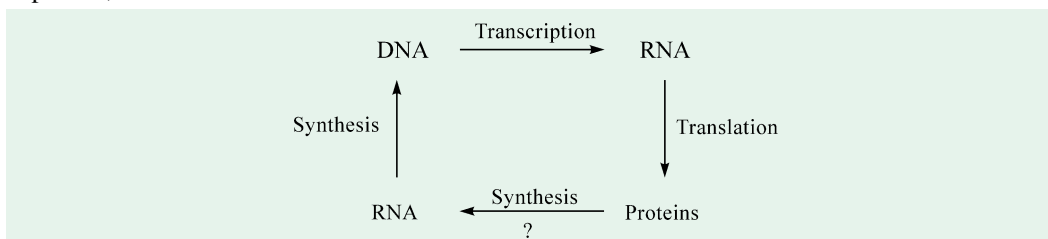
The principle depicts the general pathways of information flow *via* the processes of replication, transcription and translation. in DNA are copied precisely in the form of RNA. The RNA so formed is called messenger RNA (mRNA) as it contains the information or ‘message’ as to the sequence of nucleotides present in the

## 722 FUNDAMENTALS OF BIOCHEMISTRY

original DNA molecule. *Transcription does not involve a change of code* since DNA and the resulting mRNA are complementary. Transcription may also be defined as DNA-dependent RNA synthesis. The third step is translation, in which the genetic message coded in mRNA is translated, on the ribosomes, into a protein with a specific sequence of amino acids. In other words, reproduction of a primary polypeptide chain according to specification of mRNA is called translation. Apparently, *translation involves a change of code from nucleotide sequences to amino acid sequences*. Thus, according to this central dogma, the flow of information is one way or **unidirectional**, *i.e.*, from DNA the information is transferred to RNA (mRNA) and from RNA to protein. In other words,

**‘DNA makes RNA makes protein’.**

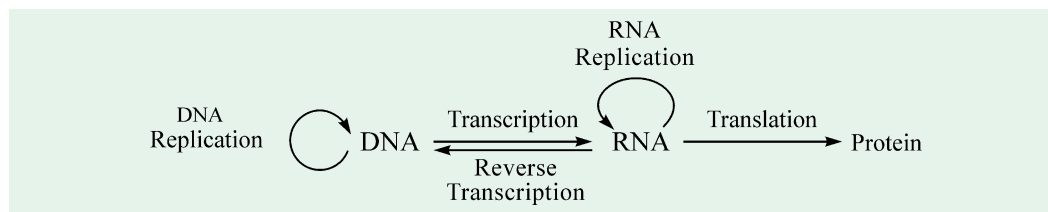
However, in 1968, Barry Commoner suggested a **circular** flow of information, *i.e.*, DNA transcribes RNA, RNA translates into proteins, and proteins synthesize RNA and RNA synthesizes DNA (Fig. 28–4). But there is no evidence whatsoever for the synthesis of RNA from protein. If this is proved, one would have to believe in Lamarckism.



**Fig. 28–4. Circular flow of information in DNA and protein syntheses**

Note that this circular flow concept is untenable as there is no proof for the synthesis of RNA from protein.

Interestingly, the existence of enzymes **RNA-dependent DNA polymerases** (or **reverse transcriptases**) in RNA viruses was predicted by an American biochemist **Howard Temin** in 1962. And the enzymes were ultimately demonstrated to be present in such viruses by Temin and his compatriot **David Baltimore** both in 1970 but independently. Their discovery aroused much interest particularly because it constituted molecular proof that genetic information can sometimes flow ‘backward’ from RNA to DNA. The RNA viruses containing reverse transcriptases are also known as retroviruses (*retro*<sup>L</sup> = backward). This exciting finding gave rise to the concept of **central dogma reverse** which means that the sequence of information flow is not necessarily from DNA to RNA to protein but can also take place from RNA to DNA (Fig. 28–5).



**Fig. 28–5. The central dogma reverse of molecular genetics**

Note that the diagram is the extension of the central dogma to include RNA-dependent synthesis of RNA and DNA.

## PHASES OF PROTEIN SYNTHESIS

Like DNA and RNA, the synthesis of polymeric protein biomolecules can be distinguished into 3 stages: initiation, elongation and termination of polypeptide chain. Besides these 3 stages or phases, the activation of amino acids and the post-translational processing of the completed polypeptide chain constitute 2 important and complex additional stages in the synthesis of protein, thus making

**HOWARD TEMIN**

(LT, 1934–1994)



**DAVID BALTIMORE**

(Born, 1938)



Temin and Baltimore, who independently discovered reverse transcriptase in 1970, were both proteges of **Renato Dulbecco**, one of the major influences in animal virology. Temin worked with Dulbecco at the California Institute of Technology in the late 1950s, Baltimore with him at the Salk Institute, San Diego, a few years later. This was the period when the possibility that cell transformation might be linked to integration of a viral genome into chromosomal DNA was first being talked about. Temin and Baltimore both attended the same undergraduate college. Swarthmore in

Pennsylvania, Temin graduating in 1955 and Baltimore in 1960. Temin spent most of his career at the University of Wisconsin at Madison; Baltimore has spent his at the Massachusetts Institute of Technology. Both Temin and Baltimore along with another fellow countryman Renato Dulbecco, working in London, shared the prestigious Nobel Prize in medicine or physiology for 1975.

Their thoughtful expressions about conducting research and integrating related branches are :

“Scientific research is fascinating because of never-ending surprises. Top quality work on any subject can suddenly lead to fundamentally important knowledge.” (Temin)

“Starting as a biochemist and later integrating genetics and cell biology into my work, I have come to realize the great power provided by joining these three streams of modern biological research.” (Baltimore)

the total number of stages to be 5. The cellular components required for each of the 5 stages in *Escherichia coli* and other bacteria are listed in Table 28–1. The requirements in eukaryotic organisms are quite similar.

**Table 28–1. Cellular components required for all the 5 stages of protein synthesis in *Escherichia coli***

Stage	Essential components
1. Activation of amino acids	20 amino acids (L type); 20 aminoacyl-tRNA synthetases; 20 or more tRNAs; ATP; Mg <sup>2+</sup>
2. Initiation	mRNA; N-formylmethionyl-tRNA (= initiator tRNA); Initiation codon in mRNA ( <i>i.e.</i> , AUG); 30S ribosomal subunit; 50 S ribosomal subunit; Initiation factors (IF-1, IF-2, IF-3); GTP; Mg <sup>2+</sup>
3. Elongation	Functional 70S ribosome (initiation complex); Aminoacyl-tRNAs specified by codons; Elongation factors (EF-Tu, EF-Ts, EF-G); Peptidyl transferase; GTP; Mg <sup>2+</sup>
4. Termination and release	70S ribosome; Termination codons in mRNA (UAA, UAG, UGA); Polypeptide release factors (RF <sub>1</sub> , RF <sub>2</sub> , RF <sub>3</sub> ); ATP
5. Folding and processing	Specific enzymes and cofactors for removal of initiating residues and signal sequences; Additional proteolytic processing; Modification of terminal residues; Attachment of phosphate, methyl, carboxyl, carbohydrate, or prosthetic groups

An overview of these stages will help a better understanding of the discussion that follows.

**Stage 1: Activation of amino acids**

This stage takes place in the cytosol (and not on the ribosomes). During this stage, each of the 20 amino acids is covalently attached to a specific tRNA at the expense of ATP energy; the reactions being catalyzed by a group of Mg<sup>2+</sup>-dependent activating enzymes called *aminoacyl-tRNA synthetases*,



**724 FUNDAMENTALS OF BIOCHEMISTRY**

each specific for one amino acid and its corresponding tRNAs. Aminoacylated tRNAs are usually referred to as being 'charged'.

**Stage 2 : Initiation**

During this stage, the mRNA, bearing the code for the polypeptide to be synthesized, binds to the smaller ribosomal subunit. This is followed by the binding of the initiating aminoacyl-tRNA and the large ribosomal subunit to form an initiation complex. The initiating aminoacyl-tRNA base-pairs with the mRNA codon AUG that signifies the beginning of the polypeptide chain. This process, which requires GTP, is promoted by specific cytosolic proteins called *initiation factors*.

**Stage 3 : Elongation**

The polypeptide chain now increases in length by covalent attachment of successive amino acid units, each carried to the ribosome and correctly positioned by its tRNA, which base-pairs to its corresponding codon in the mRNA. Elongation is promoted by cytosolic proteins called *elongation factors*. The binding of each incoming aminoacyl-tRNA and the movement of the ribosome along the mRNA are facilitated by the hydrolysis of 2 moles of GTP for each residue added to the growing polypeptide chain.

**Stage 4 : Termination and release**

The completion of the polypeptide chain is signalled by a termination codon present in the mRNA. The polypeptide chain is then released from the ribosome with the help of proteins called *release factors*.

**Stage 5 : Folding and processing**

In order to attain its biologically active form, the polypeptide chain must fold into its proper 3-dimensional conformation. Before or after folding, the new polypeptide may undergo enzymatic processing to carry out certain functions :

- (a) to remove one or two amino acids from the amino terminus,
- (b) to add acetyl, phosphate, methyl, carboxyl, or other groups to certain amino acid residues,
- (c) to cleave the protein molecule proteolytically, or
- (d) to attach oligosaccharides or prosthetic groups to the protein molecule.

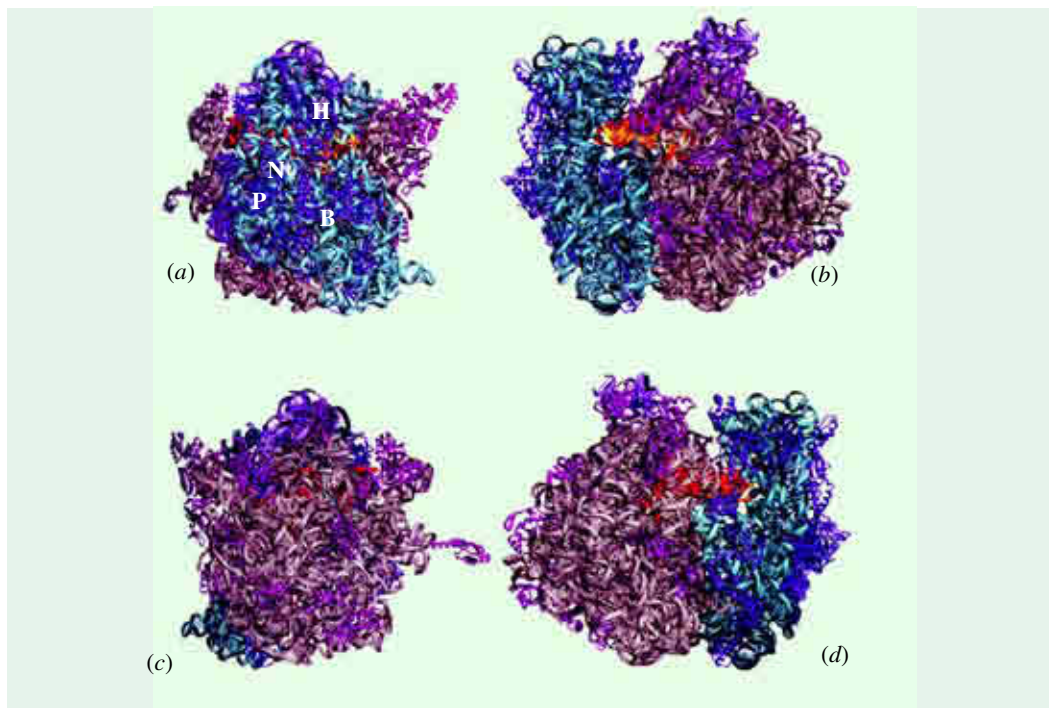
**THE TWO KEY COMPONENTS IN PROTEIN SYNTHESIS**

The two key components of protein synthesis are: ribosomes and transfer RNA.

**A. Ribosome**

The ribosomes are ribonucleoprotein cellular particles and can be regarded as the organelles of protein synthesis, just as mitochondria are the organelles of oxidative phosphorylation. A ribosome is a highly specialized and complex structure. The best-characterized ribosomes are those of *E. coli*. Each *E. coli* cell contains 15,000 or more ribosomes, which make up about 25% of the dry weight of the cell. The *E. coli* ribosomes (MW =  $2.7 \times 10^6$  or 2,500 kdal) have a diameter of about 21 nm (1 nm =  $10 \text{ \AA}$ ) and are made up of about 66% ribosomal RNA and 34% proteins. Since these particles are rich in RNA content, these were called 'ribosomes'. These have a sedimentation coefficient of 70S when  $\text{Mg}^{2+}$  ion concentration in the solution is 0.01 M. If this concentration is lowered to 0.001 M, the 70S ribosomes dissociate into 50S and 30S particles and this phenomenon is reversible, if the  $\text{Mg}^{2+}$  ion concentration is again raised.

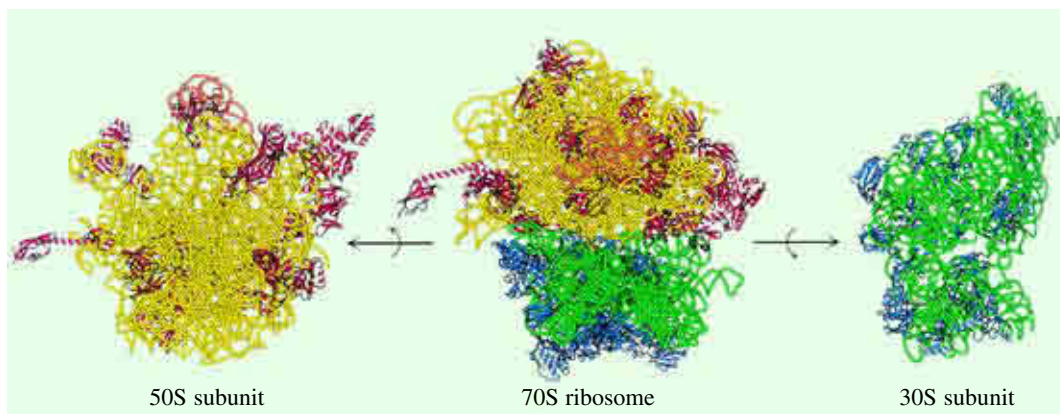
The detailed 3-dimensional structure of the bacterial ribosomes has been worked out using crystallography (Figs. 28-6 and 28-7) at different resolutions. The bacterial ribosomes consist of 2 subunits of unequal size (Fig. 28-8), the larger having a sedimentation coefficient of 50S and the smaller of 30S. The 50S subunit contains one mole of 5S rRNA (120 nucleotides), one mole of 23S rRNA (2,904 nucleotides) and 36 proteins (MW = 5,000 – 25,000). The 30S subunit contains one mole of 16S rRNA (152 nucleotides) and 21 proteins (MW = 8,000 – 26,000). The proteins are designated by numbers. Those in the large subunit are numbered L1 to L36 (L for large) and those in



**Fig. 28-6. Crystallographic structure of a 70S ribosome at 5.5 Å resolution**

The structure of the *Thermus thermophilus* ribosome is shown at 5.5 Å resolution. Successive views are rotated 90° around the vertical axis. (a) The small subunit lies atop the large subunit. Small subunit features include the head (H), connected by neck (N) to the body (B), and the platform (P) which projects toward the large subunit. The 16S RNA is coloured cyan and small subunit proteins are dark blue. (b) The large subunit is at the right; 23S RNA is grey, 5S RNA is light blue, and large subunit proteins are magenta. A-site tRNA (gold) spans the subunits. (c) The large subunit lies on top with the stalk protruding to the left. (d) The large subunit is at the left, and elements of E- and A-side tRNAs are visible in the subunit interface.

(Courtesy : Drs. A Baucom and HE Noller, 2001)



**Fig. 28-7. The ribosome at high resolution**

Detailed models of the ribosome based on the results of x-ray crystallographic studies of the 70S ribosome and the 30S and 50S subunits. 23S RNA is shown in yellow, 5S RNA in orange, 16S RNA in green, proteins of the 50S subunit in red and proteins of the 30S subunit in blue.

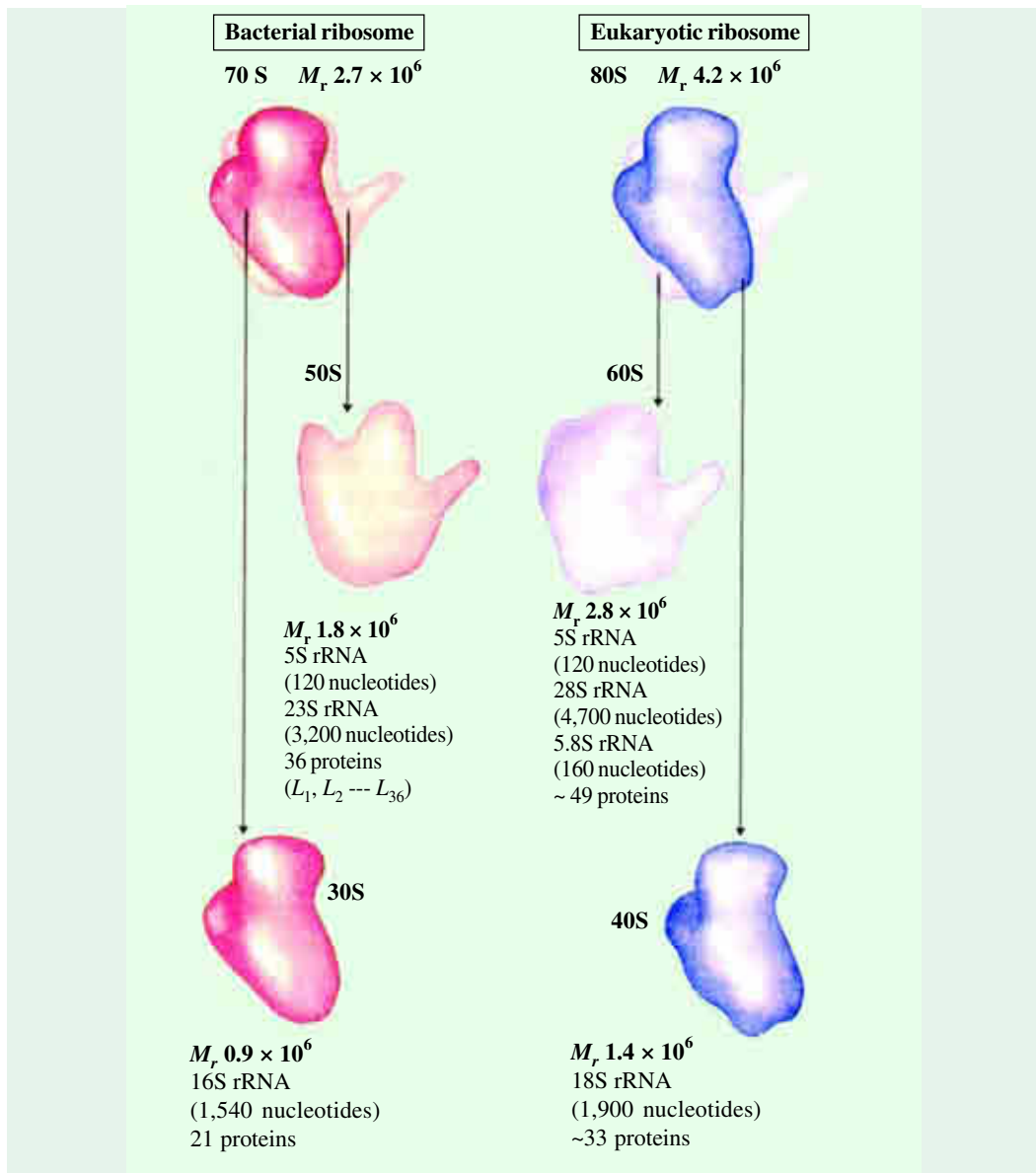
the smaller subunit S1 to S21 (S for small). Their molecular weight range between 6,000 to 75,000. The rRNAs appear to serve as a framework to which the ribosomal proteins are bound. Each of the



## 726 FUNDAMENTALS OF BIOCHEMISTRY

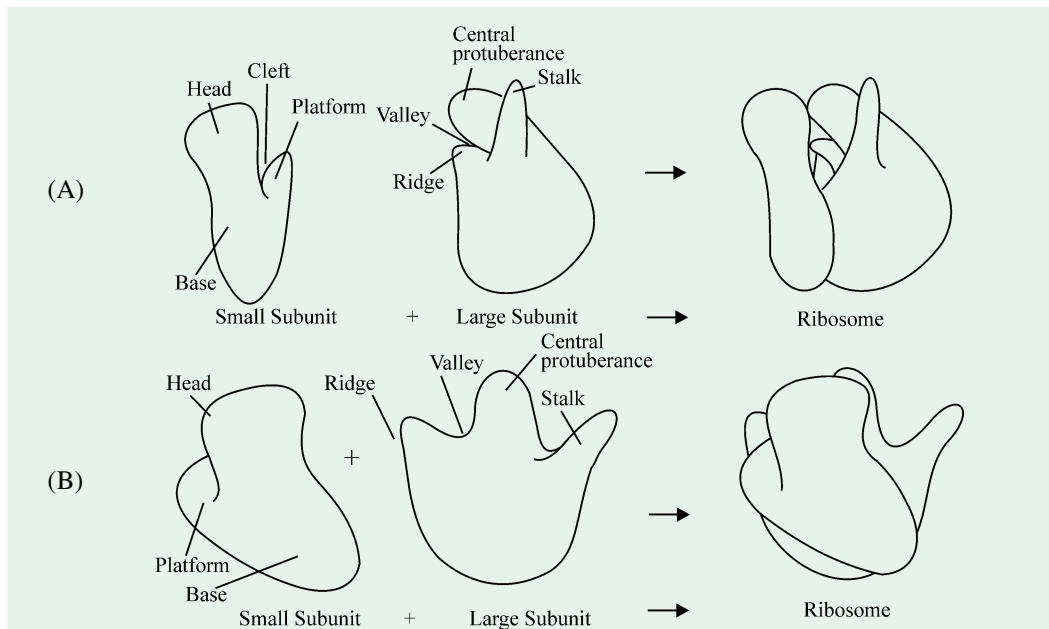
57 (36 + 21) proteins in the bacterial ribosome is believed to play a role in the synthesis of polypeptides, either as an enzyme or as a structural component in the overall process.

The two ribosomal subunits have irregular shapes. The 3-dimensional structure of the 30S and 50S subunits of *E. coli* ribosomes (Fig. 28–8) has been deduced from x-ray diffraction, electron microscopy, and other structural methods. The two oddly-shaped subunits fit together in such a way that a cleft is formed through which the mRNA passes as the ribosome moves along it during the translation process and the newly-formed polypeptide chain emerges from this cleft (Fig. 28–9).



**Fig. 28–8. Components of bacterial and eukaryotic ribosomes**

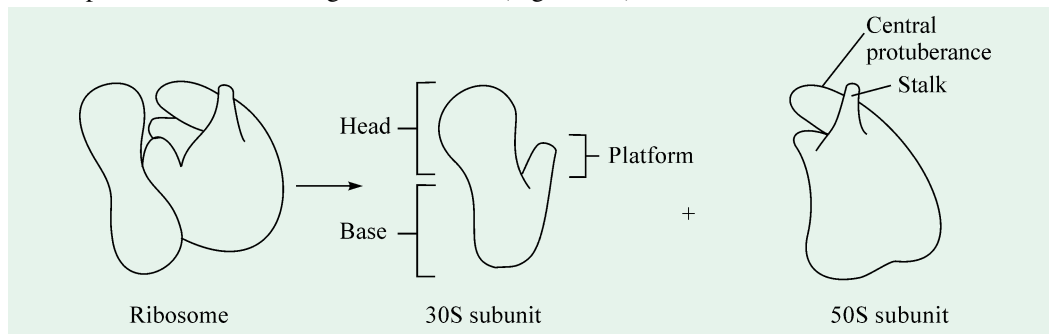
The ribosomal subunits are identified by their S (Svedberg unit) values, sedimentation coefficients that refer to their rate of sedimentation in a centrifuge. The S values (sedimentation coefficients) are not necessarily additive when the two subunits are combined because rates of sedimentation are affected by shape as well as mass.



**Fig. 28-9.** Three-dimensional model of *E.coli* ribosome and its components shown in two different orientations, (A) and (B)

The ribosomes in the cytoplasm of eukaryotic cells (other than mitochondrial and chloroplast ribosomes) are substantially larger and more complex than bacterial ribosomes (Fig. 28–8). They have a diameter of about 23 nm and a sedimentation coefficient of 80S. They also have 2 subunits, which vary in size between species but on an average are 60S and 40S. The small subunit (40S) contains a single 18S rRNA molecule and the large subunit (60S) contains a molecule each of 5S, 5.8S and 28S rRNAs. Altogether, eukaryotic ribosomes contain over 80 different proteins. Thus, an eukaryotic ribosome contains more proteins in each subunit and also has an additional RNA (5.8S) in the larger 60S subunit. The ribosomes of mitochondria and chloroplasts are different from those in the cytoplasm of eukaryotes. They are more like bacterial ribosomes. In fact, there are many similarities between protein synthesis in mitochondria, chloroplasts, and bacteria.

Polypeptide synthesis takes place on the head and platform regions of the 30S subunit and the upper half of the 50S subunit (translational domain). The mRNAs and tRNAs attach to the 30S subunit, and the peptidyl transferase site (where peptide bond formation occur) is associated with the central protuberance of the larger 50S subunit (Fig. 28–10).



**Fig. 28–10.** Schematic of the structures of *Escherichia coli* ribosome and its subunits

The arrow indicates the cleft between the subunits.

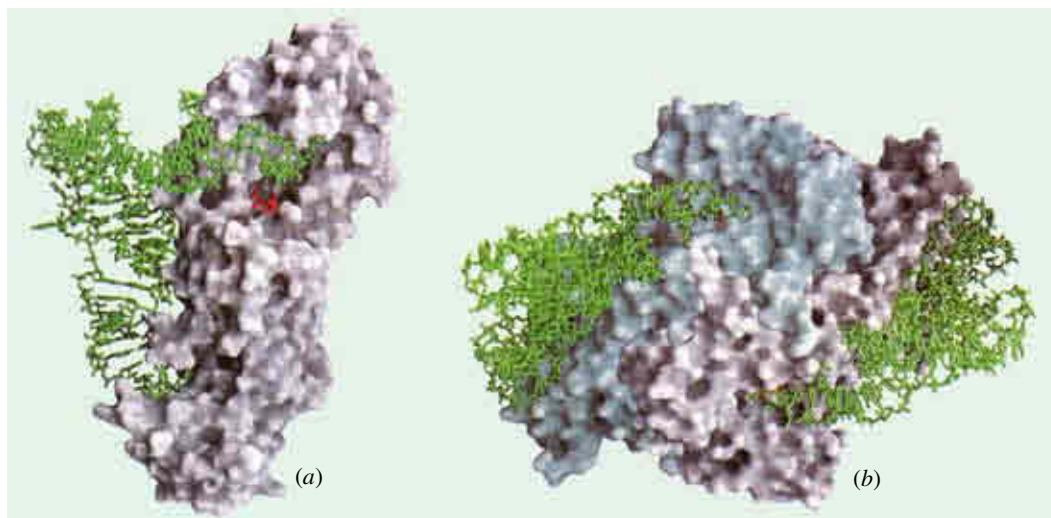
(Adapted from James A. Lake, 1985)

**B. Transfer RNA (tRNA)**

The chemical structure and the role of tRNA has already been described on pages 320-325.

**ACTIVATION OF AMINO ACIDS****A. Two Classes of Aminoacyl-tRNA Synthetases**

As noted earlier, activation of amino acids takes place in the cytosol and in it the 20 different amino acids are esterified to their corresponding tRNAs by aminoacyl-tRNA synthetases. In most organisms, there is generally one aminoacyl-tRNA synthetase (also called aminoacyl-tRNA ligase or simply activation enzyme) for each amino acid. However, for amino acids that have 2 or more corresponding tRNAs, the same aminoacyl-tRNA synthetase usually aminoacylates all of them. However, in *E. coli*, the only exception to this rule is lysine, for which there are two aminoacyl-tRNA synthetases. There is only one tRNA in *E. coli*, and the biological rationale for the presence of two Lys-tRNA synthetases is not yet understood. All the aminoacyl-tRNA synthetases have been divided into 2 classes (Fig. 28-11 and Table 28-2), based on distinctions in their structure and on differences in reaction mechanisms (described below). The overall reaction catalyzed by these enzymes is:

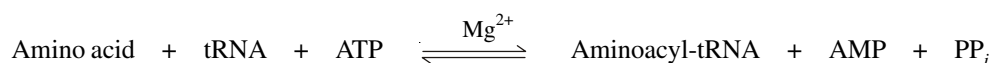


**Fig. 28-11. Aminoacyl-tRNA synthetases**

Both synthetases are complexed with their cognate tRNA (*green stick* structures). Bound ATP (*red*) pinpoints the active site near the end of the aminoacyl arm.

(a) **Gln-tRNA synthetase from *E. coli*.** This is a typical monomeric type I synthetase.

(b) **Asp-tRNA synthetase from yeast.** This is a typical dimeric type II synthetase.



The activation reaction takes place in two separate steps (Fig. 28-12) in the enzyme active site but both the steps are catalyzed by one and the same enzyme, aminoacyl-tRNA synthetase. In the *first step*, the amino acid at the active site interacts with ATP to form an enzyme-bound intermediate, aminoacyl adenylate (= aminoacyl-AMP). The reaction is analogous to that of the activation of fatty acids where an acyl-AMP is formed. In this reaction, the carboxyl group of the amino acid is bound in anhydride linkage with the 5'-phosphate group of the AMP with displacement of pyrophosphate. In the *second step*, the aminoacyl group is transferred from enzyme-bound aminoacyl-AMP to its

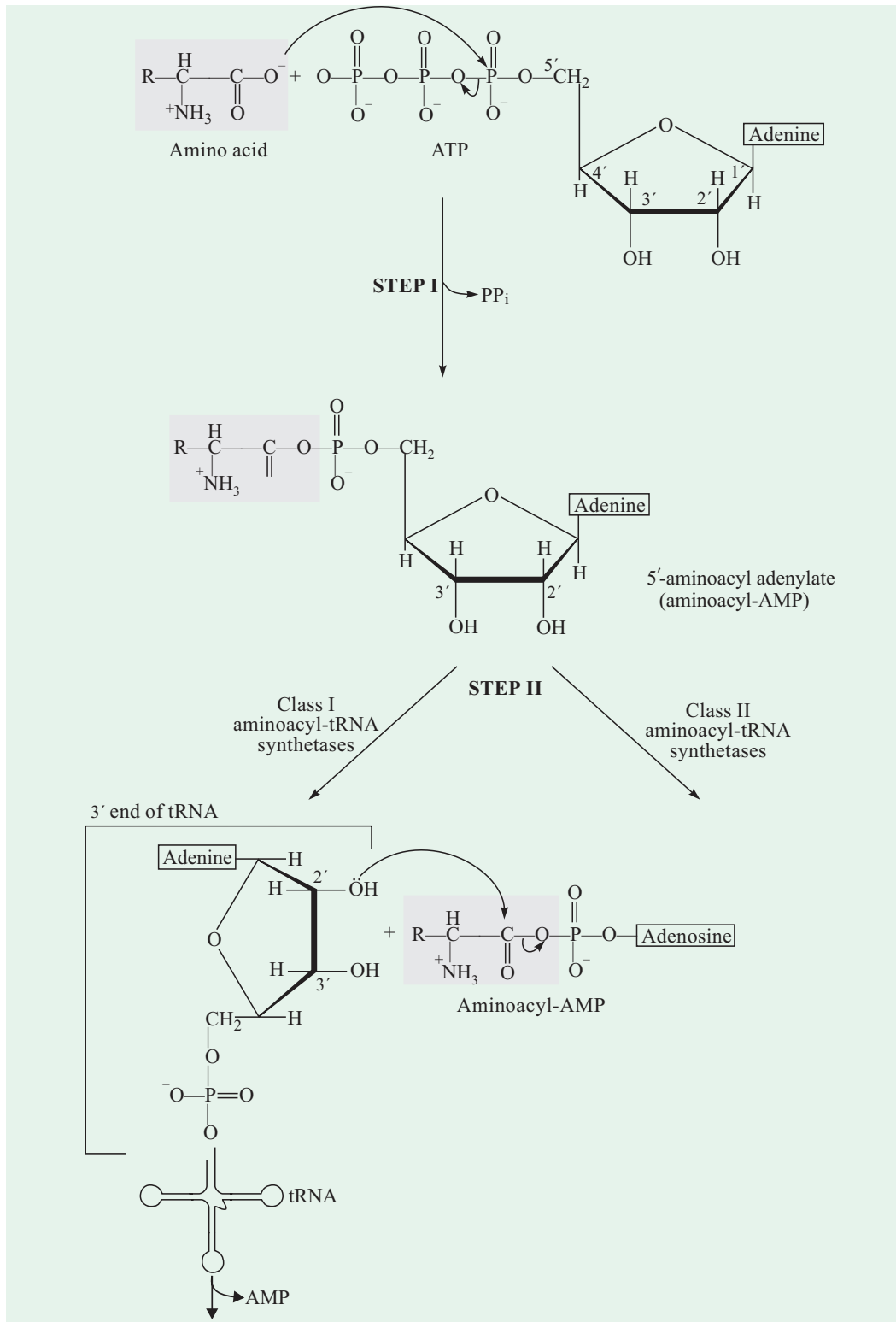
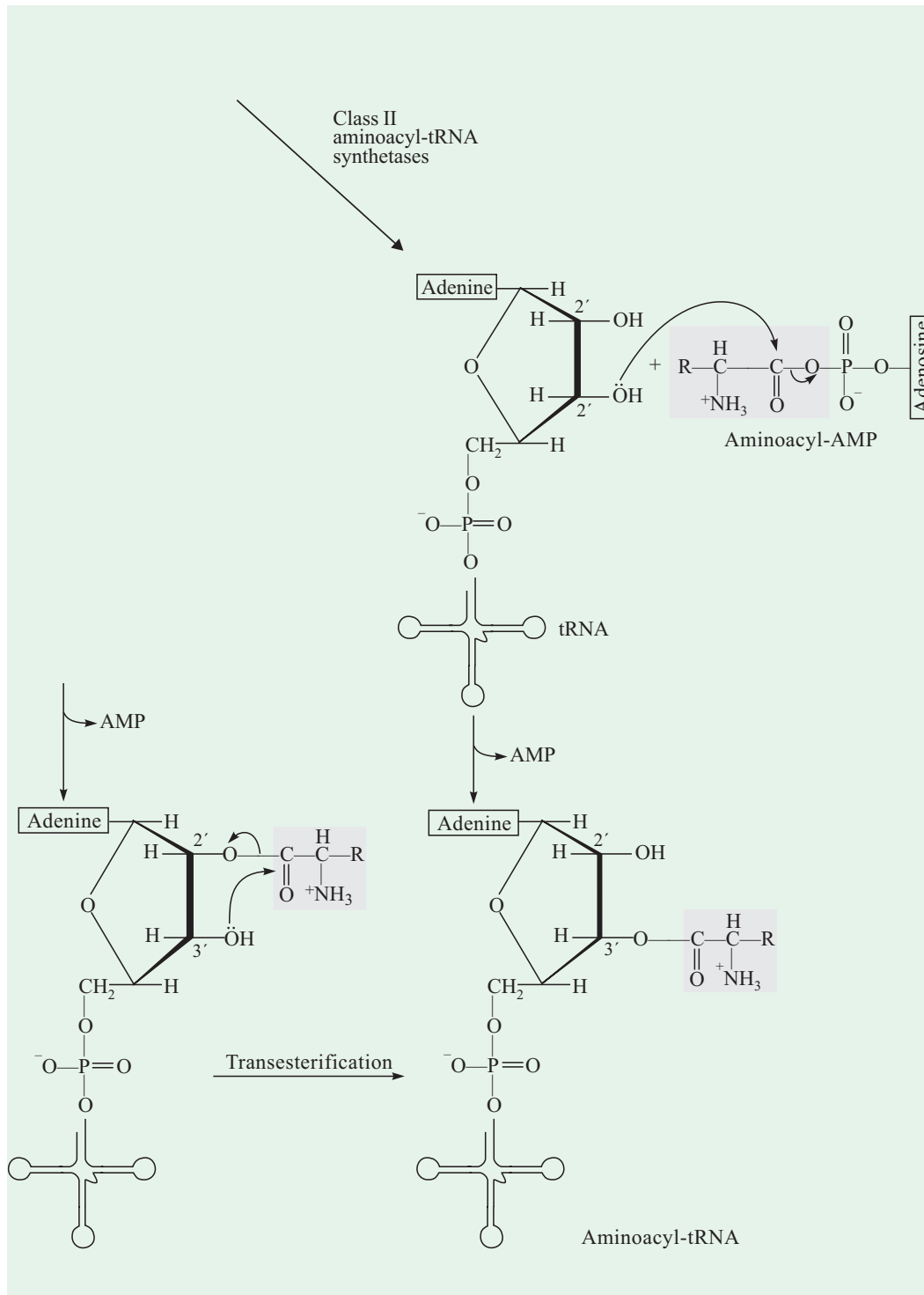


Fig. 28-12. (Cont'd.)



**Fig. 28-12. Aminoacylation of a tRNA by aminoacyl-tRNA synthetases**

(Adapted from Lehninger, Nelson and Cox, 1993)



Table 28–2. Two classes of aminoacyl-tRNA synthetases

Class I	Class II
Arginine	Alanine
Cysteine	Asparagine
Glutamic acid	Aspartic acid
Glutamine	Glycine
Isoleucine	Histidine
Leucine	Lysine
Methionine	Phenylalanine
Tryptophan	Proline
Tyrosine	Serine
Valine	Threonine

corresponding specific tRNA. The mechanism of this step is somewhat different for the 2 classes of aminoacyl-tRNA synthetase (Fig. 28–12). For **class I enzymes**, the aminoacyl group is transferred initially to the 2'-hydroxyl group of the 3'-terminal adenylate residue, then moved to the 3'-hydroxyl by a transesterification reaction. For **class II enzymes**, the aminoacyl group is transferred directly to the 3'-hydroxyl of the terminal adenylate. The reason for this mechanistic difference between the two enzyme classes is not known. The resulting ester linkage between the amino acid and the tRNA (Fig. 28–13) has a high standard free energy of hydrolysis ( $\Delta G^{0'} = -29$  kJ/mol). The pyrophosphate formed in the activation reaction undergoes hydrolysis to phosphate by inorganic phosphatase. Thus, *two* high-energy phosphate bonds are consumed in the activation of each amino acid molecule. One of the bonds is consumed in forming the ester linkage of aminoacyl-tRNA, whereas the other is consumed in driving the reaction forward. This renders the overall reaction for amino acid activation essentially irreversible :

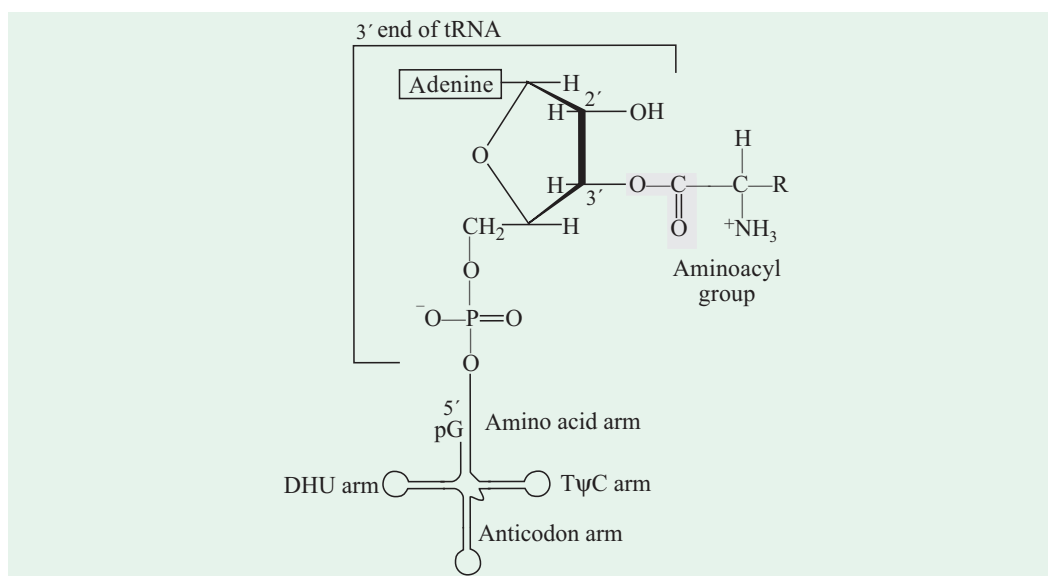
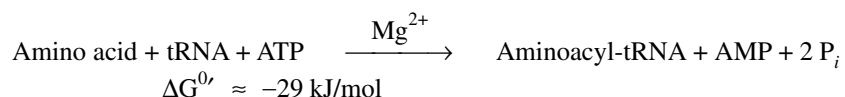


Fig. 28–13. General structure of aminoacyl-tRNAs

The aminoacyl group is esterified to the 3'-position of the terminal adenylate residue. The ester linkage, that both activates the amino acid and joins it to the tRNA, is shown enclosed in a shaded area.

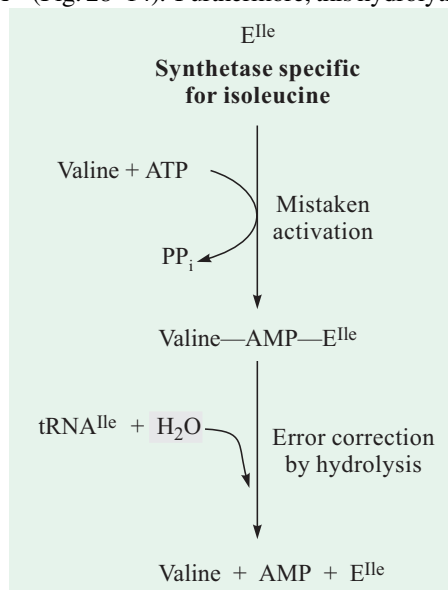
### B. Proofreading by Some Aminoacyl-tRNA Synthetases

The aminoacylation of tRNA performs two functions: the activation of an amino acid for peptide bond formation and attachment of the amino acid to an adaptor tRNA which directs its placement within a growing polypeptide. In fact, the identity of the amino acid attached to a tRNA is not checked on the ribosome and attaching the correct amino acid to each tRNA is, henceforth, essential to the fidelity of protein synthesis as a whole.

The correct translation of genetic message depends on the high degree of specificity of aminoacyl-tRNA synthetases. These enzymes are highly sensitive in their recognition of the amino acid to be activated and of the prospective tRNA acceptor. A very high specificity is indeed necessary during the Stage I (*i.e.*, activation of amino acids), in order to avoid errors in the biosynthesis of proteins, because once the aminoacyl-tRNA is formed, there is no longer any control mechanism in the cell to verify the nature of the amino acid and it is therefore not possible to reject an amino acid which would have been incorrectly bound. Consequently, the amino acid would be erroneously incorporated in the protein molecule. Therefore, the aminoacyl-tRNA synthetases, *in vivo*, must either not commit any errors or be able to rectify them.

The tRNA molecules, that accept different amino acids, have differing base sequences, and hence they can be easily recognized by the synthetases. But the synthetases must, in particular, be able to distinguish between two amino acid of very similar structure; and some of them have this capacity. For example, isoleucine (Ile) differs from valine (Val) only in having an additional methylene ( $-\text{CH}_2-$ ) group. The additional binding energy contributed by this extra  $-\text{CH}_2-$  group favours the activation of isoleucine (to form Ile-AMP) over valine by a factor of 200. The concentration of valine *in vivo* is about 5 times that of isoleucine, and so valine would be mistakenly incorporated in place of isoleucine 1 in every 40 times.

However, the observed error frequency *in vivo* is only 1 in 3,000, indicating that there must be subsequent editing steps to enhance fidelity. In fact, the Ile-tRNA synthetase corrects its own errors, *i.e.*, in the presence of tRNA<sup>Ile</sup>, the Val-AMP formed is hydrolyzed (but not Ile-AMP), thus preventing an erroneous aminoacylation (*i.e.*, a misacylation) of tRNA<sup>Ile</sup> (Fig. 28–14). Furthermore, this hydrolytic reaction frees the synthetase for the activation and transfer of Ile, the correct amino acid. Hydrolysis of Ile-AMP, the desired intermediate, is however avoided because the hydrolytic site is just large enough to accommodate Val-AMP, but too small to allow the entry of Ile-AMP. Thus, we see that most aminoacyl-tRNA synthetases contain two sites: the acylation or synthetic site and the hydrolytic site. And the entire system is forced through two successive “filters”, rather than one, whereby increasing the potential fidelity by a power of 2. The *first filter* is the **synthetic site** on **synthetases** which brings about the initial amino acid binding and activation to aminoacyl-AMP. The *second filter* is the separate active site or **hydrolytic site** on synthetases which catalyzes deacylation of incorrect aminoacyl-AMPs. The synthetic site rejects amino acids that are larger than the correct one because there is insufficient room for them, whereas the hydrolytic site destroys activated intermediates that are smaller than the correct species. Hydrolytic proofreading is central to the fidelity of many aminoacyl-tRNA synthetases, as it is to DNA polymerases. In addition to proofreading after formation of the aminoacyl-AMP intermediate, most aminoacyl-tRNA synthetases



**Fig. 28–14. Correction mechanism or proofreading on hydrolytic site, the “second filter”**

The entry of tRNA specific for isoleucine induces hydrolysis of valyl-AMP.

are also capable of hydrolyzing the ester linkage between amino acids and tRNAs in aminoacyl-tRNAs. This hydrolysis is greatly accelerated for incorrectly-charged tRNAs, providing yet a *third filter* to further enhance the fidelity of the overall process. In contrast, in a few aminoacyl-tRNA synthetases that activate amino acids that have no close structural relatives, little or no proofreading occurs; in these cases, the active site can sufficiently discriminate between the proper amino acid and incorrect amino acids. *Proofreading is costly in energy and time and hence is selected in the course of evolution only when fidelity must be enhanced.*

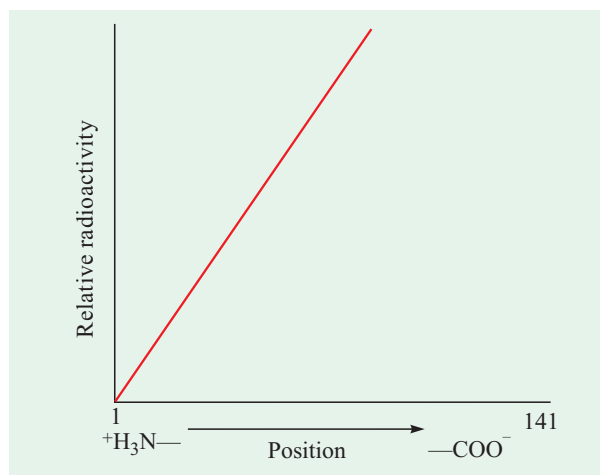
The overall error rate of protein synthesis ( $\sim 1$  mistake per  $10^4$  amino acids incorporated) is not nearly as low as for DNA replication, perhaps because a mistake in a protein is erased by destroying the protein and is not passed onto future generations. This degree of fidelity is sufficient to ensure that most proteins contain no mistakes and that the large amount of energy required to synthesize a protein is rarely wasted.

### C. The “Second Genetic Code”

An individual aminoacyl-tRNA synthetase must be specific not only for a single amino acid but for a certain tRNA as well. Discriminating among several dozen tRNAs is just as important for the overall fidelity of protein synthesis as is distinguishing among amino acids. The interaction between aminoacyl-tRNA synthetases and tRNAs has been referred to as the “**second genetic code**”, to reflect its crucial role in maintaining the accuracy of protein synthesis. The “coding” rules are apparently more complex than those in the “first” code.

### D. Direction of the Growth of Polypeptide Chain

The direction of synthesis of polypeptide chain was determined by Howard Dintzis (1961) who followed, as a function of time, the course of incorporation of tritiated leucine ( $^3\text{H}$ -leucine) in the  $\alpha$  and  $\beta$  chains of the hemoglobin, synthesized by reticulocytes (immature erythrocytes) in suspension (at low temperature to slow down protein synthesis). At regular intervals (between 4 and 60 minutes), he isolated hemoglobin, separated the  $\alpha$  and  $\beta$  chains, subjected them to a hydrolysis by trypsin and fractionated the peptides obtained. *Leucine was chosen because it occurs frequently along both the  $\alpha$ - and  $\beta$ -globin chains.* It was found that there was more radioactivity in the peptides near the carboxyl end than in those near the amino end. Thus, there was a gradient of increasing radioactivity from the amino to the carboxyl end of each chain (Fig. 28–15). The carboxyl end was the most heavily labelled because it was the last to be synthesized. Hence, the direction of chain growth is always from the amino to carboxyl end. In other words, *the proteins are synthesized in the amino-to-carboxyl direction.*



**Fig. 28–15. Distribution of  $^3\text{H}$ -leucine in a chains of hemoglobin synthesized after exposure of reticulocytes to tritiated leucine**

Note that each a chain of hemoglobin contains 141 amino acid residues. The higher radioactivity of the carboxyl ends, relative to the amino ends, indicates that the carboxyl end of each chain was synthesized last.

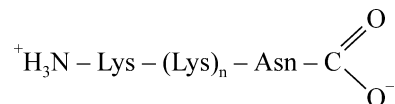
(After Lubert Stryer, 1995)

## INITIATION OF PROTEIN SYNTHESIS

### A. Translation of Messenger RNA in 5' → 3' Direction

The direction of reading of mRNA was determined by using the **synthetic polynucleotide** as the  $5' \text{A} - \text{A} - \text{A} - (\text{A} - \text{A} - \text{A})_n - \text{A} - \text{A} - \text{C} 3'$  template in a cell-free protein-synthesizing system. The triplet AAA codes for lysine (Lys), whereas AAC codes for asparagine (Asn).

The **polypeptide product** was

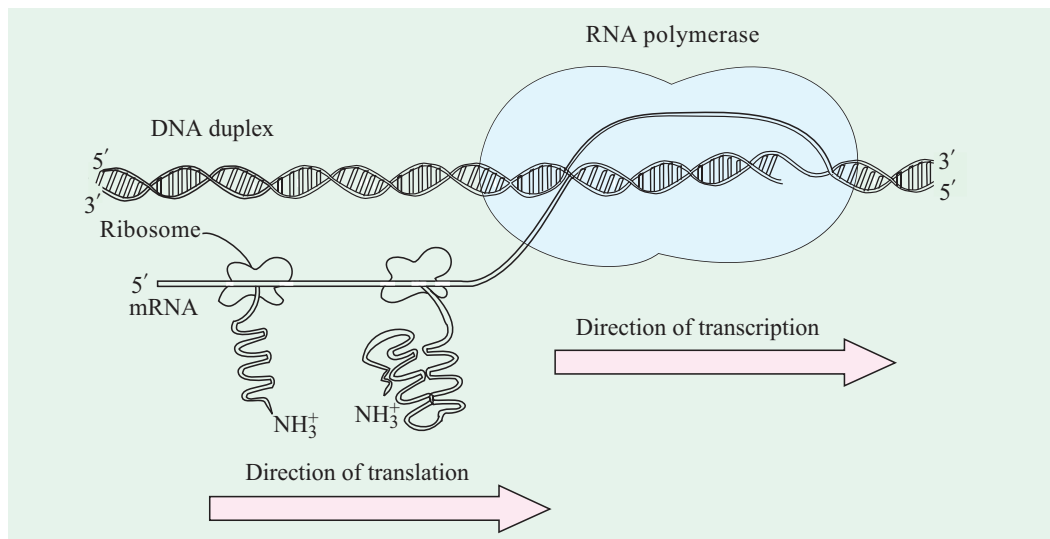


Because asparagine was the carboxyl terminal residue, the codon AAC was the last to be read. Hence, the direction of translation of mRNA is 5' → 3'.

### B. Coupling between Transcription and Translation in Bacteria

In bacteria, transcription of mRNA is closely coupled to its translation. Messenger RNAs are synthesized (or transcribed) in one direction and are translated also in the same direction. As such, the ribosomes begin translating the 5' end of the mRNA before its synthesis (or transcription) is complete (Fig. 28–16). This is possible because the mRNA in bacteria does not have to be transported from a nucleus to the cytoplasm before encountering ribosomes. The situation is somewhat different in eukaryotes, where newly-transcribed mRNAs must be transferred out of the nucleus before they can be translated.

Bacterial mRNAs usually exist for only a few minutes (half-life of bacterial mRNA is about 1.5 minutes) before they are degraded by nucleases. Therefore, in order to maintain high rates of protein synthesis, the mRNA for a given protein or a set of proteins must be formed continuously and translated with utmost efficiency. The short life of mRNA in bacteria allows synthesis of a protein to stop rapidly when it is no longer required by the cell.



**Fig. 28–16. The coupling of transcription and translation in bacteria**

The mRNA is translated by ribosome(s) while it is still being transcribed from DNA by RNA polymerase. This is possible because the mRNA in bacteria does not have to be transported from a nucleus to the cytoplasm before encountering ribosomes. In this schematic diagram, the ribosomes are depicted as smaller than the RNA polymerase. In reality, the ribosomes (MW =  $2.5 \times 10^6$ ) are larger than the RNA polymerase (MW =  $3.9 \times 10^5$ ).

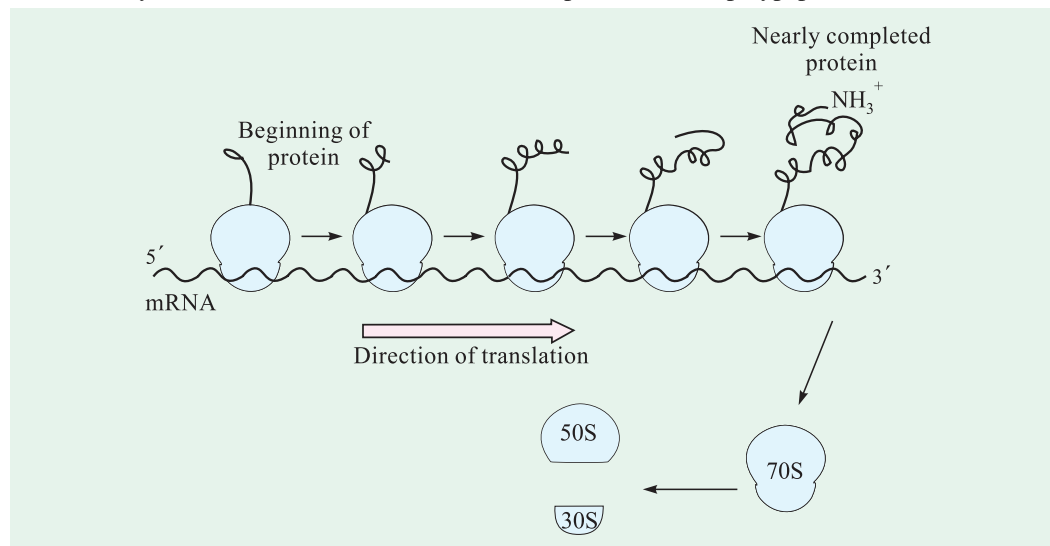
### C. Polyribosomes (or Polysomes or Ergosomes)

Many ribosomes can simultaneously translate an mRNA molecule. This parallel synthesis markedly increases the efficiency of utilization of the mRNA. The group of ribosomes bound to an mRNA molecule is called a **polyribosome** or simply **polysome** or rarely an **ergosome**. The ribosomes in this unit operate independently, each synthesizing a complete polypeptide chain. In optimal conditions for protein synthesis, there is an average one ribosome for every 80 nucleotides. In general, there are less than 10 ribosomes on the eukaryotic mRNA (which are usually *monocistronic*) while there can be several tens on prokaryotic mRNAs (which are often *polycistronic*). The polyribosomes synthesizing hemoglobin (which contains about 145 amino acids per chain, or about 500 nucleotides per mRNA) typically consist of 5 ribosomes. Ribosomes closest to the 5' end of the messenger RNA have the smallest polypeptide chains, whereas those nearest the 3' end of the mRNA have the longest chains that have almost finished synthesis (Fig. 28–17). Ribosomes dissociate into 30S and 50S subunits after the polypeptide chain is released. Polyribosomes, thus, allow rapid translocation of a single message.

### D. N-formylmethionine as Initiator of Protein Synthesis in Bacteria

The simplest possibility as to how the protein synthesis starts may be that the first three nucleotides of each mRNA serve as the first codon; no special start signal would then be needed. However, the experimental evidences indicate that translation does not start immediately at the 5' end of mRNA. In fact, the first translated codon is usually more than 25 nucleotides away from the 5' end. Furthermore, many mRNA molecules in prokaryotes are *polycistronic*, *i.e.*, they code for 2 or more polypeptide chains. These facts and findings have led to the conclusion that *all known mRNA molecules contain signals that signify the beginning and end of each encoded polypeptide chain*.

An important finding was that nearly half of the amino-terminal residues of proteins in *E. coli* are methionine, yet this amino acid is uncommon at other positions of the polypeptide chain. Furthermore,



**Fig. 28–17. A polyribosome (= polysome)**

Five ribosomes are shown translating a mRNA molecule simultaneously, moving from the 5' end to 3' the end of mRNA. They function independently of each other.

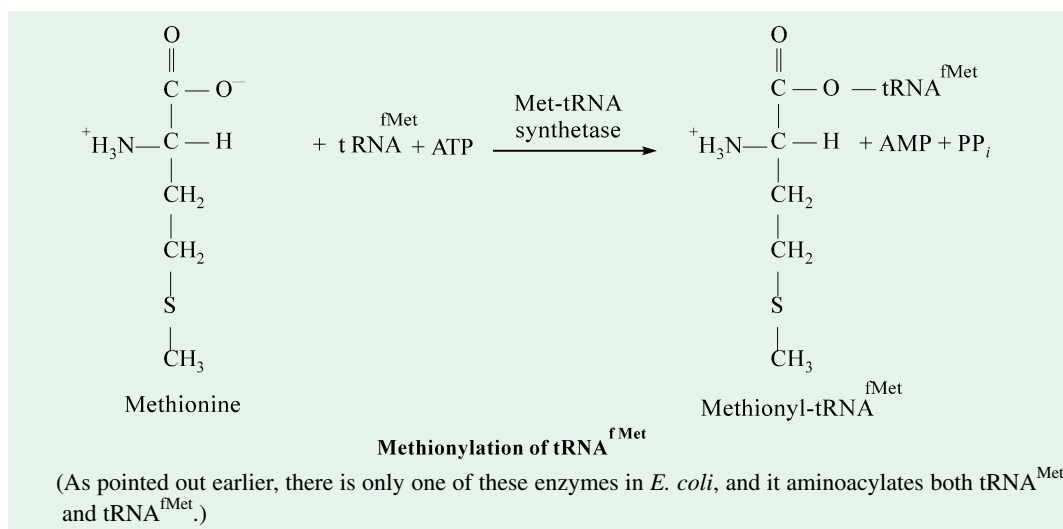
the amino end of nascent proteins is usually modified. These facts provided the clue that a derivative of methionine participates in initiation of polypeptide synthesis. In fact, *protein synthesis in bacteria starts with N-formyl methionine (fMet)*.



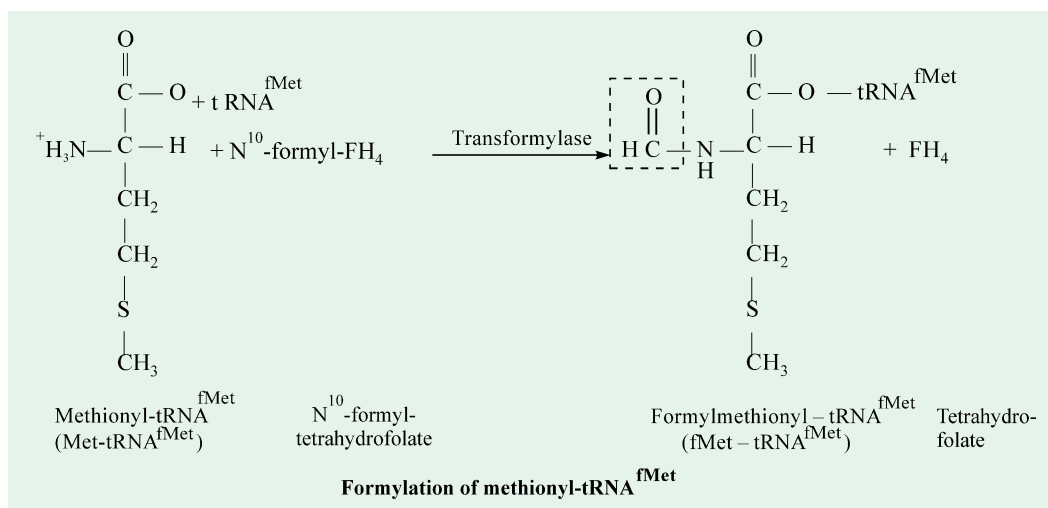
## 736 FUNDAMENTALS OF BIOCHEMISTRY

Although there is only one codon (AUG) for methionine, there are 2 tRNAs for methionine in all organisms and both have a 3' U–A–C 5' anticodon that recognizes the unique 5' A–U–G 3' codon for methionine. One tRNA, designated as 'internal' tRNA or  $\text{tRNA}^{\text{Met}}$ , acts as an adaptor for methionine residues required within the polypeptide chain; the second tRNA, designated as initiator tRNA or  $\text{tRNA}^{\text{fMet}}$  (the superscript f indicates that methionine attached to the initiator tRNA can be formylated) has a unique role in the initiation of the polypeptide chain synthesis. As already stated, the starting amino acid residue at amino-terminal is *N*-formylmethionine (fMet). It combines with  $\text{tRNA}^{\text{fMet}}$  to produce *N*-formylmethionyl-tRNA<sup>fMet</sup> (fMet–tRNA<sup>fMet</sup>) in two successive steps :

**First**, methionine is attached to  $\text{tRNA}^{\text{fMet}}$  by the Met-tRNA synthetase :



**Second**, a formyl group is transferred to the amino group of the Met residue from *N*<sup>10</sup>-formyltetrahydrofolate by a transformylase enzyme :



(The transferred formyl ( $-\overset{\text{H}}{\text{C}}=\text{O}$ ) group in fMet–tRNA<sup>fMet</sup> is shown enclosed by broken lines.)

This transformylase is more selective than the Met-tRNA synthetase, and it cannot formylate free methionine or Met residues attached to  $\text{tRNA}^{\text{Met}}$ . Instead, it is specific for Met residues attached

to tRNA<sup>Met</sup>, presumably recognizing some unique structural feature of that tRNA. The other Met-tRNA species, Met-tRNA<sup>Met</sup>, is used to insert methionine in interior positions in the polypeptide chain. Blocking of the amino group of methionine by the *N*-formyl group not only prevents it from entering interior positions but also allows fMet-tRNA<sup>fMet</sup> to be bound at a specific initiation site on the ribosome that does not accept Met-tRNA<sup>Met</sup> or any other aminoacyl-tRNA. In prokaryotes, the *N*-formylation of the Met residue on the tRNA seems to deceive the P site, leaving the A site free (refer the following Section for the two sites).

In eukaryotic cells, however, all peptides synthesized by cytosolic ribosomes begin with a Met residue (in contrast to fMet in bacteria), but again a specialized initiating tRNA is used that is distinct from the tRNA<sup>Met</sup> used at interior positions. Interestingly, one of the two types of methionyl-tRNA<sup>Mets</sup> of eukaryotes can be formylated by the *E. coli* formylase. It appears that, during the course of evolution, eukaryotes lost the ability to produce the enzyme, but one of the two methionyl-tRNAs retained the vestigial ability to be formylated. In contrast, polypeptides synthesized by the ribosomes in the mitochondria and chloroplasts of eukaryotic cells begin with *N*-formylmethionine. This and other similarities in the protein-synthesizing machinery of these organelles and bacteria strongly support the view that mitochondria and chloroplasts originated from bacterial ancestors symbiotically incorporated into the precursors of eukaryotic cells at an early stage of evolution.

### E. The Three Steps of Initiation Process

The following description of the second stage of protein synthesis, *i.e.*, initiation (and also of elongation and termination stages to follow) pertains to the protein synthesis in bacteria; the process is not well-understood in eukaryotes. The initiation of polypeptide synthesis in bacteria requires 7 components (Table 28–1): the 30S ribosomal subunit (which contains 16S rRNA), the mRNA coding for the polypeptide to be made, the initiating fMet-tRNA<sup>fMet</sup>, a set of 3 proteins called initiation factors (IF-1, IF-2, IF-3), GTP, the 50S ribosomal subunit, and Mg<sup>2+</sup>. The formation of initiation complex takes place in 3 steps (Fig. 28–18) :

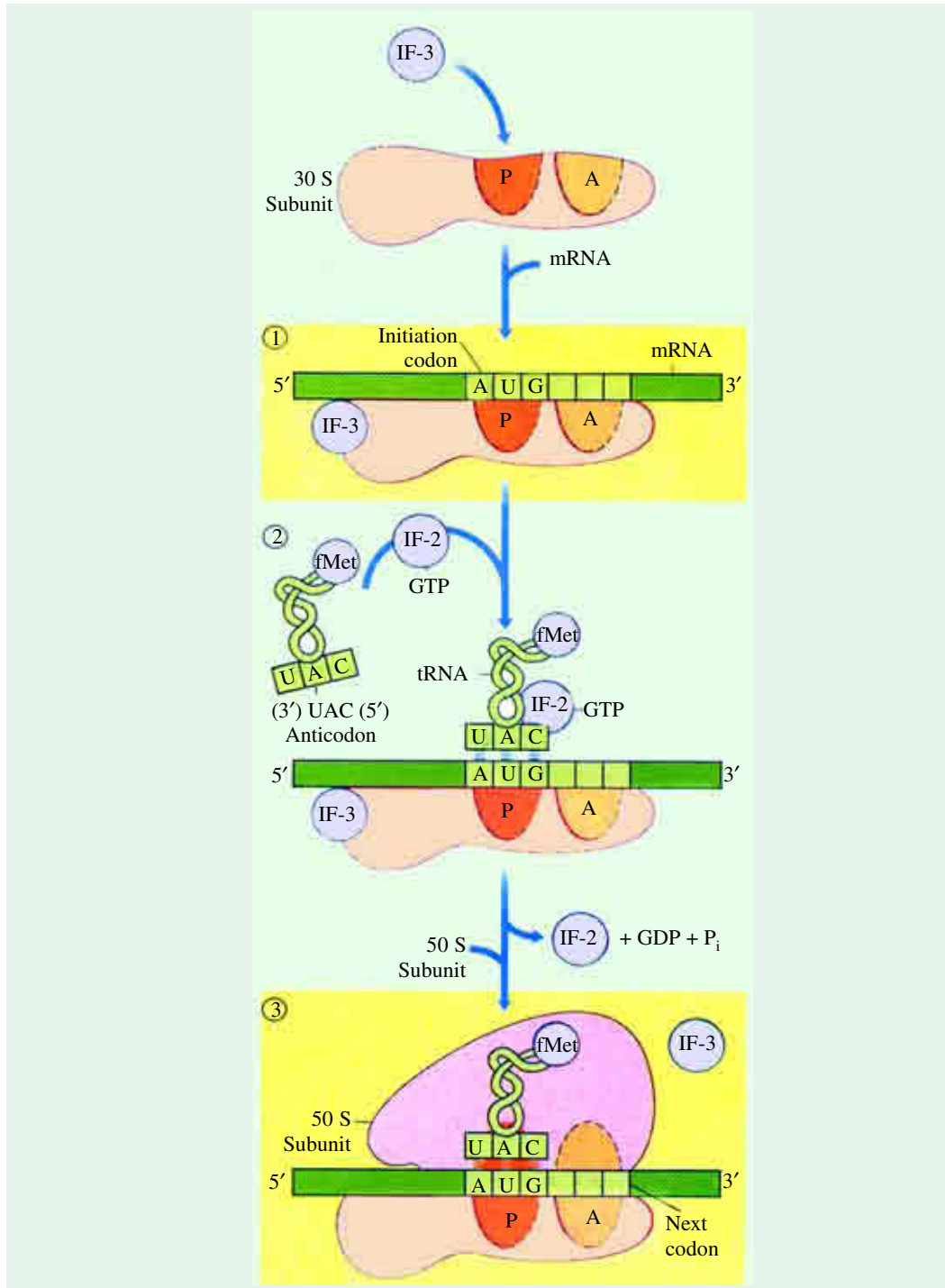
(i) In the **first step**, the 30S ribosomal subunit binds initiation factor 3 (IF-3), which prevents the 30S and 50S subunits from combining prematurely to form a dead-end 70S complex, devoid of mRNA. The mRNA binds to the 30S subunit in such a manner that the initiation codon (AUG) binds to a precise location of the 30S subunit.

Each ribosome has 2 binding sites for tRNA : the **aminoacyl** or **A site** for the incoming aminoacyl-tRNA and the **peptidyl** or **P site** for the growing polypeptidyl-tRNA. The A site is also called acceptor site, while the P site is also called **donor** or **D site**. Both the 30S and 50S subunits contribute to the characteristics of each site. The initiating AUG codon resides in the P site, which is the only site to which fMet-tRNA<sup>fMet</sup> can bind. However, fMet-tRNA<sup>fMet</sup> is the exception : during the subsequent elongation stage, all other incoming aminoacyl-tRNAs, including the Met-tRNA<sup>Met</sup> that binds to interior AUGs, bind to the A site. The P site is the site from which the ‘uncharged’ tRNAs leave during elongation of the polypeptide chain.

(ii) In the **second step** of initiation process, the complex consisting of 30S subunit, IF-3, and mRNA now forms a still larger complex by binding IF-2 which already is bound to GTP and the initiating fMet-tRNA<sup>fMet</sup>. The anticodon of this tRNA (UAC) pairs correctly with the initiation codon in this step. The protein factor IF-1, however, stimulates the activities of both IF-2, and IF-3.

(iii) In the **third step**, this large complex combines with the 50S ribosomal subunit. Simultaneously, the GTP, which is bound to IF-2, is hydrolyzed to GDP and Pi. Both GDP and Pi are released and IF-2 and IF-3 also dissociate from the ribosome.

These 3 steps in the initiation of bacterial protein synthesis result in a functional 70S ribosome called the **initiation complex**, containing the mRNA and the initiating fMet-tRNA<sup>fMet</sup>. The correct binding of the fMet-tRNA<sup>fMet</sup> to the P site in the complete 70S initiation complex is assured by two



**Fig. 28-18. Three steps of the formation of initiation complex**

The 3 steps are driven at the expense of the hydrolysis of GTP to GDP and  $P_i$ . The initiation factor IF-1, not shown here, stimulates the activities of both IF-2 and IF-3.

A = Aminoacyl site; P = Peptidyl site

points of recognition and attachment: the codon–anticodon interaction involving the initiating AUG fixed in the P site and binding interactions between the P site and the fMet-tRNA<sup>fMet</sup>. The initiation complex is now ready for the elongation process.

**Table 28–3. The roles of bacterial and eukaryotic initiation factors in protein synthesis**

<i>Bacteria</i>		<i>Eukaryotes</i>	
<i>Factor</i>	<i>Function</i>	<i>Factor</i>	<i>Function</i>
IF-1	Stimulates activities of IF-2 and IF-3	eIF2*	Facilitates binding of initiating Met-tRNA <sup>Met</sup> to 40S ribosomal subunit
IF-2	Facilitates binding of fMet-tRNA <sup>fMet</sup> to 30S ribosomal subunit	eIF3	First factors to bind 40S subunit; Facilitate subsequent steps
		eIF4C	
IF-3	Binds to 30S subunit; Prevents premature association of 50S subunit	CBP I	Binds to 5' cap of mRNA
		eIF4A	Bind to mRNA;
		eIF4B	Facilitate scanning of mRNA to locate first AUG
		eIF4F	
		eIF5	Promotes dissociation of many other initiation factors from 40S subunit as prelude to association of 60S subunit to form 80S initiation complex
		eIF6	Facilitates dissociation of inactive 80S ribosome into 40S and 60S subunits

Amazingly, the eIF2 appears to be a multifunctional proteins. Besides its role in the initiation of translation process, it also helps in the splicing of mRNA precursors in the nucleus. This finding provides an intriguing link between transcription and translation in eukaryotic cells.

(Adapted from Lehninger, Nelson and Cox, 1993)

One of the important differences between the protein synthesis in prokaryotes and eukaryotes is the presence of 9 eukaryotic initiation factors, most of them abbreviated as eIF (*e* for eukaryotic). One of these factors, called cap binding protein or CBPI, binds to the 5' cap of mRNA. CPBI facilitates formation of a complex between the mRNA and the 40S ribosomal subunit. The mRNA is then scanned to locate the first AUG codon, which signals the start of the reading frame. Many additional initiation factors are required in this *mRNA scanning reaction* and also in assembly of the complete 80S initiation complex in which the initiating Met-tRNA<sup>Met</sup> and mRNA are bound and are ready for elongation to proceed. The roles of various initiation factors in protein synthesis in both prokaryotes and eukaryotes are summarized in Table 28–3.

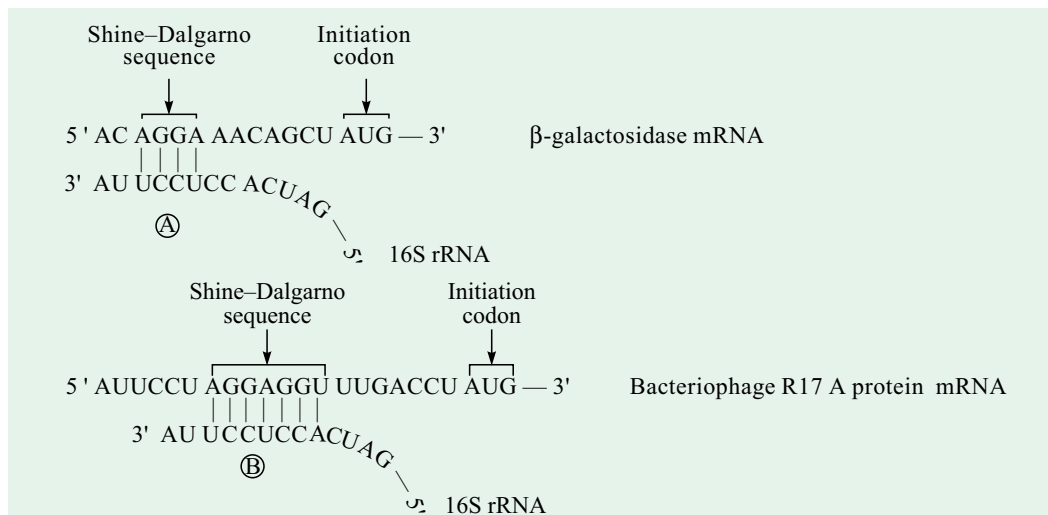
The terms **prokaryotes** and **eukaryotes** were first proposed by Edward Chatton in 1937.

#### F. Shine-Dalgarno Sequence

For the production of a 30S initiation complex in the first step of translation initiation, the 3' terminal region of the 16S RNA component of the 30S ribosomal subunit is the binding site for the mRNA. This 3' terminal region of 16S rRNA component is pyrimidine-rich and hydrogen-bonds, in an antiparallel fashion, with a complementary sequence of purine-rich region on the 5' side (upstream) of the initiating AUG codon on mRNA (Fig. 28–19). This purine-rich region, consisting of 4 to 9 residues upstream from the AUG codon on mRNA, is called **Shine-Dalgarno sequence** and is needed for the interaction of the mRNA with its *ribosomal binding site*. This particular pattern of mRNA–16S rRNA binding aligns the initiating AUG codon of the message (or the protein to be formed) for the proper binding with the anticodon (UAC) in the 30S–fMet-tRNA<sup>fMet</sup> complex. Thus, the specific

## 740 FUNDAMENTALS OF BIOCHEMISTRY

AUG codon where fMet-tRNA<sup>fMet</sup> is to be bound is thereby distinguished from interior methionine codons (AUGs) by its proximity to the Shine-Dalgarno sequence in the mRNA. This very well explains as to how the single codon, 5'-AUG serves to identify both the starting *N*-formylmethionine (or methionine in the case of eukaryotes) and those methionine residues that occur in interior



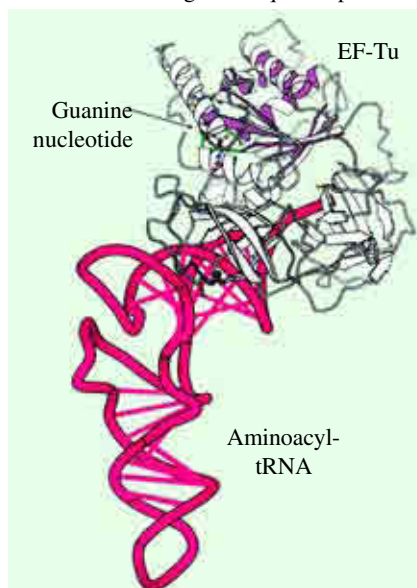
**Fig. 28-19.** Binding sequence of two mRNAs (Shine-Dalgarno sequence) which base-pairs with a sequence near the 3' end of the 16S rRNA

The portions of the two mRNA transcripts shown here code for b-galactosidase A and the A protein of R17 phage B. The AUG codon defines the start of the polypeptide chain. Shine-Dalgarno sequence pairs with 16S rRNA, whereas the initiator codon pairs with fMet-tRNA<sup>fMet</sup>, *i.e.*, initiator tRNA.

positions in the polypeptide chains. In a polycistronic mRNA, containing 2 or more messages, there is a ribosomal binding site to the 5' side of each message; hence, such mRNAs can accommodate as many initiation complexes as there are messages to be translated. It may, thus, be inferred that two kinds of interactions determine where protein synthesis starts: the pairing of mRNA bases with the 3' end of 16S rRNA, and the pairing of the initiator codon on mRNA with the anticodon of fMet initiator tRNA.

### ELONGATION OF THE POLYPEPTIDE CHAIN

The third stage of protein synthesis is elongation, *i.e.*, the stepwise addition of amino acids to the polypeptide chain. The process of elongation requires 4 components: the initiation complex (described above), the next aminoacyl-tRNA specified by the next codon in the mRNA, a set of 3 soluble cytosolic proteins called *elongation factors* (EF-Tu, EF-Ts, EF-G), and GTP. Three steps take place in the addition of each amino acid residue, and this cycle is repeated as many times as there are amino acid residues to be added. The three steps involved (Figs. 28-22, 28-23, 28-24) are :



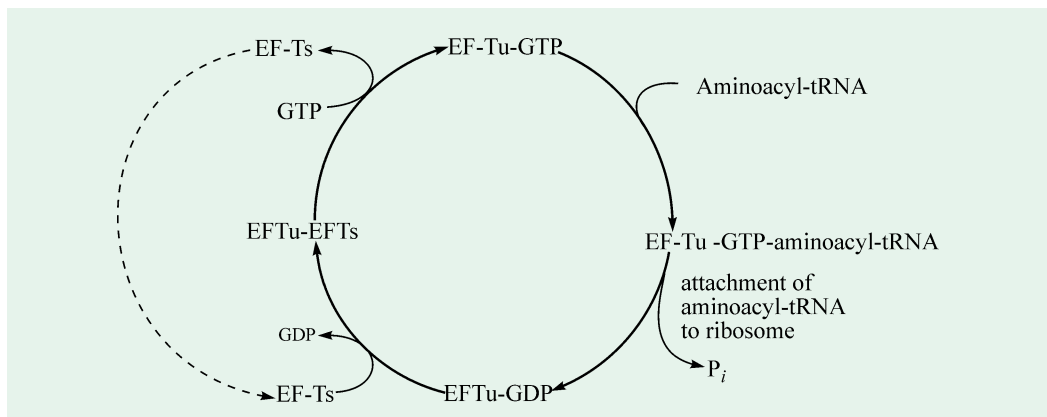
**Fig. 28-20.** Structure of elongation factor, Tu

The structure of a complex between elongation factor Tu (EF-Tu) and an aminoacyl-tRNA. The amino terminal domain of EF-Tu is a P-loop NTPase domain, similar to those in other G proteins.



### Step 1: Codon recognition (= Binding of an aminoacyl-tRNA at site A of ribosome)

In this first step (Fig. 28–22), the next aminoacyl-tRNA (aa-tRNA) is first bound to a complex of EF-Tu (Tu for temperature-unstable) (Fig. 28–20) containing a molecule of bound GTP. The resulting aminoacyl-tRNA – EF-Tu · GTP complex is then bound to the A site of the 70S initiation complex. The GTP is hydrolyzed to GDP and  $P_i$ , and an EF-Tu · GDP complex leaves the 70S ribosome. The bound GDP is dissociated when the EF-Tu · GDP complex binds to the second elongation factor, EF-Ts (Ts for temperature-stable), and EF-Ts is subsequently released when another molecule of GTP becomes bound to EF-Tu. EF-Tu containing bound GTP is now ready to pick up another aminoacyl-tRNA and deliver it to the A site of the ribosome (Fig. 28–21).



**Fig. 28.21. Reaction cycle of elongation factor Tu (EF-Tu)**

The diagram shows the regeneration of EF-Tu · GTP complex.

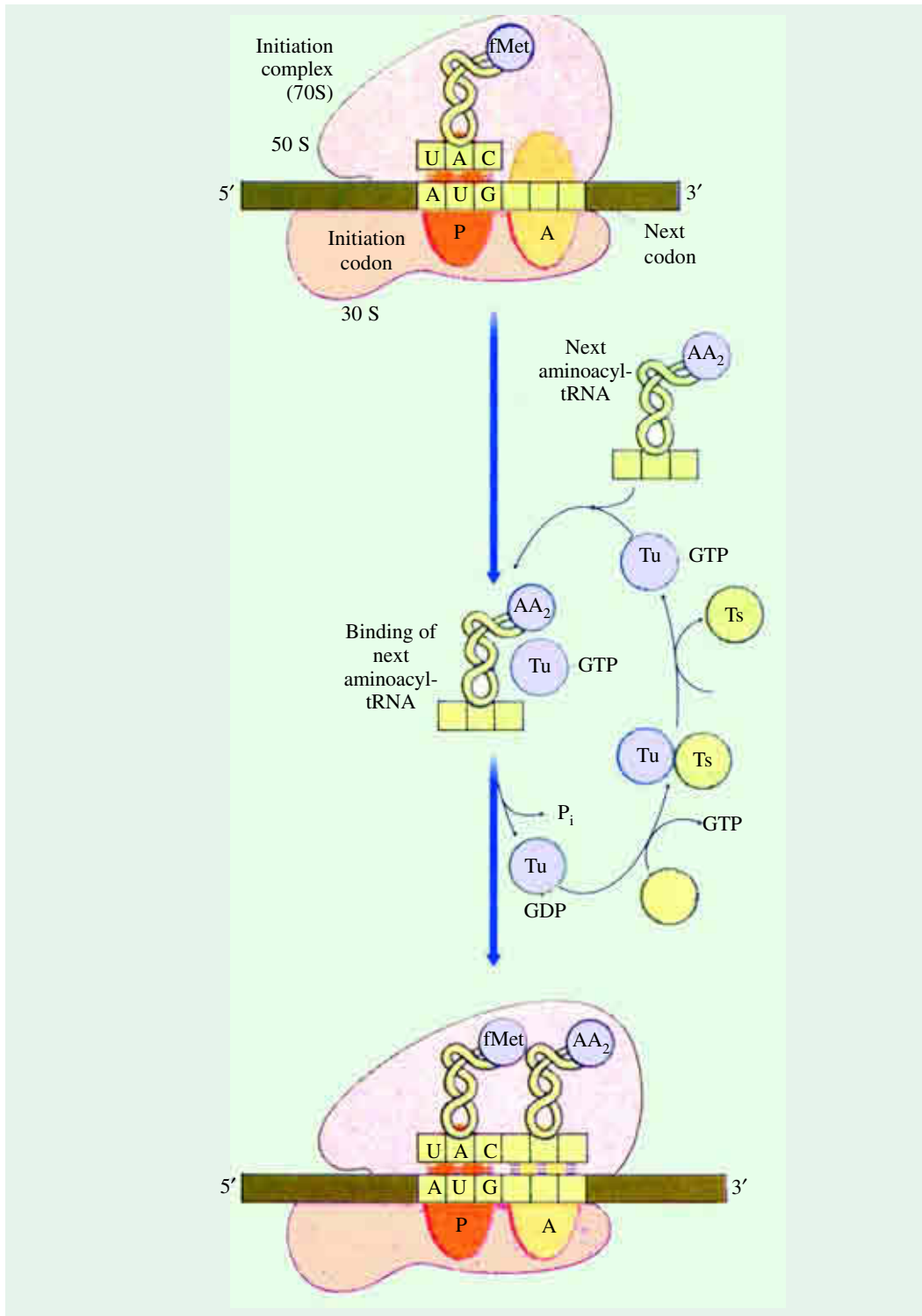
It is worth noting that *EF-Tu does not interact with fMet-tRNA<sub>f</sub>*. Hence, this initiator is not delivered to the A site. On the contrary, Met-tRNA<sub>m</sub>, like all other aminoacyl-tRNAs, does bind to EF-Tu. These findings point out that *the internal AUG codons are not read by the initiator tRNA*. Conversely, initiation factor 2 recognizes fMet-tRNA<sub>f</sub> but no other tRNA. The GTP-GDP cycle of EF-Tu is reminiscent of that of transducin in vision, the stimulatory G protein in hormone action and the *ras* protein in growth control. Indeed, the amino-terminal domain (the G domain) of EF-Tu is structurally similar to the GTP-binding subunit of these signal-transducing proteins. The other two domains of the tripartite EF-Tu are distinctive — they mediate interactions with aminoacyl-tRNA and the ribosome.

### Step 2 : Peptide bond formation.

In the second step, a new peptide bond is created between the amino acids bound by their tRNAs to the A and P sites on the ribosome (Fig. 28–24). During the process, the initiating *N*-formylmethionyl group from its tRNA is transferred to the amino group of the second amino acid now in the A site. The  $\alpha$ -amino group of the amino acid in the A site displaces the tRNA in the P site to form the peptide bond. This reaction produces a dipeptidyl-tRNA in the A site and the now ‘uncharged’ (or deacylated) tRNA<sup>fMet</sup> remains bound to the P site. The peptide bond formation was till now being referred to have been catalyzed by **peptidyl transferase** (also called **peptide synthase**). But, in 1992, Harry F. Noller *et al* discovered that this activity was catalyzed not by a protein but *by a highly conserved domain of 23S rRNA that forms the peptidyl transferase active site*. This has added another critical biological function for ribozymes.

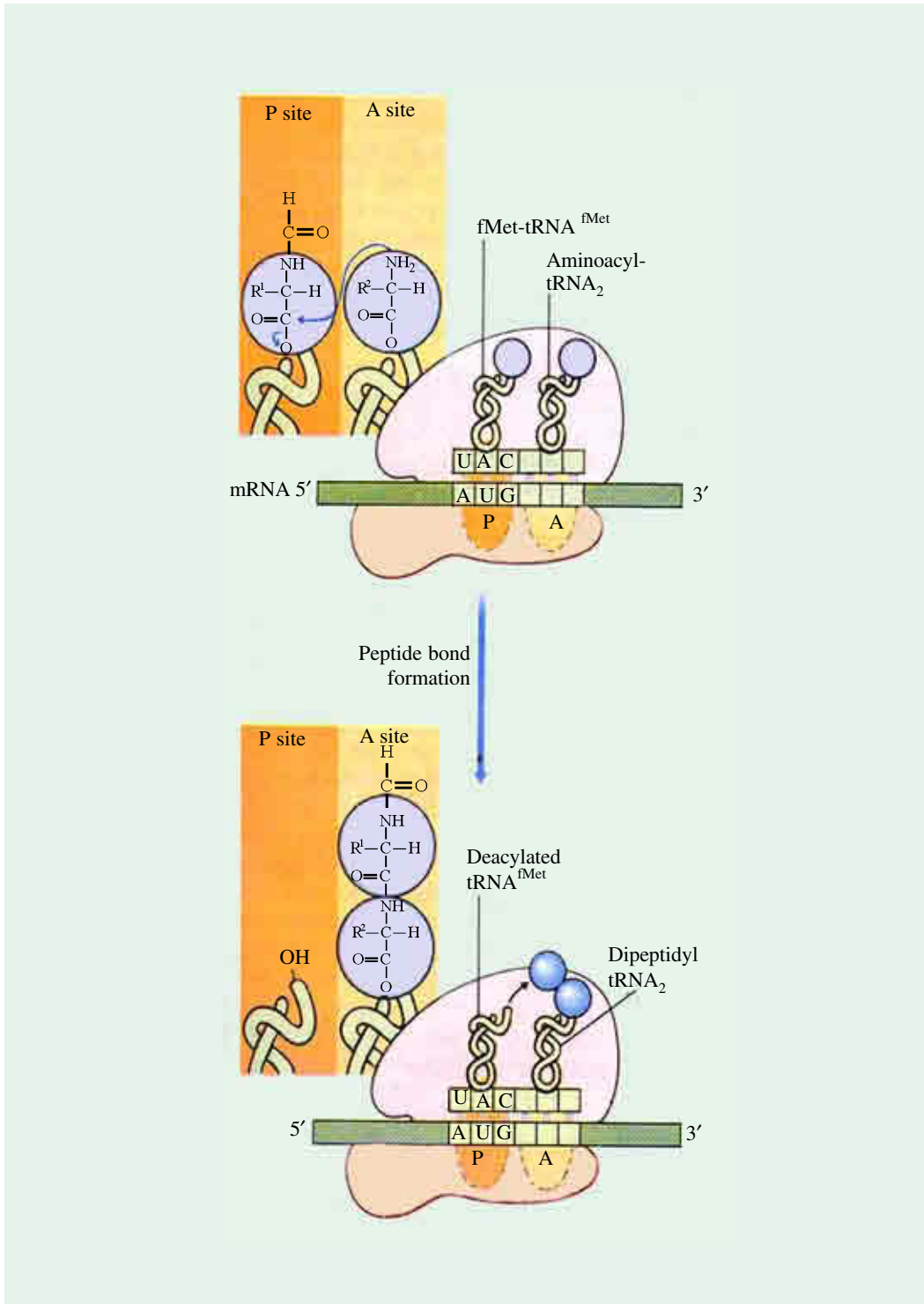
The dipeptidyl-tRNA present at A site is now translocated to P site. Two *models* are now available to explain for the translocation of dipeptidyl-tRNA from A to P site:

(a) **Two sites (A, P) model**. According to this model, deacylated tRNA is released from P site, and with the help of one GTP mole and an elongation factor, EF-G, the peptidyl-tRNA is translocated

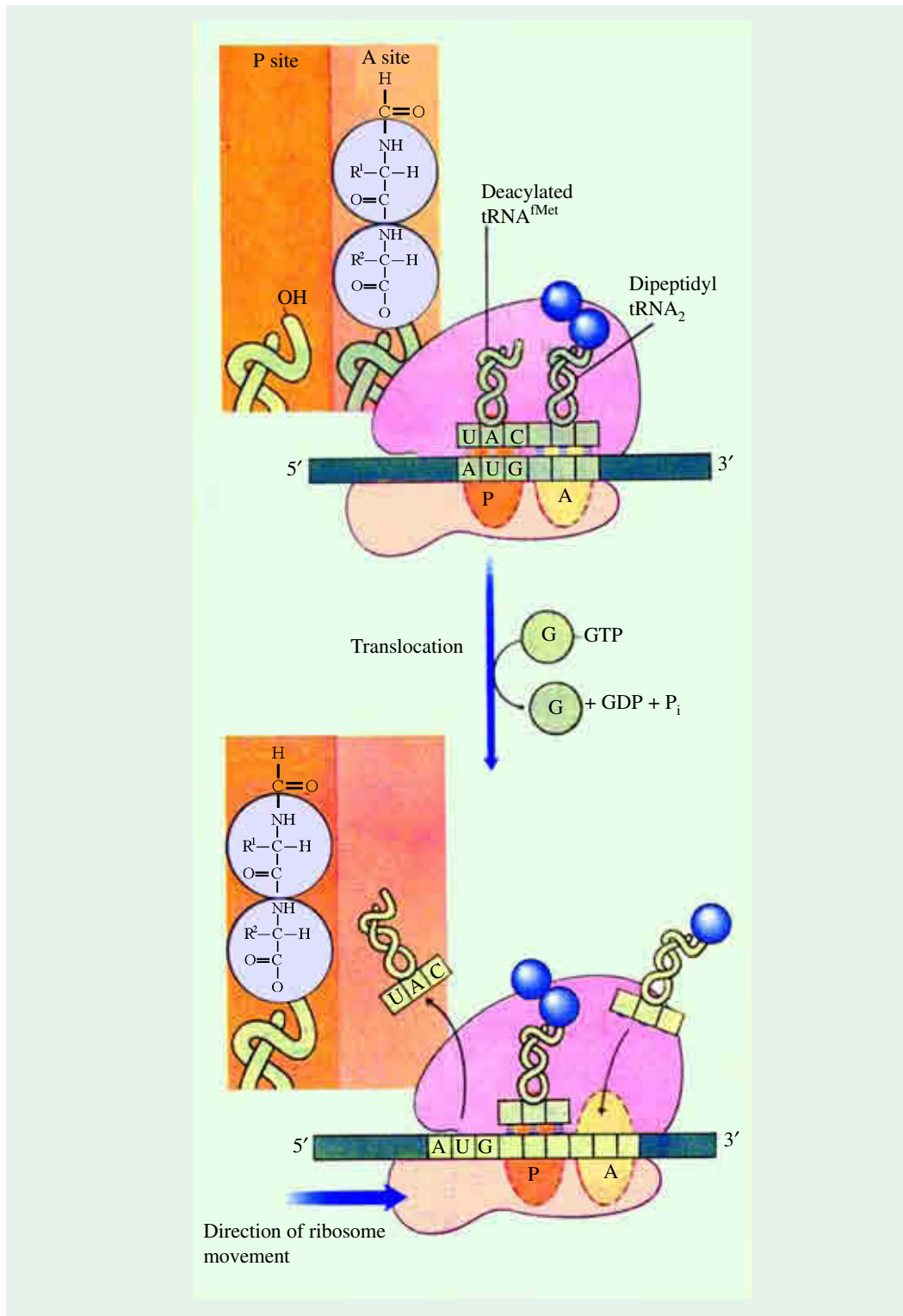


**Fig. 28–22.** First step in elongation (*i.e.*, the binding of the second aminoacyl-tRNA) in protein synthesis in a bacterium

H



**Fig. 28–23.** Second step in elongation (*i.e.*, formation of the first peptide bond) in bacterial protein synthesis



**Fig. 28–24.** Third step in elongation (*i.e.*, translocation of the peptidyl-tRNA from A to P site of the ribosome) in bacterial protein synthesis

from A to P site. Thus, according to this model, tRNA is either entirely in the A site or entirely in the P site.

(b) **Three sites (A, P, E) model.** According to this newer model, initially the aminoacyl end of the tRNA bound to A site moves to the P site on 50S subunit at the time of peptide transfer (before translocation), but only later during translocation, the anticodon end of this tRNA moves from A to P site on 30S subunit. Only this later step requires action of elongation factor, EF-G. In this model, thus, there is an intermediate state, when anticodon of tRNA is still on A site (on 30S subunit), while the aminoacyl end occupies P site (on 50S subunit). A third tRNA binding site, E (exit) was also recognized, through which tRNA leaves the ribosome. E site is found mainly on 50S subunit, and deacylated tRNA interacts with it through its CCA sequence at the end. The movement of deacylated tRNA from P to E site also occurs in 2 steps: the aminoacyl end moves to E site during peptide bond formation, and the anticodon end leaves the P site (on 30S subunit) only during translocation. This three site model suggests that *tRNAs interact mainly at their ends*. It also assumes that a tRNA remains bound to ribosome at one or the other of its two ends (aminoacyl end, anticodon end) and does not leave ribosome at both its end simultaneously (either on A or P site). The two-step transfer of tRNAs from A to P site and from P to E site could result from reciprocating motions of the 2 subunits of a ribosome. It means that *50S and 30S subunits move alternately rather than simultaneously*.

### Step 3 : Translocation of peptidyl-tRNA from A to P site.

In the third step of elongation cycle (Fig. 28–24), 3 movements take place:

- (a) the ribosome moves by the distance of one codon (*i.e.*, three nucleotides) towards the 3' end of mRNA,
- (b) because the dipeptidyl-tRNA is still attached to the second codon of the mRNA, the movement of the ribosome shifts the dipeptidyl-tRNA from the A site (on the 30S subunit) to the P site (on the 30S subunit), and
- (c) the deacylated tRNA is released from the initial P site (on the 30 S subunit) back into the cytosol.

The result is that the third codon of the mRNA is now in the A site and the second codon in the P site. This shift of the ribosome along the mRNA requires **elongation factor G, EF-G** (called so since it was isolated as GTP-requiring factor; also earlier called **translocase**) and the energy provided by hydrolysis of another molecule of GTP. EF-G, like IF-2 and EF-Tu, cycles between a GTP and a GDP form. The GTP form of EF-G is the one that drives the translocation step. After translocation, the A site is empty, ready to bind an aminoacyl-tRNA to start another round of elongation. The filling of the A site induces the release of deacylated tRNA from the E site; *the A and E sites cannot be simultaneously occupied*.

Thus, the ribosome with its attached dipeptidyl-tRNA and mRNA is now ready for another elongation cycle to attach the third amino acid residue. This occurs precisely in the same way as the addition of the second residue. Thus for each amino acid residue added to the chain, two GTPs are hydrolyzed to GDP and P<sub>i</sub>. The ribosome moves from codon to codon along the mRNA toward the end, adding one amino acid residue at a time to the growing chain.

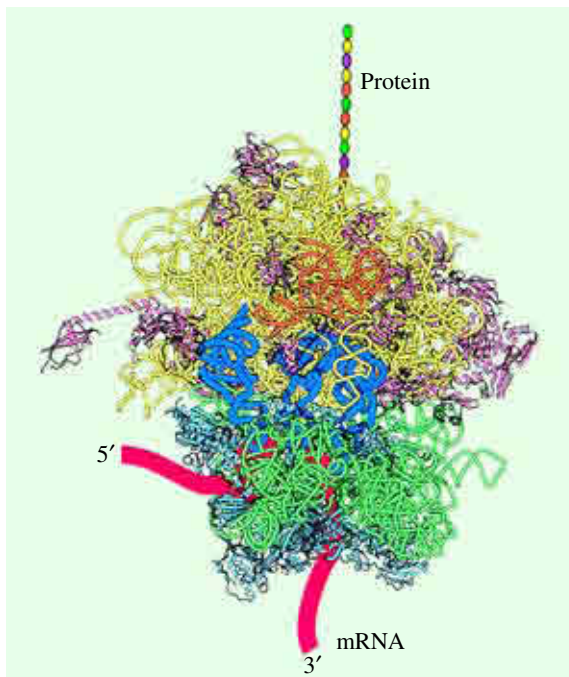
The polypeptide chain always remains attached to the tRNA of the last amino acid to have been inserted. This continued attachment to a tRNA is the chemical glue that makes the entire process work. The ester linkage between the tRNA and the carboxyl terminus of the polypeptide activates the terminal carboxyl group for nucleophilic attack by the incoming amino acid to form a new peptide bond (second step of elongation). At the same time, this tRNA represents the only link between the

## 746 FUNDAMENTALS OF BIOCHEMISTRY

growing polypeptide and the information in the mRNA. A 3-dimensional, x-ray structure of the entire gamut of ribosome, mRNA, tRNA and the nascent polypeptide chain or the '**protein assembly**', as it is called, is presented in Fig. 28–25.

### GTPase Rate of EF-Tu as the Pace Setter of Protein Synthesis

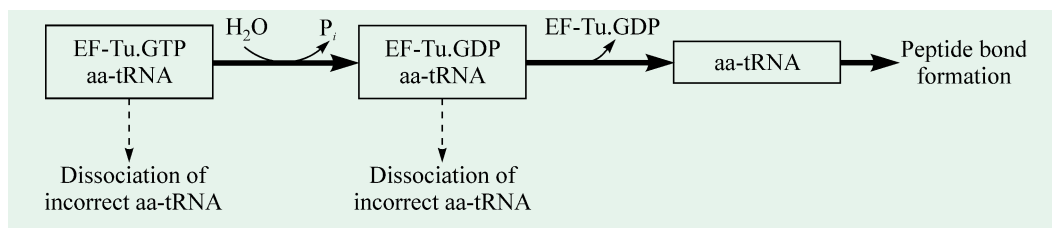
The GTPase activity of EF-Tu makes an important contribution to the rate and fidelity of the overall process of protein synthesis. The EF-GTP complex exists for a few milliseconds, and the EF-Tu-GDP complex also exists for a similar period before it dissociates (Fig. 28–26). Both these intervals provide an opportunity for the codon-anticodon interactions to be verified (*i.e.*, proofread). An incorrect aminoacyl-tRNA usually leaves the ribosome during one of these intervals, whereas the correct one stays bound. The correct aminoacyl-tRNA interacts strongly with mRNA in both states, but an incorrect one does not. In effect, the codon-anticodon interaction is scrutinized twice in different ways to achieve higher accuracy, just as proofreading of a manuscript by two readers for an hour each is much more efficient than two hours of proofreading by one. This proofreading mechanism verifies only that the proper codon-anticodon pairing has taken place. The identity of the amino acids attached to tRNAs is not checked at all on the ribosome. This was demonstrated experimentally in 1962 by two groups led by Fritz Lipmann and Semour Benzer. This finding also provided timely proof for **Crick's adaptor hypothesis** (Fig. 28-27).



**Fig. 28–25. Protein assembly**

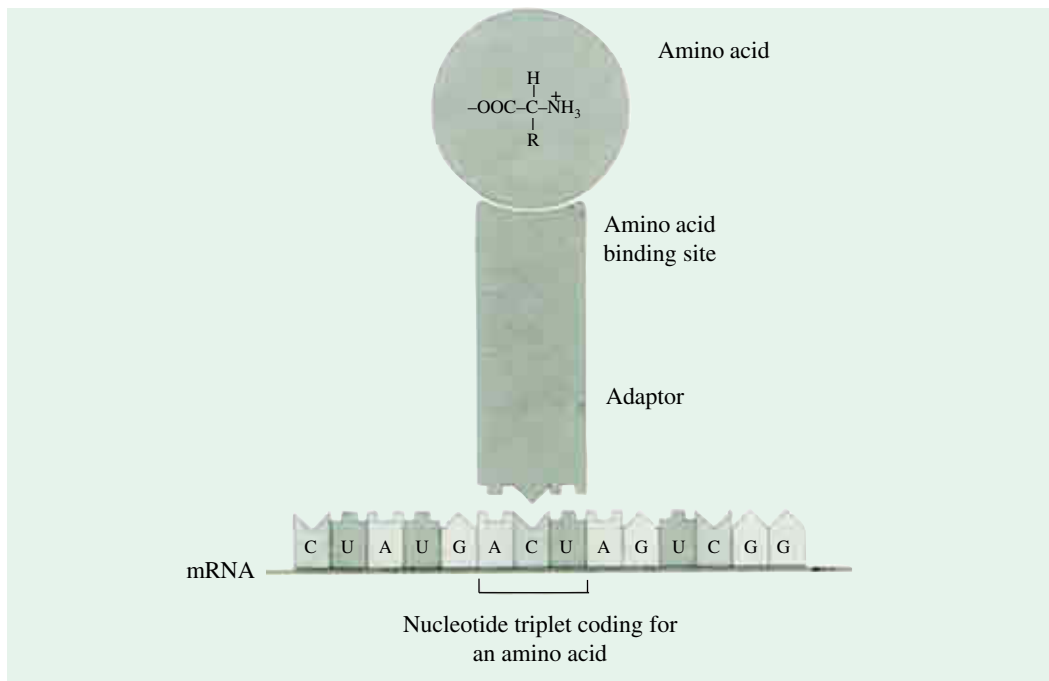
The ribosome is a factory for the manufacture of polypeptides. Amino acids are carried into the ribosome, one at a time connected to transfer RNA molecules (*blue*). Each amino acid is joined to the growing polypeptide chain, which detaches from the ribosome only once it is completed. This assembly line approach allows even very long polypeptide chains to be assembled rapidly and with impressive accuracy.

(Courtesy : Doug Martins)



**Fig. 28–26. Proofreading of the aminoacyl-tRNA occupying the A site of the ribosome**

Proofreading occurs both before and after hydrolysis of GTP bound to EF-Tu. Incorrect aminoacyl-tRNAs dissociate from the A site (steps shown by broken lines), whereas the correct aminoacyl-tRNA stays bound and a peptide bond is formed (thick arrows).



**Fig. 28–27. Crick’s adaptor hypothesis**

The hypothesis asserts that the amino acid is covalently bound at the 3' end of a tRNA molecule and that a specific nucleotide triplet elsewhere in the tRNA interacts with a particular triplet codon in mRNA through hydrogen bonding of complementary bases.

The probability  $p$  of forming a protein without any error depends on  $n$ , the number of amino acid residues, and  $\epsilon$ , the frequency of inserting a wrong amino acid:

$$p = (1 - \epsilon)^n$$

It is apparent from Table 28–4 that an error frequency of  $10^{-2}$  would be intolerable, even for small proteins. An  $\epsilon$  of  $10^{-3}$  would actually lead to the error-free synthesis of a 300-residue protein (~ 33 kd) but not of a 1000-residue protein (~ 110 kd). Thus, an error-frequency of the order of no more than  $10^{-4}$  per residue is needed to effectively produce the larger proteins. In fact, the observed values of  $\epsilon$  are close to  $10^{-4}$ . *An error frequency of about  $10^{-4}$  per residue has been selected in the course of evolution to produce the greatest number of functional proteins in the shortest time.*

**Table 28–4. Accuracy of protein synthesis**

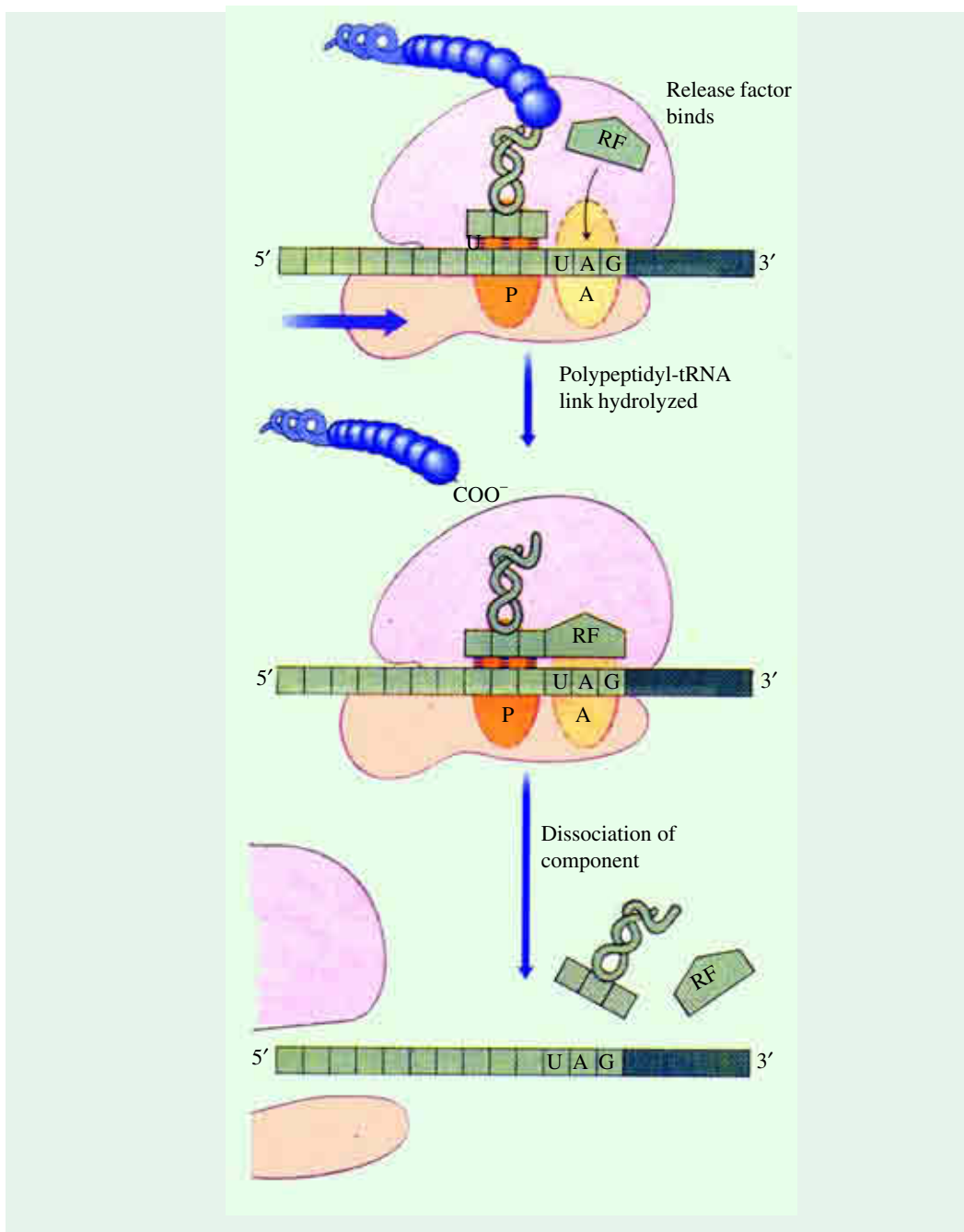
Frequency of inserting an incorrect amino acid	Probability of synthesizing an error-free protein with		
	100 AARs*	300 AARs	1000 AARs
$10^{-2}$	0.366	0.049	0.000
$10^{-3}$	0.905	0.741	0.368
$10^{-4}$	0.990	0.970	0.905
$10^{-5}$	0.999	0.997	0.990

\*AARs = amino acid residues

### TERMINATION OF POLYPEPTIDE SYNTHESIS AND RELEASE OF POLYPEPTIDE CHAIN

The process of elongation continues till the ribosome adds the last amino acid, completing the





**Fig. 28–28. The termination of protein synthesis in bacteria in response to a termination codon in the A site of mRNA**

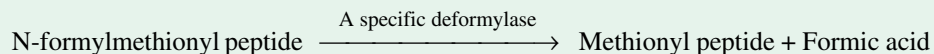
RF = Release factor RF<sub>1</sub> or RF<sub>2</sub>, depending on which termination codon is present.

polypeptide coded by the mRNA. Termination, the 4th stage of polypeptide synthesis, is signalled by one of the 3 termination codons or stop signals in the mRNA (UAA, UAG, UGA) immediately following the last amino acid codon (Fig. 28–28). Normal cells do not contain tRNAs with anticodons complementary to these stop signals. Instead, stop codons are recognized by release factors (RF<sub>1</sub>,

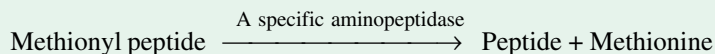
RF<sub>2</sub> and RF<sub>3</sub>), which are proteins. RF<sub>1</sub> recognizes the termination codons UAA and UAG, and RF<sub>2</sub> recognizes UGA and UAA. First, either RF<sub>1</sub> or RF<sub>2</sub> (depending on which codon is present) binds at a termination codon to the A site of the mRNA. This leads, in the second step, to the activation of peptidyl transferase so that it hydrolyzes the bond between the polypeptide and the tRNA in the P site. *The specificity of peptidyl transferase is altered by the release factor so that water rather than an amino group is the acceptor of the activated peptidyl moiety.* The completed polypeptide is, thus, detached and leaves the ribosome. In the third and final step, the mRNA, deacylated tRNA and release factor leave the ribosome and the ribosome itself dissociates into 30S and 50S components, which are now ready to start a new cycle of polypeptide synthesis. The specific function of RF<sub>3</sub> has not been decisively established. It is probably involved in the release of the completed polypeptide chain from its attachment site on the last tRNA molecule. In eukaryotes, however, a single release factor eRF recognizes all 3 termination codons.

The completed polypeptide chain has a formylmethionyl-amino terminal. Such a protein is called *nascent protein*. Before it becomes functional, the formylmethionyl-amino terminal must be removed which is accomplished by enzymes in 2 steps:

*Step I :*



*Step II :*



*Possible Role of GTP in Translation:*

Whereas ATP is selectively used for the activation of amino acids, *GTP is the only nucleoside triphosphate that is utilized for ribosomal protein synthesis.* In fact, GTP is needed for each of the 3 phases of polypeptide synthesis. It is doubtful that the hydrolysis of GTP, that repeatedly occurs during polypeptide synthesis, provides energy to drive reactions, since there is no evidence for the nucleotide's utilization to form covalent bonds. Its main role appears to be in the noncovalent binding of the various factors to a ribosome, *e.g.*, EFT and RFs. One explanation is that GTP, in each of the translation phases, allows for the specific ribosomal binding of individual elongation factors and that its hydrolysis to GDP and Pi is required for the release of the bound factors, freeing them for participation in protein synthesis.

## FOLDING AND PROCESSING OF POLYPEPTIDE CHAIN

In the fifth and final step of protein synthesis, the newly-formed peptide chain is folded and processed into its biologically-active form. At some point of time, during or after protein synthesis, the polypeptide chain spontaneously assumes its native conformation by forming sufficient number of hydrogen bonds and van der Waals, ionic, and hydrophobic interactions. In this way, the linear (or one dimensional) genetic message encoded in mRNA is converted into the 3-dimensional structure of the protein. However, there are some other nascent proteins which undergo one or more processing reactions called **posttranslational modifications**, for their conversion to the active forms. Such modifications occur in both eukaryotes and prokaryotes and include the following:

### 1. N-terminal and C-terminal Modifications

All polypeptides begin with a residue of *N*-formylmethionine (in bacteria) or methionine (in eukaryotes). However, the formyl group, the terminal methionine residue, and often additional *N*-terminal or *C*-terminal residues must be removed enzymatically before they convert into the final functional proteins. The formyl group at the *N*-terminus of bacterial proteins is hydrolyzed by a *deformylase*. One or more *N*-terminal residues may be removed by *aminopeptidases*. In about half

**750 FUNDAMENTALS OF BIOCHEMISTRY**

of the eukaryotic proteins, the amino group of the N-terminal residue is acetylated after translation. The C-terminal residues are also sometimes modified.

**2. Loss of Signal Sequences**

In certain proteins, some (15 to 30) amino acid residues at the N-terminus play a role in directing the protein to its ultimate destination in the cell. Such **signal sequences**, as they are called, are ultimately removed by specific *peptidases*.

**3. Modification of Individual Amino Acids**

Certain amino acid side chains may be specifically modified. For instance, the hydroxyl groups of certain serine, threonine, tyrosine residues of some proteins undergo enzymatic phosphorylation by ATP [Fig. 29(a)]; the phosphate groups add negative charge to these polypeptides. The functional significance of this modification varies from one protein to the other. For example, the milk protein *casein* has many phosphoserine groups, which function to bind  $\text{Ca}^{2+}$ . Given that  $\text{Ca}^{2+}$  and phosphate, as well as amino acids, are required by suckling young, casein provides three essential nutrients. The phosphorylation and dephosphorylation of the OH group of certain serine residues regulate the activity of some enzymes, such as *glycogen phosphorylase*.

Sometimes, additional carboxyl groups are added to Asp and Glu residues of some proteins. For instance, the blood clotting protein prothrombin contains many  $\gamma$ -carboxyglutamate residues [Fig. 28–29 (b)] in its N-terminal region. These groups bind  $\text{Ca}^{2+}$  which is required to initiate the clotting mechanism.

In some proteins, certain lysine residues are methylated enzymatically [Fig. 28–29(c)]. Monomethyl- and dimethyllysine residues are present in some muscle proteins and in *cytochrome c*. *Calmodulin* of most organisms contains one trimethyllysine residue at a specific position. In other proteins, the carboxyl groups of some Glu residues undergo methylation, which removes their negative charge.

Some proline and lysine residues in *collagen* are hydroxylated.

**4. Formation of Disulfide Cross-links**

Some proteins after acquiring native conformations are often covalently cross-linked by the formation of disulfide bridges between cysteine residues (refer Fig. 9– ). These cross-links help to protect the native conformation of the protein molecule from denaturation in an extracellular environment that is quite different from that inside the cell.

**5. Attachment of Carbohydrate Side Chains**

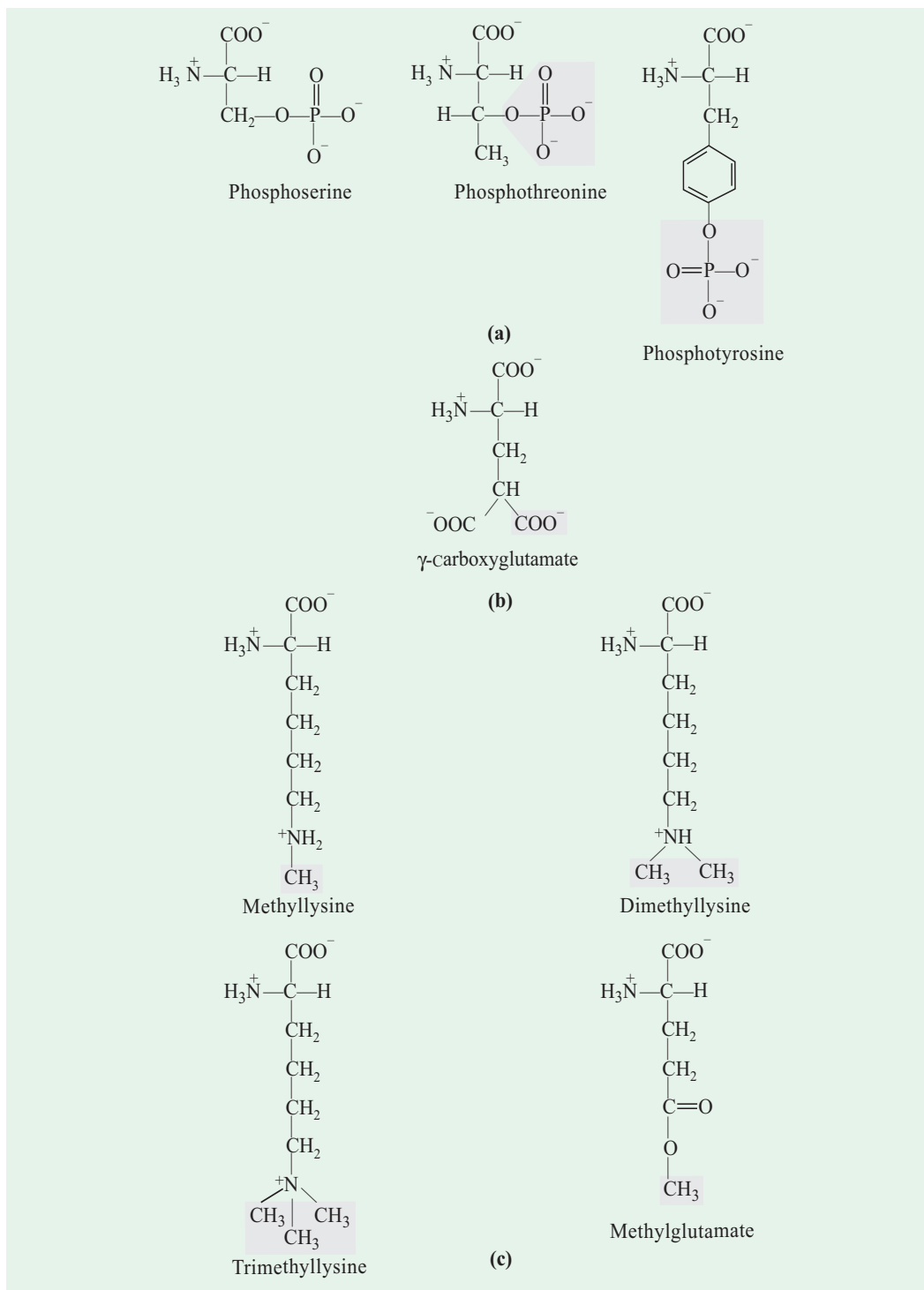
In glycoproteins, the carbohydrate side chains are attached covalently during or after the synthesis of polypeptide chain. In some glycoproteins, the carbohydrate side chain is attached enzymatically to Asn residues (*N*-linked oligosaccharides), in others to Ser or Thr residues (*O*-linked oligosaccharides). Many proteins that function extracellularly contain oligosaccharide side chains.

**6. Addition of Prosthetic Groups**

Many prokaryotic and eukaryotic proteins require for their activity covalently-bound prosthetic groups. These groups become attached to the polypeptide chain after it leaves the ribosome. The two significant examples are the covalently-bound biotin molecule in *acetyl-CoA carboxylase* and the heme group of *cytochrome c*.

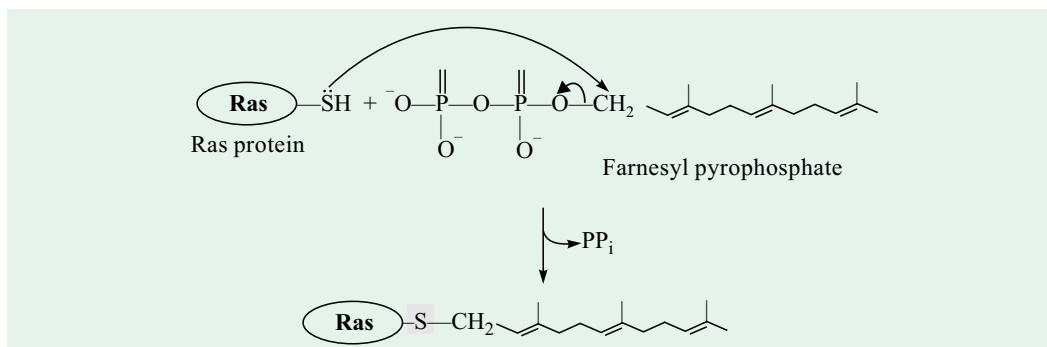
**7. Addition of Isoprenyl Groups**

Many eukaryotic proteins are isoprenylated; a thioester bond is formed between the isoprenyl group and a cysteine residue of the protein. The isoprenyl groups are derived from pyrophosphate intermediates of the cholesterol biosynthetic pathway, such as farnesyl pyrophosphate (Fig. 28–30). Proteins so modified include the products of the *ras* oncogenes and proto-oncogenes, *G proteins*, and proteins called *lamins*, found in the nuclear matrix.



**Fig. 28-29. Some modified amino acid residues**

(a) Three phosphorylated amino acids (b) A carboxylated amino acid (c) Four methylated amino acids



**Fig. 28–30. Farnesylation of cysteine residue on a protein**

The *ras* protein is the product of the *ras* oncogene. The thioether linkage is shown in a shaded rectangle.

### 8. Proteolytic Trimming

Many proteins (*insulin, collagen*) and proteases (*trypsin, chymotrypsin*) are initially synthesized as larger, inactive precursor proteins. These precursors are proteolytically trimmed to produce their final, active forms. Some animal viruses, notably poliovirus, synthesize long polycistronic proteins from one long mRNA molecule. These protein molecules are subsequently cleaved at specific sites to provide the several specific proteins required for viral function.

### ENERGY REQUIREMENTS FOR PEPTIDE BOND FORMATION

The formation of a peptide bond in a small peptide requires the utilization of ATP. In the case of ribosome-dependent peptide bond synthesis, the equivalent of 4 ATPs are required per peptide bond formed. The break-up of these 4 ATP equivalents is as follows:

- Two moles of ATP are required for the activation of each amino acid to form an aminoacyl-tRNA.
- One mole of GTP is cleaved to GDP and P<sub>i</sub> during the first step of elongation, *i.e.*, binding of an aminoacyl-tRNA to the A site.
- Another mole of GTP is hydrolyzed in the translocation of the peptidyl-tRNA from the P to A site.

Additional ATPs are used each time incorrectly-activated amino acids are hydrolyzed by the deacylation activity of some *aminoacyl-tRNA synthetases*. Henceforth, a total of at least 4 high-energy bonds is ultimately required for the formation of each peptide bond of the completed polypeptide chain.

This represents an exceedingly large thermodynamic ‘push’ in the direction of synthesis: at least  $4 \times 30.5 = 122$  kJ/mol of phosphodiester bond energy is required to produce a peptide bond having a standard free energy of hydrolysis of only about  $-21$  kJ/mol. The net free energy change in peptide bond synthesis is thus  $-101$  kJ/mol. Thus, *the fidelity in protein synthesis is energetically expensive*. Although, this large energy expenditure may appear wasteful, it is important to remember that proteins are information-containing polymers. The biochemical problem is not simply the formation of a peptide bond, but the formation of a peptide bond between *specific* amino acids. This greater energy requirement for the synthesis of peptide bonds of large proteins probably reflects the need for greater *sequence specificity* in these macromolecules. This energy makes possible the nearly perfect fidelity in biological translation of the genetic message of mRNA into the amino acid sequence of proteins.

### INHIBITORS OF PROTEIN SYNTHESIS

As in the investigation of other complex metabolic processes, specific inhibitors have played a major role in separating the steps in the biosynthesis of nucleic acids and proteins. Some of these

inhibitors are synthetic compounds; others were first isolated from the culture filtrates of various organisms as antibiotics in efforts to control infectious diseases or to inhibit the growth of malignant tissue.

Protein synthesis is a pivotal function in cellular physiology and as such it is a primary target of a number of naturally-occurring antibiotics and toxins. Most of these antibiotics inhibit protein synthesis in bacteria (*i.e.*, prokaryotes) but are relatively harmless to eukaryotic cells. The inhibitory effects of antibiotics on bacterial growth led to their application in studies designed to determine the intermediary steps in protein synthesis. *Nearly every step in protein synthesis can be specifically inhibited by one antibiotic or the other* (Table 28–5).

**Antibiotics** are substances produced by bacteria or fungi that inhibit the growth of other organisms. They are known to inhibit a variety of essential biological processes, such as DNA replication (e.g., novobiocin), transcription (e.g., rifamycin) and bacterial wall synthesis (e.g., penicillin). However, the majority of known antibiotics, including a great variety of medically-useful substances block translation. This situation is presumably a consequence of the translation machinery's enormous complexity, which makes it vulnerable to disruption in many ways.

**Table 28–5. Inhibitors of protein or RNA synthesis**

Antibiotics	Organisms	Site of action	Specific action*
Streptomycin	Prokaryotes	30S	Inhibits movement of initiation complex and causes misreading of mRNA
Tetracycline	Prokaryotes	30S	Blocks recognition site and inhibits binding of aminoacyl-tRNAs at A site
Chloramphenicol	Prokaryotes	50S	Inhibits the peptidyl transferase activity on 70S ribosomes
Cycloheximide	Eukaryotes	60S	Inhibits the peptidyl transferase activity on 80S ribosomes
Erythromycin	Prokaryotes	50S	Binds to the 50S subunit and inhibits translocation reaction on ribosomes
Puromycin	Prokaryotes, Eukaryotes	50S, 60S	Causes premature chain termination by acting as analogue of aminoacyl-tRNA
Fusidic acid	Prokaryotes, Eukaryotes	50S, 60S	Blocks the release of EF-G and GDP during elongation cycle
Rifamycin	Prokaryotes	Transcription	Blocks initiation of RNA chains by binding to RNA polymerase
Streptoligidin	Prokaryotes	Transcription	Inhibits elongation of the peptide chain
Actinomycin D	Prokaryotes, Eukaryotes	—	Binds to DNA and blocks the movement of RNA polymerase
Anisomycin	Eukaryotes	—	Blocks the peptidyl transferase reaction on ribosomes
$\alpha$ -amanitin	Eukaryotes	—	Blocks mRNA synthesis

\* The ribosomes of eukaryotic mitochondria (and chloroplasts) often resemble those of prokaryotes in their sensitivity to inhibitors. Therefore, some of these antibiotics can have a deleterious effect on human mitochondria.

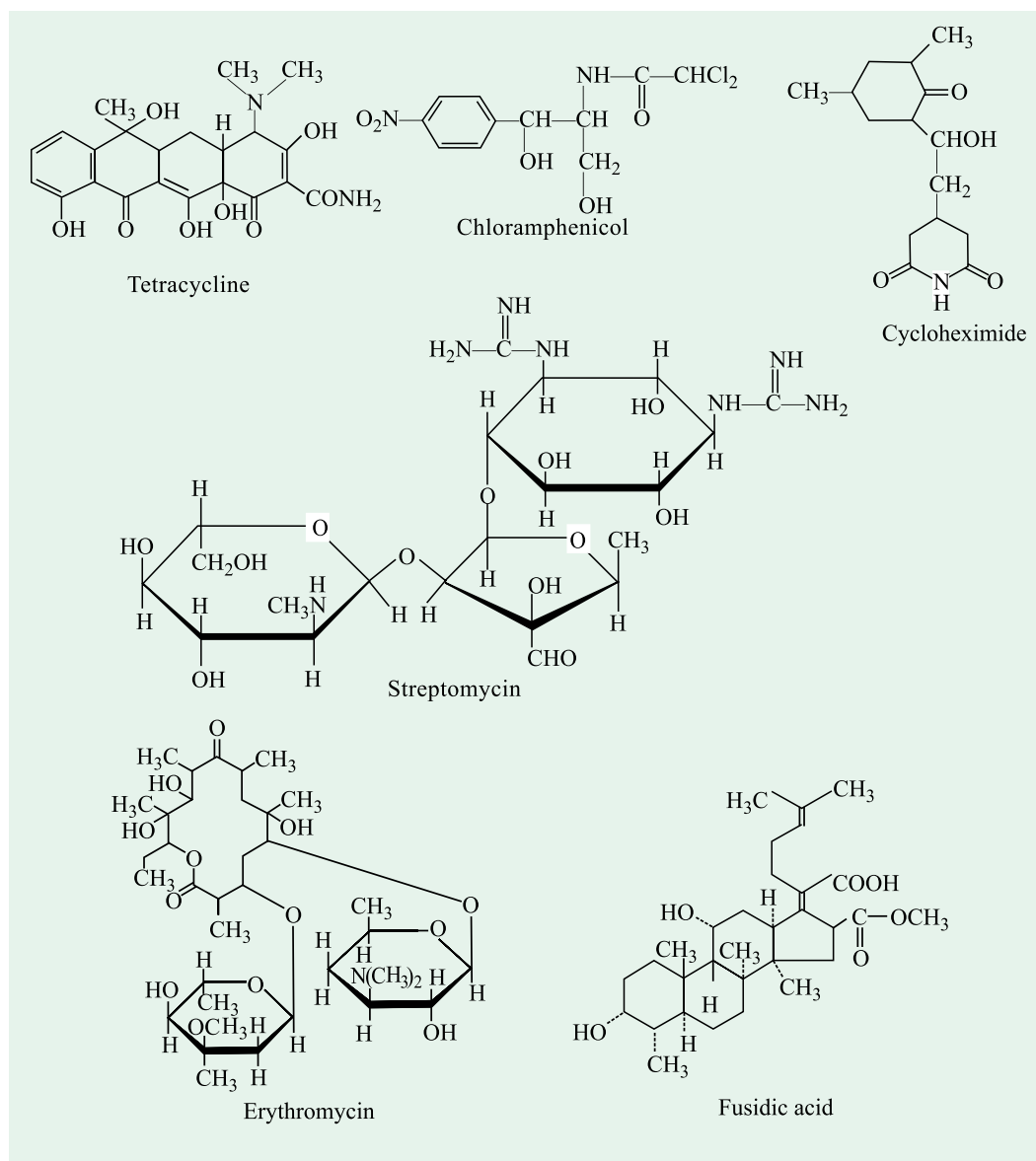
Some known inhibitors of protein synthesis are described below (Fig. 28–25):

**1. Streptomycin.** Streptomycin, which was discovered by Selman Wakesman in 1944, is a medically important member of a family of antibiotics known as *aminoglycosides* that inhibit prokaryotic ribosomes in a variety of ways. It is a highly basic trisaccharide and, at higher

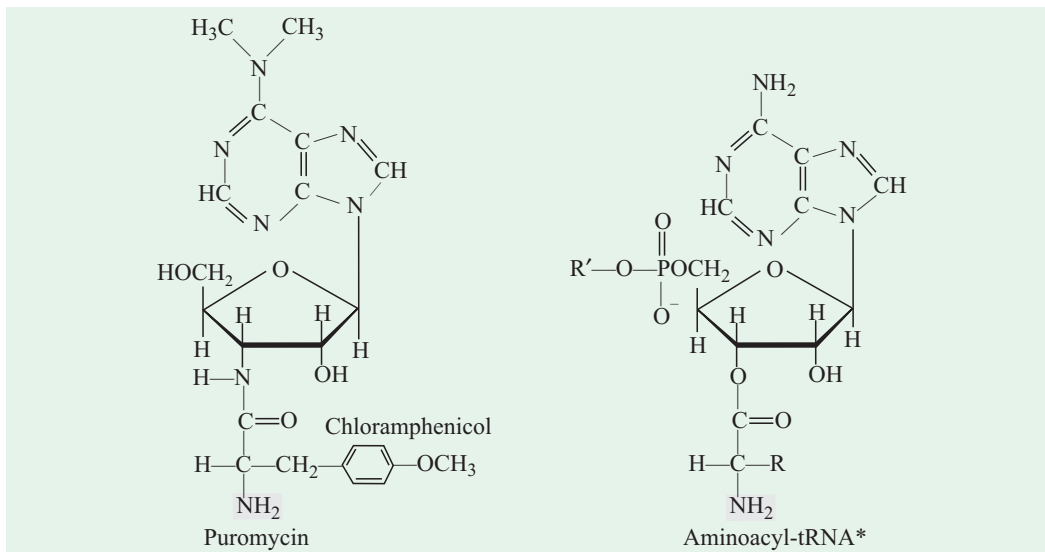
## 754 FUNDAMENTALS OF BIOCHEMISTRY

concentrations, interferes with the binding of fMet-tRNA to ribosomes and thereby prevents the correct initiation of protein synthesis. And at relatively low concentrations, streptomycin also leads to a misreading of the genetic code on the mRNA and inhibit initiation of the polypeptide chain. If poly U is the template, Ile (AUU) is incorporated in addition to Phe (UUU). An extensive series of experiments revealed that a single protein in the 30S subunit, namely protein S12, is the determinant of streptomycin sensitivity.

**2. Tetracyclines.** Tetracycline and its derivatives are broad-spectrum antibiotics that inhibit protein synthesis by blocking the A site on the ribosome so that the binding of aminoacyl-tRNAs is inhibited; the nascent polypeptide chain remains in the P site and can react normally with pyromycin, another antibiotic inhibitor.







**Fig. 28–31. Some known inhibitors of bacterial and eukaryotic protein syntheses**

The structure of aminoacyl-tRNA is shown here because its terminal aminoacyl-adenosine portion is a structural analogue of puromycin.

**3. Chloramphenicol, CAP (= Chloromycetin).** Chloramphenicol, the first of the “broad-spectrum” antibiotics, inhibits peptidyl transferase activity on the large subunit of prokaryotic ribosomes. However, its clinical uses are limited to only severe infections because of its toxic side effects, which are caused, at least in part, by the chloramphenicol sensitivity of mitochondrial ribosomes. It is a classic inhibitor of protein synthesis in bacteria and acts, at relatively low concentrations on bacterial (also mitochondrial and chloroplast) ribosomes by blocking peptidyl transfer by interfering with the interactions of ribosomes with A site-bound aminoacyl-tRNAs, but does not affect cytosolic protein synthesis in eukaryotes. Of the various possible optical isomers, only the D (–) *threo* form shows significant inhibitory activity.

**4. Cycloheximide (= Actidione).** It is a potent fungicide antibiotic and blocks the peptidyl transferase of 80S eukaryotic ribosomes but not that of 70S bacterial (also mitochondrial and chloroplast) ribosomes. Contrary to chloramphenicol, cycloheximide affects only ribosomes in the cytosol. The difference in the sensitivity of protein synthesis to these two drugs provides a powerful way to determine in which cell compartment a particular protein is translated.

**Cycloheximide** is also spelt as cyclohexamide (G. Karp, 1984)

**5. Erythromycin.** It binds to the bacterial 50S ribosomal subunit and blocks the translocation step, thereby “freezing” the peptidyl-tRNA in the A site.

**6. Fusidic acid.** It is a steroid and affects the translocation step in eukaryotic ribosomes after formation of the peptide bond, possibly by preventing cleavage of GTP in the eEF2-mediated cleavage-translocation reaction.

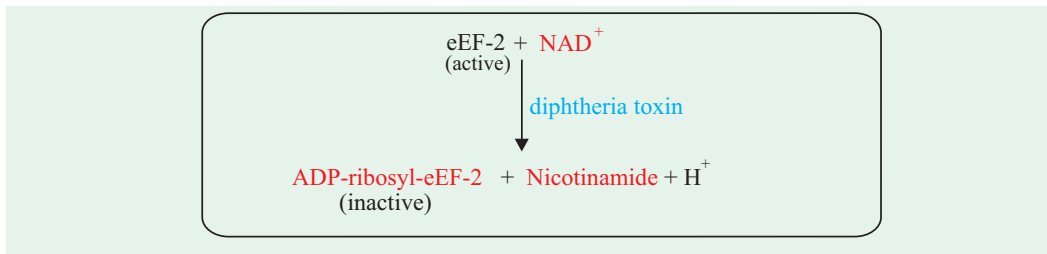
**7. Diphtheria toxin :** Diphtheria is a disease resulting from bacterial infection by *Corynebacterium diphtheriae* that is infected with a specific lysogenic phage called *corynephage*  $\beta$ . (Diphtheria was a leading cause of childhood death until the late 1920s when immunization became prevalent). Although the bacterial infection is usually confined to the upper respiratory tract, the bacteria secrete a phage-encoded protein, called *diphtheria toxin*, that is responsible for the disease’s lethal effects. *Diphtheria toxin specifically inactivates the eukaryotic elongation factor eEF-2, thereby inhibiting eukaryotic protein synthesis.*

Diphtheria toxin acts in a particularly interesting way. It is a monomeric 535-residue protein

## 756 FUNDAMENTALS OF BIOCHEMISTRY

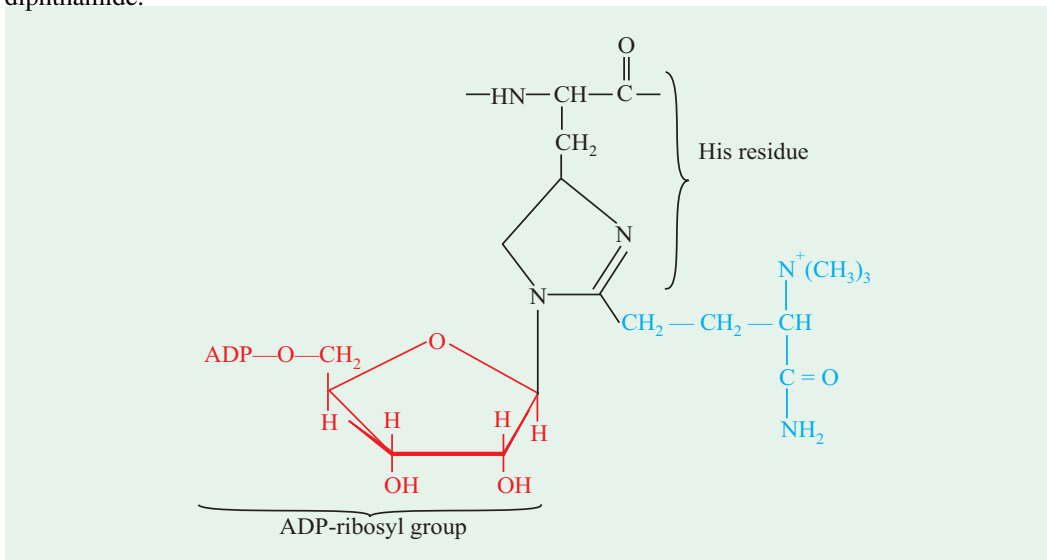
that is cleaved past Arg residues 190, 192 and 193 by trypsin and trypsin-like enzymes. This hydrolysis occurs around the time diphtheria toxin encounters its target cell, yielding 2 fragments A and B, which nevertheless remain linked by a disulfide bond. The fragment B binds to a specific receptor on the plasma membrane of susceptible cells, whereupon it facilitates A fragment's cytosolic uptake *via*, receptor-mediated endocytosis (free fragment A is devoid of toxic activity). The intracellular reducing environment then cleaves the disulfide bond linking the A and B fragments.

Within the cytosol the fragment A catalyzes the ADP-ribosylation of eEF-2 by  $\text{NAD}^+$ , thereby inactivating this elongation factor. Since the fragment A acts catalytically, *one molecule is sufficient*



to ADP-ribosylate all of a cell's eEF-2s, which halts protein synthesis and kills the cells. Only a few micrograms of diphtheria toxin are, hence, sufficient to kill an unimmunized individual.

Diphtheria toxin specifically ADP-ribosylates a modified. His residue on eEF-2 known as diphthamide.



Diphthamide occurs only in eEF-2 (not even in its bacterial counterpart, EF-G), which accounts for the specificity of diphtheria toxin in exclusively modifying eEF-2. Since diphthamide occurs in all eukaryotic eEF-2 is, it probably is essential to eEF-3 activity.

**8. Ricin or Abrin.** It is an extremely toxic protein and inactivates the 60S subunit of eukaryotic ribosomes.

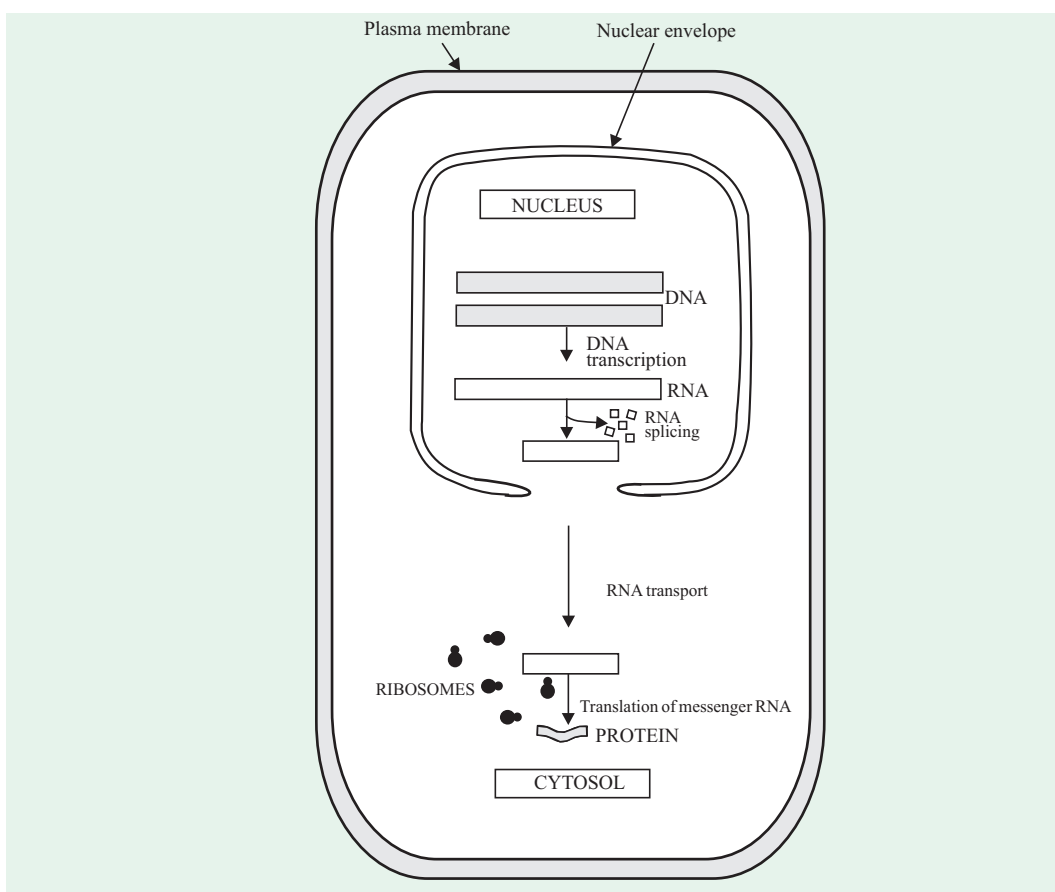
**9. Puromycin.** It is one of the best-understood inhibitory antibiotics and is obtained from a mould *Streptomyces alboniger*. Puromycin is a structural analogue of the terminal aminoacyl-adenosine portion of aminoacyl-tRNA. It binds to the A site on the ribosome and inhibits the entry of aminoacyl-tRNA. Furthermore, puromycin contains an  $\alpha$ -amino group which, like the one on aminoacyl-tRNA, forms a peptide bond with the carboxyl group of the growing peptide chain in a reaction catalyzed by peptidyl transferase. The product is a peptide having a covalently-attached puromycin residue at its

carboxyl end. Peptidyl puromycin then dissociates from the ribosome. Thus, normal chain growth is aborted, and incomplete peptide chains bearing carboxyl-terminal puromycin are released from the ribosomes. Puromycin has been used to ascertain the functional state of ribosomes. In fact, the concept of A and P sites resulted from the use of puromycin to ascertain the location of peptidyl-tRNA. When peptidyl-tRNA is in the A site (before translocation), it cannot react with puromycin.

### EUKARYOTIC PROTEIN SYNTHESIS

The eukaryotic protein synthesis, which has been in essence diagrammatically represented in Fig. 28–26, follows essentially the same pattern as that described for *Escherichia coli* (the bacterium most ideal for *in vitro* experimentation) with certain distinct differences, already mentioned at places of their reference. These may, however, be summed up as follows:

**1. Chromatin Material.** In prokaryotes, there is no distinct nucleus bound by a nuclear membrane. As such, the two processes of transcription and translation take place *simultaneously* in the cell; whereas in eukaryotes, transcription of mRNA takes place inside the nucleus and followed



**Fig. 28–32. Protein synthesis (DNA → RNA → Protein) in eukaryotes**

Eukaryotic cells have evolved many membrane-bound compartments that separate their various chemical reactions so as to make them more efficient, and the nucleus is one such compartment. The nuclear envelope keeps functional ribosomes out of the nucleus and thus prevents RNA transcripts from being translated into protein until they have been processed (spliced) and transported and transport steps are interposed between DNA transcription and RNA translation.

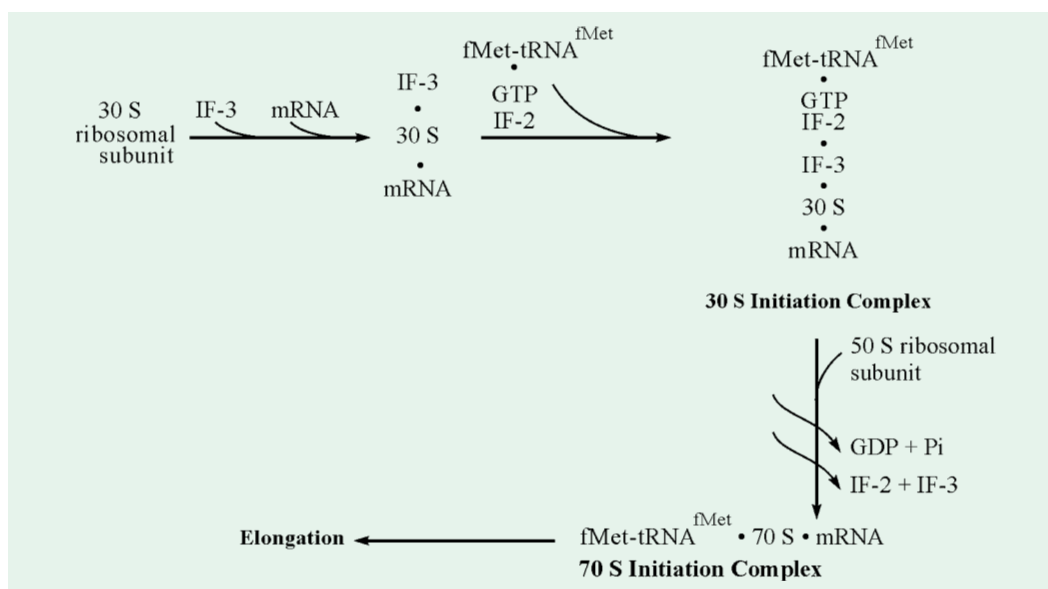
## 758 FUNDAMENTALS OF BIOCHEMISTRY

by its translocation in the cytoplasm, *i.e.*, the two processes take place *successively*. In other words, in eukaryotic cells, transcription and translation are separated in time and space; whereas in prokaryotes, they are not. However, in both transcription occurs at the site of DNA representing a gene. Prokaryotes lack a nuclear envelope. Since mRNAs are synthesized (or transcribed) in a 5' → 3' direction, and translation also proceeds in the same direction, there can be, in prokaryotes, a coupling between transcription and translation, *i.e.*, translation of a mRNA can begin even before its synthesis is completed. Contrarily, in eukaryotes, such a coupling is not possible because the mRNAs are synthesized in the nucleus (where the DNA is localized) and must then cross the nuclear membrane and reach the cytoplasm before translation begins.

**2. Ribosomes.** The eukaryotic ribosomes are larger than those of the prokaryotes. They consist of a 60S larger subunit and a 40S smaller subunit, come together to form an 80S particle having a mass of 4,200 kd, in comparison to 2,700 kd for the prokaryotic 70S ribosome. The 40S subunit contains an 18S RNA that is homologous to the prokaryotic 16S RNA. The 60S subunit contains 3 RNAs: the 5S and 23S are homologous to the prokaryotic 5S and 23S; its 5.8S is unique to eukaryotes and has no counterpart in prokaryotes.

**3. Initiator tRNA.** Methionine initiates eukaryotic protein synthesis; whereas the starting amino acid residue at the amino-terminus in prokaryotes is N-formylmethionine. However, as in prokaryotes, a special tRNA participates in initiation. This aminoacyl-tRNA is called Met-tRNA<sub>f</sub> or Met-tRNA<sub>i</sub> (the subscript f indicates that it can be formylated *in vitro*, and i stands for initiation).

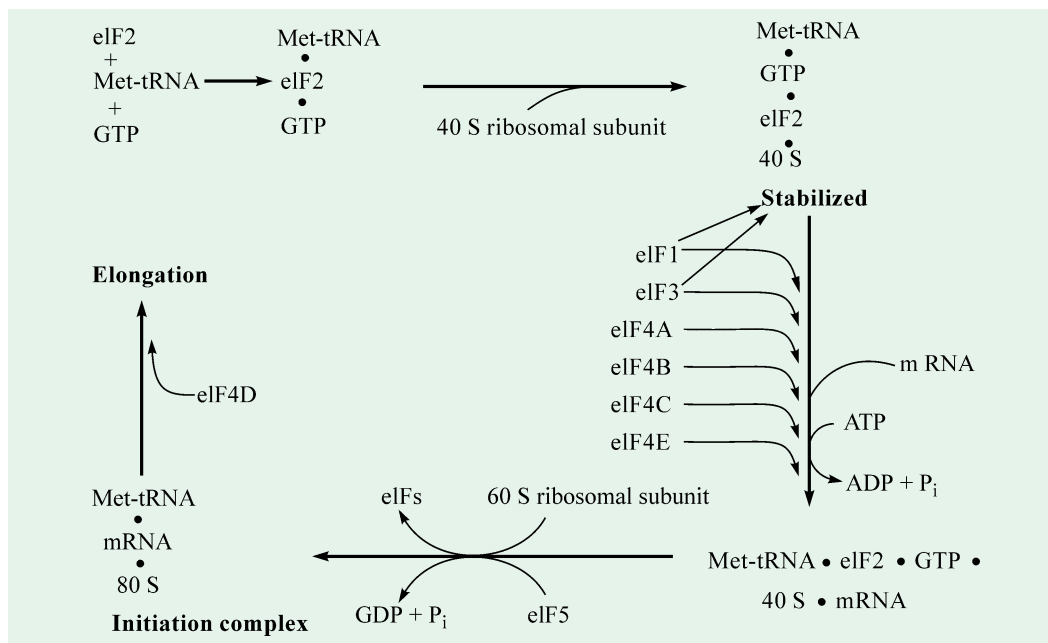
**4. Initiation Complexes.** The most enigmatic difference to date, however, is with regard to the number of initiation factors needed for the formation of the eukaryotic initiation complex. Whereas the *E. coli* system requires 3 factors (IF1, IF2, IF3), the eukaryotic initiation process requires at least 9 (eIF1, eIF2, eIF3, eIF4A, eIF4B, eIF4C, eIF4D, eIF4E, eIF5). This list does not include factors that enhance the activity of eIF2 (Co-eIF2 factors) or those that enhance the dissociation of 80S ribosomes into their subunits (ribosome-dissociation factors). Also, in prokaryotes, the smaller ribosomal subunit (30S) first combines with mRNA and the 30S-mRNA complex then unites with fMet-tRNA<sup>fMet</sup> (Fig. 28–33). But in eukaryotic protein synthesis, the smaller ribosomal subunit (40S) associates with Met-tRNA<sup>fMet</sup> without the help of mRNA. The scheme for the assemblage of



**Fig. 28–33. Proposed assembly of the initiation complex for prokaryotic protein synthesis**

The initiation factor 1 (IF-1) plays a role in stimulating activities of IF-2 and IF-3.

the eukaryotic initiation complex is presented in Fig. 28–34. The step-by-step process is similar to that described for the formation of the *E. coli* complex, *i.e.*, complexing of an initiating tRNA, the smaller ribosomal subunit, an mRNA, and, finally, the larger ribosomal subunit. With its requirement for 6 of the 9 initiation factors, the addition of mRNA to the complex is the most complex molecular step in the process. ATP hydrolysis is also required for mRNA addition. Thus, *for the eukaryotic system, both ATP and GTP are necessary, whereas the comparable E. coli process requires only GTP.*



**Fig. 28–34. Proposed assembly of the initiation complex for eukaryotic protein synthesis**

**5. Start Signals.** The initiating codon in eukaryotic protein synthesis is always AUG. Eukaryotes, as opposed to prokaryotes, do not use a purine-rich sequence on the 5' side to distinguish initiator AUGs from internal AUGs. Instead, the AUG nearest the 5' end of mRNA is normally selected as the initiation site. A eukaryotic mRNA has a *single initiation site* and hence is the template for a single protein. On the contrary, a prokaryotic mRNA can have *multiple initiation sites* and it can serve as the template for the synthesis of many proteins.

**6. Elongation and Termination Factors.** For eukaryotic chain termination, only one release factor (eRF, e for eukaryotes), which is a GTP-driven protein and recognizes all 3 termination codons (UAA, UAG, UGA), is required; whereas *two release factors* ( $RF_1$ ,  $RF_2$ ) are needed for termination of polypeptide chain in prokaryotic systems such as *E. coli*. Finally, eIF3, like its prokaryotic counterpart IF3, prevents the reassociation of ribosomal subunits in the absence of an initiation complex.

## PROTEIN SYNTHESIS IN MITOCHONDRIA AND CHLOROPLASTS

The translation apparatuses in both mitochondria and chloroplasts differ with that in cytoplasm in the following respects:

1. Ribosomes in these organelles are smaller in size (70S) than those in the cytoplasm (80S).
2. The processes of translation starts with formylmethionyl-tRNA in both mitochondria and chloroplasts, whereas formylation does not occur in the cytoplasm.
3. Translation in both mitochondria and chloroplasts can be inhibited by chloramphenicol, as in bacteria, since the 70S ribosomes are sensitive to chloramphenicol; on the contrary, the

## 760 FUNDAMENTALS OF BIOCHEMISTRY

translation in cytoplasm is inhibited by cycloheximide since 80S ribosomes are sensitive to cycloheximide.

The mitochondria contain circular DNA molecules which differ from nuclear DNA in size, base composition, synthesis and turnover rate. The DNA is located in many discrete regions in the mitochondrial matrix, which codes for fewer proteins. The mitochondria contain ribosomes of 70S type and the ribosomal species resemble those of prokaryotes, rather than their counterparts in the cytoplasm which are 80S type. Mitochondria also possess many organelle-specific tRNAs and aminoacyl-tRNA synthetases, which catalyze the attachment of amino acids to tRNA molecules. The mitochondrial tRNA hybridizes with mitochondrial DNA and hence may be made within the organelle on mitochondrial DNA template. The enzymes DNA and RNA polymerases, capable of acting on mitochondrial DNA, are also found in mitochondria.

The **mitochondrial** genome of various eukaryotic cells (esp. yeast cells and human cells) has been extensively studied in the recent past. The mitochondrial genome of humans, which is a circular double-stranded DNA, comprises 16,569 base pairs (as opposite to 78,000 base pairs in the case of mitochondrial genome of the yeast *Saccharomyces cerevisiae*) and codes for the following:

1. two rRNAs with 1,559 and 954 nucleotides respectively (3,200 and 1,660 nucleotides in yeast)
2. twenty two tRNAs (24 in yeast). This number of tRNAs is definitely smaller than the minimum number of tRNAs required to 'read' the 61 'sense' codons, taking into consideration the wobble hypothesis, *i.e.*, 31 tRNAs (plus the initiation tRNA).
3. about thirteen polypeptides which are part of oligomeric membrane proteins: cytochrome oxidase, apocytochrome b, ATPase, NADH dehydrogenase. The cytochrome oxidase contains seven different types of polypeptides, of which 3 are synthesized on mitochondrial ribosomes and the remaining 4 on the cytoplasmic ribosomes.

It has been observed that human mitochondrial genes are not separated by intergenic sequences (punctuation is carried out by the tRNA genes) and that the mRNAs contain neither the leader sequence nor the tail sequence. Moreover, there seems to be only one promoter in each strand of human mitochondrial DNA, which suggests that the genome is transcribed completely and symmetrically. In the yeast mitochondrion, on the contrary, transcription seems to take place from several different promoters.

Although the coding sequences in both the genes of apocytochrome b in human and yeast mitochondria are of the same size, there is no intron in the human mitochondrial gene as opposed to 5 introns in yeast. Besides, while in man, there is neither leader sequence in 5' nor tail sequence in, in yeast there is a leader sequence of about 1,000 nucleotides and a tail sequence of about 50 nucleotides.

Amazingly, it is also revealed that the genetic code used by the mitochondria is different from the normal genetic code which was thought to be universal. For example, the UGA codes for tryptophan (instead of acting as nonsense codon) and the AUA codes for methionine (instead of coding for isoleucine). Also, AUU and AUA can be used as initiation codons (instead of being codons for isoleucine), and AGA and AGG can be used either as termination codons in the case of vertebrates or as codons for serine in the case of *Drosophila* (instead of being codons for arginine). Surprisingly, the CUN (N = any nucleotide) codes for tryptophan in yeast (instead of coding for arginine).

The mitochondrial genome of plants is also circular but larger than that of either mammals (ca 16 kbp) or yeasts (ca 78 kbp) and that its size varies, according to the species, from 200 to 2,000 kbp. In fact, only about 15 genes of proteins have been identified and sequenced at present: these are some genes coding for subunits of cytochrome oxidase, NADH dehydrogenase, ATPase and cytochrome b, and some other genes coding for proteins of the small ribosomal particles S<sub>2</sub>, S<sub>3</sub>, S<sub>4</sub>, etc. A characteristic feature of plant mitochondria is that a number of mitochondrial tRNAs are coded by the mitochondrial genome but the other mitochondrial tRNAs are coded by the nuclear genome and are imported into the mitochondrion. Moreover, a heterogeneity is observed in the size of mitochondrial DNA molecules of the same plant. This is due to the presence, in the mitochondrial genome, of *repeated sequences* which can generate sub-genomic DNA molecules by recombination.

Recently, it has been shown that in plant mitochondria, there exist a phenomenon of correction of

genetic information at the level of mRNAs, called RNA editing. It consists in modifying some cytosines and uracils and thus changing some codons. This results in the synthesis of a protein whose amino acid sequence is different from the sequence which could be deduced from that of the gene, but has more similarity in the sequence of amino acids of the corresponding mitochondrial protein in other organisms such as mammals and yeasts.

The **chloroplasts**, like mitochondria, also possess a level of autonomy. The chloroplast DNA is much larger than mitochondrial DNA and hence codes for a greater number of proteins. Each chloroplast has more than one DNA molecule and the number varies with the size of the organelle and the type of organism. As opposed to mitochondria, the size of the chloroplast genomes, which are also circular, does not vary considerably from one species to another and is about 120 to 190 kbp. In most of the species studied, the chloroplast DNA comprises a region of about 10 to 25 kbp present as two copies in opposite orientations (*i.e.*, *inverted repeats*). These inverted repeats contain the genes of the rRNAs and are separated by a large single copy region and a small single copy region.

Recently, the complete sequence of the chloroplast genome of two plants namely tobacco, *Nicotiana tabacum* (155 kbp) and a liverwort, *Marchantia* (120 kbp) was determined. The chloroplast DNA codes for the following:

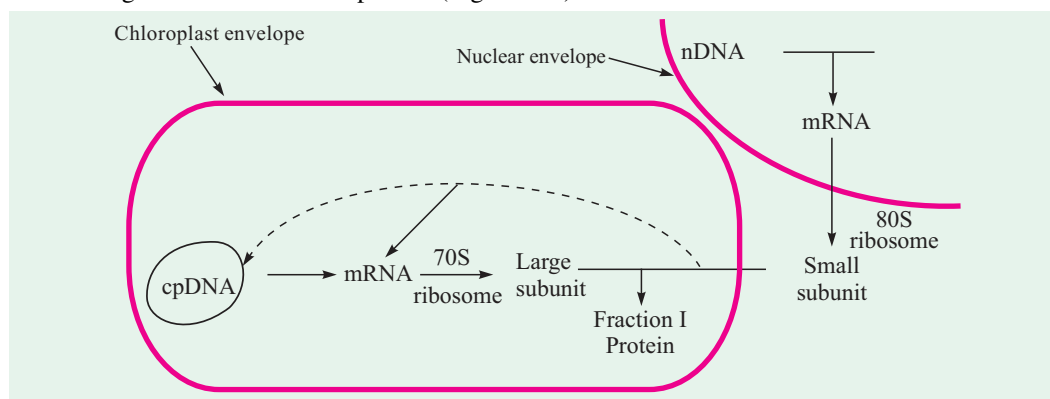
1. the chloroplast 23S, 16S, 5S and 4.5S rRNAs
2. thirty chloroplast tRNAs
3. nineteen chloroplast ribosomal proteins (11 proteins of the 30S particle + 8 proteins of the 50S particle)
4. the translation initiation factor, IF-1
5. three subunits of the RNA polymerase
6. two proteins of photosystem I, five proteins of photosystem II, six polypeptides forming part of the ATP-synthase complex, and three proteins involved in electron transport.

Besides, the genes for these 40 or so identified proteins, the chloroplast genome contains some 40 *open reading frames* (ORF) which could also code for polypeptides, but most of the chloroplast proteins are coded by the nuclear genome.

Ellis (1975) studied protein synthesis in the isolated chloroplasts of a leguminous plant, pea. The isolated chloroplasts could synthesize following proteins:

1. large subunit of Fraction I protein
2. five unidentified proteins of the internal lamellar system
3. two or three unidentified polypeptides of the envelope.

Ellis also demonstrated that while the large subunit of Fraction I protein is synthesized under the influence of chloroplast DNA, the small subunit is synthesized in the cytoplasm under the influence of nuclear DNA. Small subunit is then transported to the chloroplast where it associates with large subunit to give rise to Fraction I protein (Fig. 28–35).



**Fig. 28–35.** A model of cooperativity between nuclear DNA (nDNA) and chloroplast DNA (cpDNA) in pea

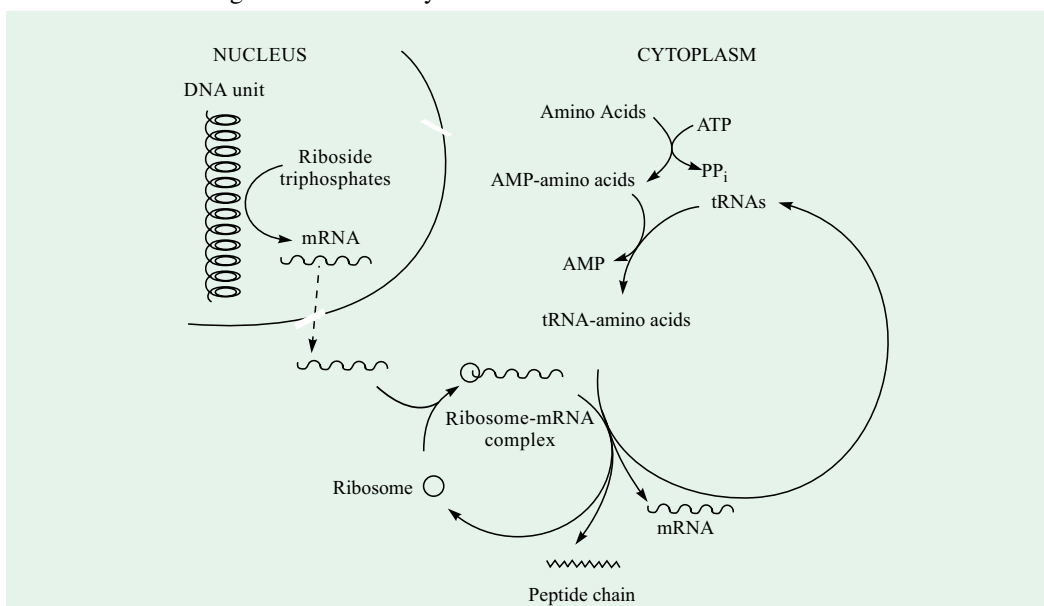
(Redrawn from Ellis RJ, 1975)



### SALIENT FEATURES OF RIBOSOMAL PROTEIN SYNTHESIS

The protein synthesis carried on ribosomes, which is summarily represented in Fig. 28–36, is a highly complex, anabolic process and occurs continuously within the cells, *i.e.*, intracellularly. Certain salient features of ribosomal protein synthesis are:

1. *DNA molecule never leaves the chromosome* during the course of eukaryotic protein synthesis. It acts through the intermediary of mRNA.



**Fig. 28–36.** A summary of the process of protein synthesis

(Adapted from Fairley JL and Kilgour GL, 1974)

2. The process of protein synthesis is *extremely accurate* and *highly fool-proof*. In fact, there is less than one error for every 10,000 amino acids polymerized into protein, *i.e.*, the error rate is about 0.01%.
3. The mechanism of protein synthesis is *cyclical*. Proteins are linear polymers of many monomer units. Therefore, this type of structure can be synthesized by a mechanism in which monomer units are added sequentially to the growing polymer, with essentially the same events of each step. It would be less efficient and certainly wasteful of genetic material if an entirely separate mechanism were used to add each amino acid onto the nascent polypeptide.

### BIOSYNTHESIS OF SHORT PEPTIDES (= ANTIBIOTIC POLYPEPTIDE SYNTHESIS)

Certain short peptides are not synthesized by ribosomes and a different mechanism of peptide bond formation operates in them. The biosynthesis of short peptides that are not under genetic control, insofar as determination of the sequence of their amino acid residues is concerned, is known in at least two instances.

#### A. Synthesis of Gramicidin S

Gramicidin S, an antibiotic produced by certain strains of *Bacillus*

Gramicidin S is also spelt as gramacidin S.

BIOSYNTHESIS OF PROTEINS 763

*brevis*, is a cyclic decapeptide made up of two identical pentapeptides joined head-to-tail (Fig. 28–37). A characteristic feature of this molecule is that it contains a D-amino acid (phenylalanine) and a nonprotein amino acid ornithine, which is used in urea cycle but not in protein synthesis.

Gramicidin S biosynthesis requires a much simpler synthetic apparatus consisting only of two enzymes,  $E_I$  and  $E_{II}$ . D-phenylalanine is activated by  $E_{II}$ , whereas the remaining 4 amino acids (L-Pro, L-Val, L-Orn, L-Leu) of the pentapeptide unit are activated by  $E_I$ . Both enzymes also participate in peptide-bond formation.

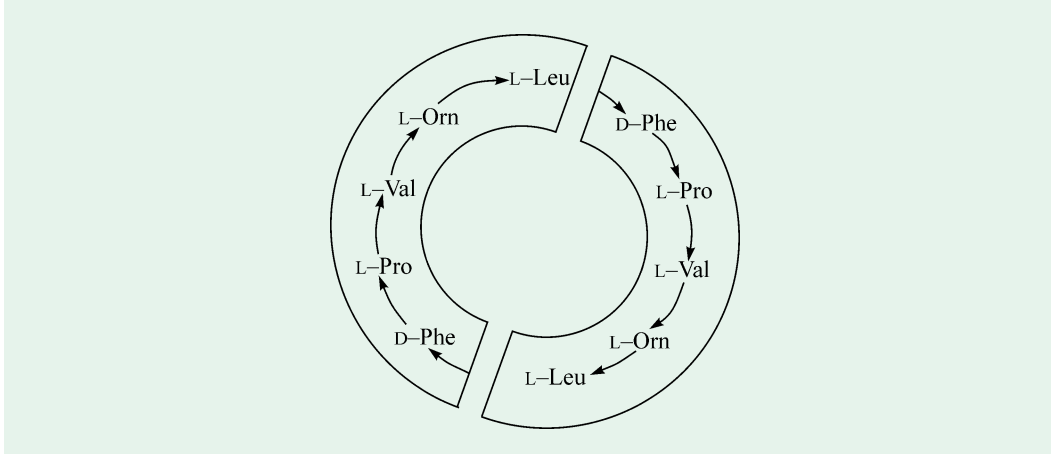
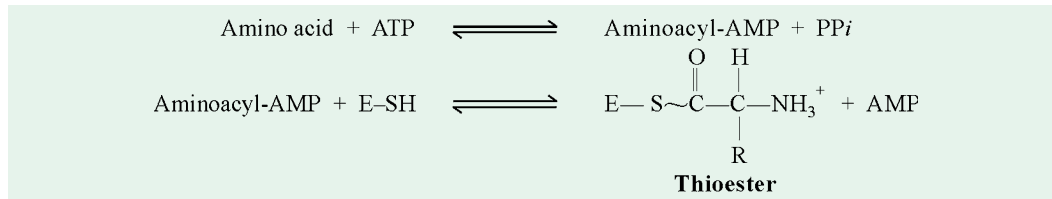


Fig. 28–37. Gramicidin S

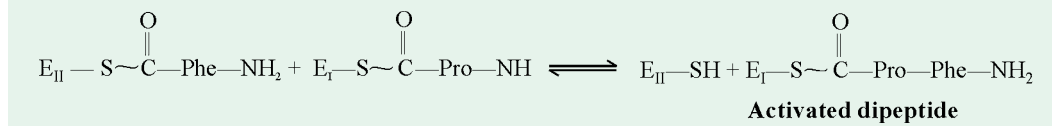
This is a cyclic peptide made of two identical pentapeptide units, joined in a head-to-tail manner. The arrows denote the polarity of the polypeptide chain (amino to carboxyl direction).

In the biosynthetic process, amino acids are activated by the formation of enzyme-bound thioesters. The activated amino acids are attached to a sulfhydryl (–SH) group of  $E_I$  or  $E_{II}$ , instead of the terminal hydroxyl group of a tRNA :



L-Proline, L-valine, L-ornithine and L-leucine form thioester linkages with specific sulfhydryls of  $E_I$  when they are incubated in the presence of ATP. Similarly, D-phenylalanine forms a thioester bond with  $E_{II}$  in the presence of ATP.

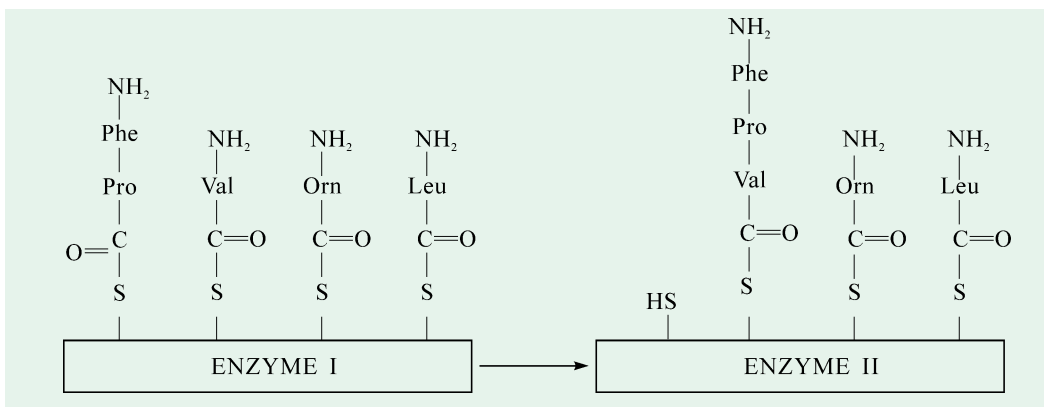
The synthesis of gramicidin S starts with the interaction of  $E_I$  and  $E_{II}$ . The D-phenylalanine residue on  $E_{II}$  is transferred to the imino (>NH) group of the L-proline residue on  $E_I$  to form a dipeptide:



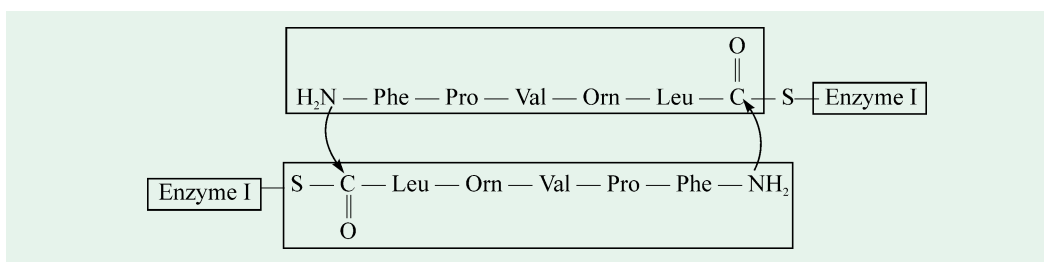
The subsequent reactions require only  $E_I$ . The activated carbonyl group of the proline residue in the dipeptide reacts with the amino group of the valine residue on the same enzyme to yield a tripeptide. This process is repeated with ornithine and then with leucine to produce an enzyme-bound pentapeptide.

## 764 FUNDAMENTALS OF BIOCHEMISTRY

The growing peptide is transferred to a different sulfhydryl each time a peptide bond is formed.



Finally, the activated pentapeptides attached to two different  $E_1$  molecules react with each other to form cyclic gramicidin S.

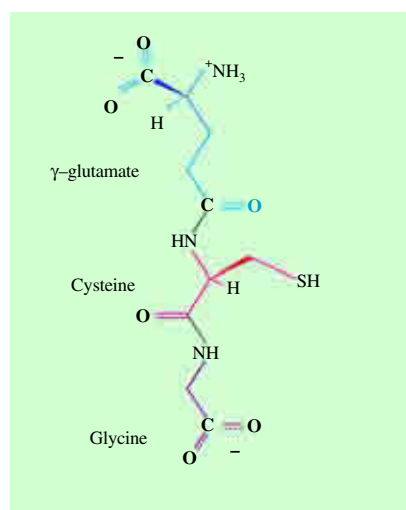


Two noteworthy features of this biosynthetic pathway are:

- The amino acid sequence of gramicidin S is determined by the spatial arrangement and specificity of the enzymes I and II. There is at least one protein subunit for each peptide bond formed. Therefore, *this mode of synthesis is uneconomical as compared to the ribosomal protein synthesis*. With the result, a peptide containing more than about 15 amino acid residues is not synthesized by this mechanism.
- The synthesis of gramicidin S resembles that of fatty acid*, because the activated intermediates in both processes are thioesters. Furthermore, E contains a covalently-attached phosphopantetheine residue. This thiol probably carries the growing peptide chain from one site to the next in  $E_1$ . *This antibiotic polypeptide synthesis may be a surviving relic of a primitive mechanism of protein synthesis used early in evolution* (Fritz Lipmann, 1971). *Ribosomal protein synthesis may have evolved from fatty acid synthesis*.

### B. Synthesis of Glutathione

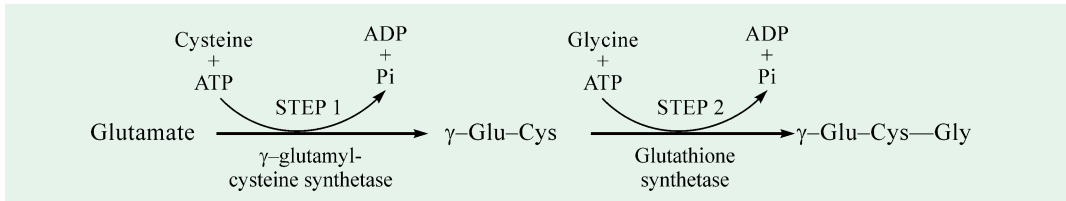
Glutathione (Fig. 28–38), a tripeptide containing a sulfhydryl group, is a highly distinctive amino acid



**Fig. 28-38. Glutathione (reduced)**  
γ-glutamylcysteinylglycine

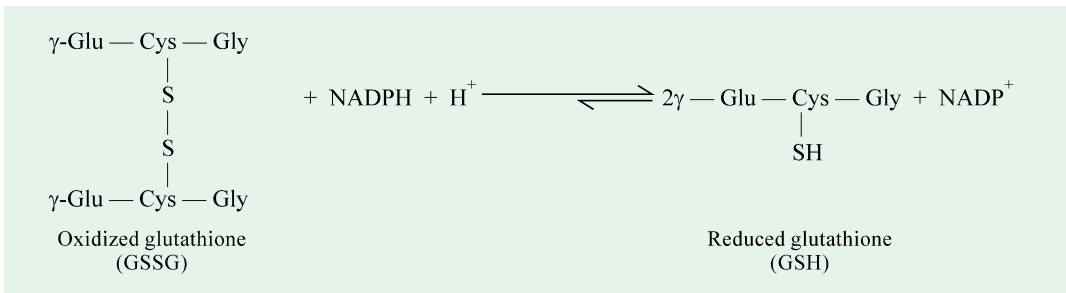
BIOSYNTHESIS OF PROTEINS 765

derivative performing a number of important physiological functions. It was isolated in 1921 by Frederick Gowland Hopkins. The discovery of this sulfur-containing compound gave a great impetus to the study of the complicated nature of cellular oxidation and metabolism. The synthesis of glutathione is a two-step process (Fig. 28–39). Although both the steps involve participation of a mole of ATP each, there is no evidence for any other requirements. In the **first step**, a peptide linkage is formed between the  $\alpha$ -carboxyl group of glutamate and the amino group of cysteine, the reaction being catalyzed by  $\gamma$ -glutamylcysteine synthetase. Formation of this peptide bond requires activation of the  $\gamma$ -carboxyl group, which function is performed by ATP. In the **second step**, a second mole of ATP activates the carboxyl group of cysteine to enable it to condense with the amino group of glycine, the reaction being catalyzed by glutathione synthetase.

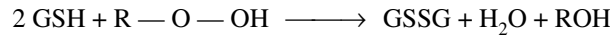


**Fig. 28–39. Synthesis of glutathione, GSH**  
( $\gamma$ -glutamylcysteinylglycine)

Glutathione cycles between a reduced sulfhydryl form (GSH) and an oxidized disulfide form (GSSG) in which two tripeptides are linked by a disulfide bond. GSSG is reduced to GSH by *glutathione reductase* (Fig. 28–40), a flavoprotein that utilizes NADPH as the electron donor. Normally, the ratio of the reduced (GSH) to oxidized (GSSG) form of glutathione in red cell is greater than 500. *The reduced (GSH) form of glutathione serves as a sulfhydryl buffer* that maintains the cysteine residues of hemoglobin and other red-cell proteins in the reduced state.

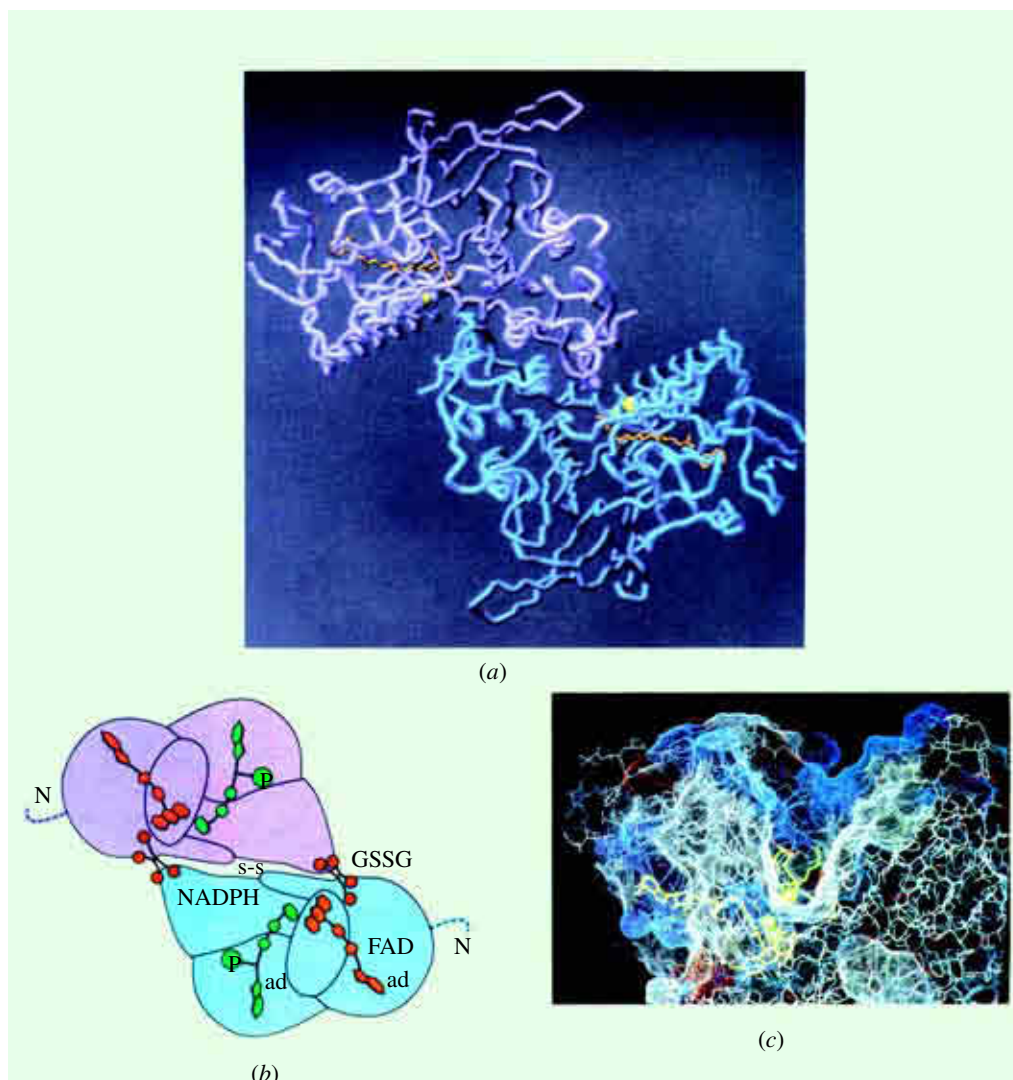


*The reduced form of glutathione also plays a key role in detoxication by reacting with  $H_2O_2$  and organic peroxides, the harmful byproducts of aerobic life. Glutathione peroxidase, the enzyme catalyzing this reaction, is remarkable in having a covalently-attached selenium (Se) atom. Its active site contains the Se analogue of cysteine, in which selenium has replaced sulfur (Epp, Ladenstein and Wendel, 1983).*



**EVOLUTION OF PROTEIN SYNTHESIS**

The mechanism underlying protein synthesis seems extremely complex. Protein synthesis in the present-day organisms centres on the ribosome, which consists of proteins arranged around a core of ribosomal RNA (rRNA) molecules. How did rRNA molecules come to play such a dominant role in the ribosome?



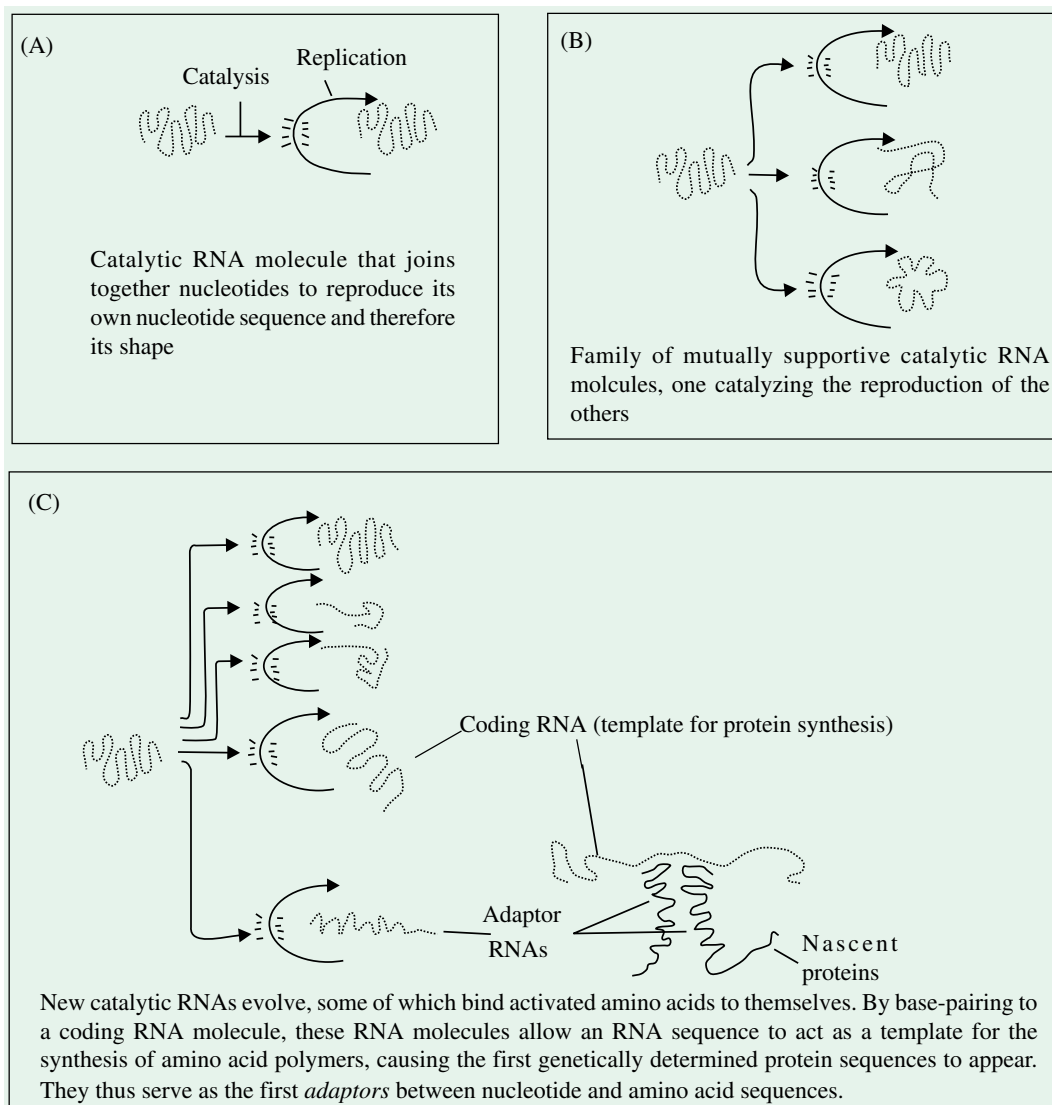
**Fig. 28–40.** The x-ray structure of the dimeric enzyme glutathione reductase as viewed along the molecule's twofold axis of symmetry

(a) The  $C_{\alpha}$  backbone with the two identical subunits shown in different colours. The S atoms of the redox-active disulfides are represented by yellow spheres and the FAD prosthetic groups are shown in orange (each flavin residue is near an active disulfide group).

(b) An interpretive diagram of Part a showing how each subunit is organized into five domains. The 18-residue N-terminal domain (dashed lines) is not visible in the x-ray structure presumably because it is flexibly linked to the rest of the protein. The binding sites of NADPH and GSSG [not shown in Part a], as well as those of FAD, are indicated. The two subunits are covalently linked by a disulfide bridge across the molecular twofold axis.

(c) The active site region of glutathione reductase showing the FAD and GSSR positions. Acidic residues (Asp and Glu) are red, basic residues (Arg, Lys, and His) are blue and all other residues are white. The dot surface, which is colored according to the nearest residue, represents the protein's solvent accessible surface. The FAD has thick yellow bonds, the GSSR has thick green bonds, and the redoxactive sulfur atoms of both the enzyme (below) and substrate (above) are represented by green spheres.

[Courtesy: (a) Arthur Olson, (b) EF Pai and GE Schulz, and (c) John Kuriyan]



**Fig. 28–41. Three successive steps in the evolution of a self-replicating system of RNA molecules, capable of directing protein synthesis**

(Adapted from Alberts, Bray, Lewis, Raff, Roberts and Watson, 1994)

Prior to the discovery of mRNA in the early 1960s, it was presumed that the large amount of RNA in ribosomes served a “messenger” function, carrying genetic information from DNA to proteins. However, it is now well known that all of the ribosomes in a cell contain an identical set of rRNA molecules that have no such information role. In bacterial ribosomes, rRNA molecules act as a catalyst in protein synthesis. As mentioned earlier, the major rRNA of the large ribosomal subunit appears to be the peptidyl transferase; besides, the rRNA of the small ribosomal subunit forms short base-paired helix with the initiation site sequence on bacterial mRNA molecules. Similarly, a variety of specific base-pair interactions form between tRNA molecules and bacterial rRNAs. This suggests that complex set of interactions take place that depend on the tertiary structure of the rRNA.

The process of protein synthesis relies heavily on a large number of proteins that are bound to the rRNAs in a ribosome. The complexity of a number of interacting components has made biochemists despair as to the understanding of the pathway by which protein synthesis evolved. The discovery that RNA molecules can act as enzymes, however, has given a new direction of viewing the pathway. Perhaps, early biological reactions used RNA molecules rather than protein molecules as catalysts. In the most primitive cells, tRNA molecules on their own may have formed catalytic surfaces that allowed them to bind and activate specific amino acids without requiring aminoacyl-tRNA synthetase enzymes. Similarly, rRNA molecules may have served by themselves as the entire "ribosome", folding up in complex ways to produce a complex set of surfaces that guided pairing of tRNA with mRNA codons as well as catalyzed the polymerization of the tRNA-linked amino acids (Fig. 28–41). During the course of evolution, individual proteins have been added to this complex machinery, each one making the process more accurate and efficient or adding regulatory controls. In this view, the large amount of RNA in present-day ribosomes is a remnant of a very early stage in evolution, before proteins dominated biological catalysis.

## REFERENCES

- Altman S (editor)** : Transfer RNA. *MIT Press, Cambridge*. 1978.
- Altman S** : Transfer RNA processing enzymes. *Cell*. **23** : 3, 1981.
- Arnstein HRV (editor)** : Synthesis of amino acids and proteins vol. 7. In MTP International Review of Sciences. Biochemistry Section. *University Park Press, Baltimore, Md.*, 1975.
- Aspirion D (editor)** : Processing of RNA. *CRC Press, Boca Raton*. 1984.
- Basavappa R, Sigler PB** : The 3 Å crystal structure of yeast initiator tRNA: Functional implications in initiator/elongator discrimination. *EMBO J*. **10** : 3105-3111, 1991.
- Berchtold H, Reshetnikova L, Reiser COA, Schirmer NK, Sprinzl M, Hilgenfeld R** : Crystal structure of active elongation factor Tu reveals major domain rearrangements. *Nature*. **365** : 126-132, 1991.
- Bermek E (editor)** : Mechanisms of Protein Synthesis. Structure-Function Relations, Control Mechanisms, and Evolutionary Aspects. *Springer-Verlag, New York*. 1985.
- Björk GR, Ericson JU, Gustafsson CED, Hagervall TG, Jönsson YH, Wikström PM** : Transfer RNA modification. *Ann. Rev. Biochem.* **56** : 263-268, 1987.
- Blobel G et al** : Translocation of proteins across membranes: the signal hypothesis and beyond. *Symp. Soc. Exp. Biol.* **33** : 9-36, 1979.
- Bradshaw RA** : Protein translocation and turnover in eukaryotic cells. *Microbiol. Rev.* **55**: 543-560, 1989.
- Brimacombe R, Stöffler G, Wittman HG** : Ribosome structure. *Ann. Rev. Biochem.* **47**: 217, 1978.
- Burbaum JJ, Schimmel P** : Structural relationships and the classification of aminoacyl-tRNA synthetases. *J. Biol. Chem.* **266**: 16965-16968, 1991.
- Calender R, Gold L (editors)** : Sequence Specificity in Transcription and Translation. *Alan R. Liss, New York*. 1985.
- Carter CW Jr.** : Cognition, mechanism and evolutionary relationships in aminoacyl-tRNA synthetases. *Ann. Rev. Biochem.* **62** : 715-748, 1993.
- Caskey CT** : Peptide chain termination. *Trends Biochem. Sci.* **5** : 234, 1980.
- Cavarelli J, Moras D** : Recognition of tRNAs by aminoacyl-tRNA synthetases. *FASEB J.* **7** : 79-86, 1993.
- Cech TR** : RNA as an enzyme. *Sci. Amer.* **255(5)**: 64-75, 1986.



18. **Chambliss G (editor)** : Ribosomes: Structure, Function and Genetics. *University Park Press, Baltimore*. 1980.
19. **Chambon P** : Eukaryotic nuclear RNA polymerases. *Ann. Rev. Biochem.* **44**: 613, 1975.
20. **Choe S, Bennett MJ, Fujii G, Curmi PM, Kantardjieff KA, Collier RJ, Eisenberg D** : The crystal structure of diphtheria toxin. *Nature.* **357**: 216-222, 1992.
21. **Clark B** : The elongation of protein synthesis. *Trends Biochem. Sci.* **5** : 207, 1980.
22. **Craigie WJ, Lee CC, Caskey CT** : Recent advances in peptide chain termination. *Mol. Microbiol.* **4** : 861-865, 1990.
23. **Dahlberg AE** : The functional role of ribosomal RNA in protein synthesis. *Cell.* **57**: 525-529, 1989.
24. **Freifelder D** : Molecular Biology. *2nd edition. Jones and Bartlett*. 1987.
25. **Gesteland RF, Atkins JF (editors)** : The RNA World. *Cold Spring Harbour Laboratory Press, New York*. 1993.
26. **Gluzman Y (editor)** : Eucaryotic Transcription. *Cold Spring Harbour Press, New York*. 1985.
27. **Grunberg-Manago M, Gros F** : Initiation mechanism of protein synthesis. *Prog. Nucleic Acid Res. Mol. Biol.* **20** : 209, 1977.
28. **Gualerzi CO, Pon CL** : Initiation of mRNA translation in prokaryotes, *Biochemistry.* **29**: 5881-5889, 1990.
29. **Hardesty B, Kramer G** : Structure, Function and Genetics of Ribosomes. *Springer-Verlag, New York*. 1986.
30. **Hershey JWB** : Overview: Phosphorylation and transport control. *Enzyme.* **44 (1-4)** : 17-27, 1990.
31. **Hill WE, Dahlberg E, Garrett RA, Moore PB, Schlessinger D, Warner JR (editors)** : The Ribosome: Structure, Function, and Evolution. *The American Society for Microbiology, Washington, DC*. 1990.
32. **Holmes WW, Platt T, Rosenberg M** : Termination of transcription in *E.coli*. *Cell.* **32**: 1029, 1983.
33. **Hunt T** : The initiation of protein synthesis. *Trends Biochem. Sci.* **5** : 178-181, 1980.
34. **Hunt T, Prentis S, Tooze J (editors)** : DNA makes RNA makes Protein. *Elsevier, New York*. 1983.
35. **Hunter T, Karin M** : The regulation of transcription by phosphorylation. *Cell.* **70** : 375-387, 1992.
36. **Jacques N, Dreyfus M** : Translation initiation in *Escherichia coli*: Old and new questions. *Mol. Microbiol.* **4 (7)** : 1063-1067, 1990.
37. **Jiménez A** : Inhibitors of translation. *Trends Biochem. Sci.* **1** : 28-30, 1988.
38. **Jukes T** : Evolution of the amino acid code: Interferences from mitochondrial codes. *J. Mol. Evol.* **19** : 219, 1983.
39. **Kirkwood TBL, Rosenberger RF, Galas DJ (editors)** : Accuracy in Molecular Processes. Its Control and Relevance to Living Systems. *Chapman and Hall*. 1986.
40. **Kjeldgaard M, Nissen P, Thirup S, Nyborg J** : The crystal structure of elongation factor EF-Tu from *Thermus aquaticus* in the GTP conformation. *Structure* **1** : 35-50, 1993.
41. **Kozak M** : Evolution of the 'Scanning Model' for initiation of protein synthesis in eucaryotes. *Cell.* **22** : 7, 1980.
42. **Kozak M** : Comparison of initiation of P.S. in procaryotes, eucaryotes, and organelles. *Microbiol. Rev.* **47** : 1-45, 1983.

## 770 FUNDAMENTALS OF BIOCHEMISTRY

43. **Kozak M** : Regulation of translation in eukaryotic systems. *Ann. Rev. Cell Biol.* **8** : 197-225, 1992.
44. **Kurland CG** : Translational accuracy and the fitness of bacteria. *Ann. Rev. Genet.* **26** : 29-50, 1992.
45. **Lake JA** : The ribosome. *Sci. Amer.* **245(2)** : 84-97, 1981.
46. **Losick R, Chamberlin M (editors)** : RNA Polymerases. *Cold Spring Harbour, New York.* 1976.
47. **Maitra U, Stringer EA, Chaudhuri A** : Initiation factors in protein biosynthesis. *Ann. Rev. Biochem.* **51** : 869-900, 1982
48. **Merrick WC** : Overview: Mechanism of translation initiation in eukaryotes. *Enzyme.* **44(1-4)** : 7-16, 1990.
49. **Merrick WC** : Mechanism and regulation of eukaryotic protein synthesis. *Microbiol. Rev.* **56(2)** : 291-315, 1992.
50. **Moldave K (editor)** : RNA and Protein Synthesis. *Academic Press Inc., New York.* 1981.
51. **Moldave K** : Eukaryotic protein synthesis. *Ann. Rev. Biochem.* **54** : 1109-1149, 1985.
52. **Moldave K, Grossman L (editors)** : Nucleic Acids and Protein Synthesis. Parts G and H. *Methods Enzymol.* Vols. 59 and 60. *Academic Press Inc., New York,* 1979.
53. **Moller W** : Functional aspects of ribosomal proteins. *Biochimie.* **73** : 1093-1100, 1991.
54. **Moore P** : Ribosomes. *Curr. Opin. Struct. Biol.* **1** : 258-263, 1991.
55. **Moras D** : Structural and functional relationships between aminoacyl-tRNA synthetases. *Trends Biochem. Sci.* **17** : 159-164, 1992.
56. **Nathans D** : Puromycin inhibition of protein synthesis: Incorporation of puromycin into peptide chains. *Proc. Natn. Acad. Sci., USA.* **51**: 485, 1964.
57. **Nierhaus KH** : The allosteric three-site model for the ribosomal elongation cycle: Features and future. *Biochemistry.* **24** : 4997-5008, 1990.
58. **Noller HF** : Structure of ribosomal RNA. *Ann. Rev. Biochem.* **53** : 119, 1984.
59. **Noller HF** : Ribosomal RNA and translation. *Ann. Rev. Biochem.* **60** : 191-227, 1991.
60. **Noller HF** : tRNA-rRNA interactions and peptidyl transferase. *FASEB J.* **7** : 87-89, 1993.
61. **Noller HF** : Peptidyl transferase: protein, ribonucleoprotein or RNA? *J. Bacteriol.* **175(17)**: 5297-5300, 1993.
62. **Nomura M** : The control of ribosome synthesis. *Sci. Amer.* **250(1)** : 102-114, 1984.
63. **Nomura M, Tissieres A, Lengyel P (editors)** : Ribosomes. *Cold Spring Harbour, New York.* 1974.
64. **Normanly J, Abelson J** : Transfer RNA identity. *Ann. Rev. Biochem.* **58** : 1029-1049, 1989.
65. **Olsen GJ, Woese CR** : Ribosomal RNA: A key to phylogeny. *FASEB J.* **7** : 113-123, 1993.
66. **Powers T, Noller HF** : Evidence for functional interaction between elongation factor Tu and 16S ribosomal RNA. *Proc. Nat. Acad. Sci.* **90** : 1364-1368, 1993.
67. **Proud CG** : Protein phosphorylation in translational control. *Curr. Top. Cell Regul.* **32** : 243-369, 1992.
68. **Rich A** : Polyribosomes. *Sci. Amer.* **209(6)** : 44-53, 1963.
69. **Rich A** : Transfer RNA: Structure and biological function. *Trends Biochem. Sci.* **3** : 34, 1978.
70. **Rich A, Kim SH** : The three-dimensional structure of transfer RNA. *Sci. Amer.* **238(1)** : 52-62, 1978.

71. **Riis B, Rattan SI, Clark BF, Merrick WC** : Eukaryotic protein elongation factors. *Trends Biochem. Sci.* **15(11)** : 420-424, 1990.
72. **Roberts JM** : Turning DNA replication on and off. *Curr. Opin. Cell Biol.* **5** : 201-206, 1993.
73. **Sachs AB** : Messenger RNA degradation in eukaryotes. *Cell.* **74** : 413-421, 1993.
74. **Samuel CE** : The eIF-2 alpha protein kinases, regulators of translation in eukaryotes from yeasts to humans. *J. Biol. Chem.* **268** : 7603-7606, 1993.
75. **Schimmel P** : Parameters for the molecular recognition of transfer RNAs. *Biochemistry.* **28** : 2747-2759, 1989.
76. **Schimmel P, Söll D, Albelson J (editors)** : Transfer RNA: Structure, Properties and Recognition. *Cold Spring Harbour, New York.* 1979.
77. **Sentenac A** : Eukaryotic RNA polymerases. *CRC Crit. Rev. Biochem.* **18** : 31-91, 1985.
78. **Shatkin AJ** : mRNA cap binding proteins: Essential factors for initiating translation. *Cell.* **40** : 223-224, 1985.
79. **Söll D** : The accuracy of aminoacylation— ensuring the fidelity of the genetic code. *Experimentia.* **46(11-12)** : 1089-1096, 1990.
80. **Söll D, Albelson J, Schimmel P (editors)** : Transfer RNA: Biological Aspects. *Cold Spring Harbour, New York.* 1980.
81. **Spirin AS** : Ribosomal translocation: Factors and Models. *Prog. Nucleic Acid Res. Mol. Biol.* **32** : 75, 1985.
82. **Sprin AS** : Ribosome Structure and Protein Biosynthesis. *Benjamin/Cummings, Menlo Park, California.* 1986.
83. **Sprinzi M** : Elongation factor Tu: A regulatory GTPase with an integrated effector. *Trends Biochem. Sci.* **19** : 245-250, 1994.
84. **Stewart PR, Letham DS (editors)** : The Ribonucleic Acids. *2nd ed. Springer-Verlag, New York.* 1977.
85. **Szekely M** : From DNA to Protein. *John Wiley and Sons, New York.* 1980.
86. **Tai P-C, Wallace BJ, Davis BD** : Streptomycin causes misreading of natural messenger by interacting with ribosomes after initiation. *Proc. Nat. Acad. Sci.* **75** : 275-279, 1978.
87. **Varshavsky A** : The N-end rule. *Cell.* **69** : 725-735, 1992.
88. **Von Hippel PH, Bear DG, Morgan WD, McSwiggen JA** : Protein–nucleic acid interactions in transcription: A molecular analysis. *Ann. Rev. Biochem.* **53** : 389, 1984.
89. **Watson JD** : The involvement of RNA in the synthesis of proteins. *Science.* **140** : 17-26, 1963.
90. **Yonath A** : Three dimensional crystals of ribosomal particles. *Trends Biochem. Sci.* **9** : 227, 1984.
91. **Young RA** : RNA polymerase II. *Ann. Rev. Biochem.* **60** : 689-715, 1991.
92. **Zamecnik P** : The machinery of protein synthesis. *Trends Biochem. Sci.* **9** : 464-466, 1984.

## PROBLEMS

1. What is the smallest number of molecules of ATP and GTP consumed in the synthesis of a 200-residue protein, starting from amino acids ? Assume that the hydrolysis of  $PP_i$  is equivalent to the hydrolysis of ATP for this calculation.
2. EF-Tu, a member of the G-protein family, plays a crucial role in the elongation process of translation. Suppose that a slowly hydrolyzable analogue of GTP were added to an elongating system. What would be the effect on rate of protein synthesis ?

## 772 FUNDAMENTALS OF BIOCHEMISTRY

3. What is the nucleophile in the reaction catalyzed by peptidyl transferase ? Write out a plausible mechanism for this reaction.
4. Which protein in G-protein cascades plays a role similar to that of elongation factor Ts ?
5. Eukaryotic elongation factor 2 is inhibited by ADP ribosylation catalyzed by diphtheria toxin. What other G proteins are sensitive to this mode of inhibition ?
6. The E site may not require codon recognition. why ?
7. Suppose that the probability of making a mistake in translation at each translational step is a small number,  $\delta$ . Show that the probability,  $p$ , that a given protein molecule, containing  $n$  residues, will be completely error-free is  $(1 - \delta)^n$ .
8. Assume that the translational error frequency,  $\delta$ , is  $1 \times 10^{-4}$ .
  - (a) Calculate the probability of making a perfect protein of 100 residues.
  - (b) Repeat for a 1,000-residue protein.
9. A given sequence of bases in an mRNA will code for one and only one sequence of amino acids in a polypeptide, if the reading frame is specified. From a given sequence of amino acid residues in a protein such as cytochrome *c*, can we predict the base sequence of the unique mRNA that coded for it ? Give reasons for your answer.
10. The chemical mechanisms used to avoid errors in protein synthesis are different from those used during DNA replication. DNA polymerases utilize a 3'→5' exonuclease proofreading activity to remove mispaired nucleotides incorrectly inserted into a growing DNA strand. There is no analogous proofreading function on ribosomes; and, in fact, the identity of amino acids attached to incoming tRNAs and added to the growing polypeptide is never checked. A proofreading step that hydrolyzed the last peptide bond formed when an incorrect amino acid was inserted into a growing polypeptide (analogous to the proofreading step of DNA polymerases) would actually be chemically impractical. Why ? (Hint : Consider how the link between the growing polypeptide and the mRNA is maintained during the elongation phase of protein synthesis; see Figs. 28–23 and 28–24).
11. Although the Shine–Dalgarno sequences vary considerably in different genes, they include examples like GAGGGG that could serve as code– in this case, for Glu–Gly. Does this imply that the sequence Glu–Gly cannot ever occur in a protein, lest it be read as a Shine–Dalgarno sequence ? Speculate.