CHAPTER 17
FROM GENE TO PROTEIN

Section A: The Connection Between Genes and Proteins

1. The study of metabolic defects provided evidence that genes specify proteins
2. Transcription and translation are the two main processing linking gene to protein: *an overview*
3. In the genetic code, nucleotide triplets specify amino acids
4. The genetic code must have evolved very early in the history of life

Introduction

- The information content of DNA is in the form of specific sequences of nucleotides along the DNA strands.
- The DNA inherited by an organism leads to specific traits by dictating the synthesis of proteins.
- Proteins are the links between genotype and phenotype.
  - For example, Mendel's dwarf pea plants lack a functioning copy of the gene that specifies the synthesis of a key protein, gibberellins.
  - Gibberellins stimulate the normal elongation of stems.

- Research refined the one gene - one enzyme hypothesis.
- First, it became clear that not all proteins are enzymes and yet their synthesis depends on specific genes.
  - This tweaked the hypothesis to *one gene - one protein*.
- Later research demonstrated that many proteins are composed of several polypeptides, each of which has its own gene.
- Therefore: the **one gene - one polypeptide hypothesis**.

2. Transcription and translation are the two main processes linking gene to protein: an overview

- Genes provide the instructions for making specific proteins.
- The bridge between DNA and protein synthesis is RNA.
- RNA is chemically similar to DNA, except that it contains ribose as its sugar and substitutes the nitrogenous base uracil for thymine.
  - An RNA molecule almost always consists of a single strand.

- In DNA or RNA, the four nucleotide monomers act like the letters of the alphabet to communicate information.
- The specific sequence of hundreds or thousands of nucleotides in each gene carries the information for the primary structure of a protein, the linear order of the 20 possible amino acids.
- To get from DNA, written in one chemical language, to protein, written in another, requires two major stages, transcription and translation.

- During **transcription**, a DNA strand provides a template for the synthesis of a complementary RNA strand.
  - This process is used to synthesize any type of RNA from a DNA template.
- Transcription of a gene produces a **messenger RNA** (**mRNA**) molecule.
- During **translation**, the information contained in the order of nucleotides in mRNA is used to determine the amino acid sequence of a polypeptide.
  - Translation occurs at ribosomes.

- The basic mechanics of transcription and translation are similar in eukaryotes and prokaryotes.
- Because bacteria lack nuclei, transcription and translation are coupled.
- Ribosomes attach to the leading end of a mRNA molecule while transcription is still in progress.

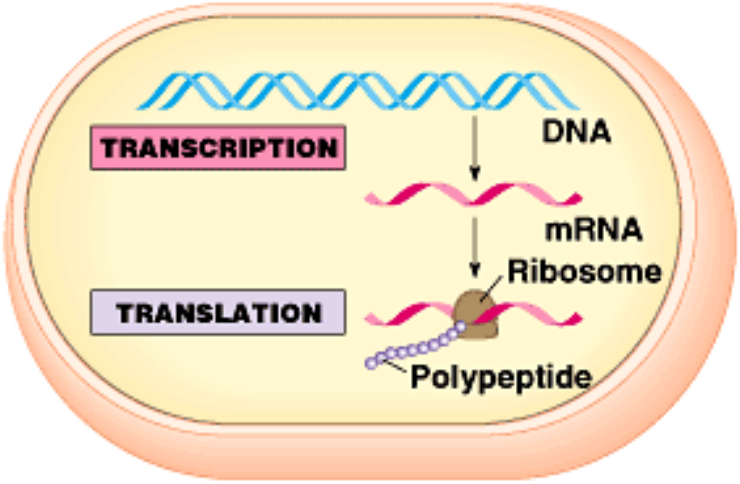


Fig. 17.2a (a) Prokaryotic cell

- In a eukaryotic cell, almost all transcription occurs in the nucleus and translation occurs mainly at ribosomes in the cytoplasm.
- In addition, before the **primary transcript** can leave the nucleus it is modified in various ways during **RNA processing** before the finished mRNA is exported to the cytoplasm.

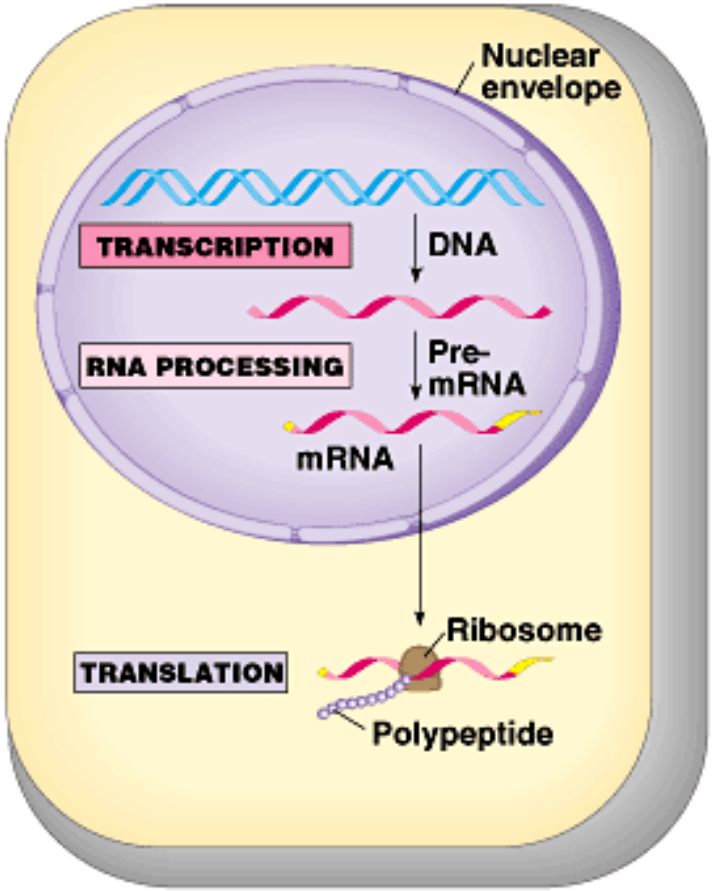


Fig. 17.2b (b) Eukaryotic cell

- Which cell type can make protein faster, prokaryote or eukaryote? Why?

- To summarize, genes program protein synthesis via genetic messenger RNA.
- The molecular chain of command in a cell is :

DNA→ RNA→ protein

This flow of information in a cell is the:

Central Dogma of Biology

3. In the genetic code, nucleotide triplets specify amino acids

- If the genetic code consisted of a single nucleotide or even pairs of nucleotides per amino acid, there would not be enough combinations (4 and 16 respectively) to code for all 20 amino acids.
- Triplets of nucleotide bases are the smallest units of uniform length that can code for all the amino acids.
- In the **triplet code**, three consecutive bases specify an amino acid, creating 4^3 (64) possible code words.
- The genetic instructions for a polypeptide chain are written in DNA as a series of three-nucleotide words.

- During transcription, one DNA strand, the **template strand**, provides a template for ordering the sequence of nucleotides in an RNA transcript.
  - The complementary RNA molecule is synthesized according to base-pairing rules, except that uracil is the complementary base to adenine.
- During translation, blocks of three nucleotides, **codons**, are decoded into a sequence of amino acids.

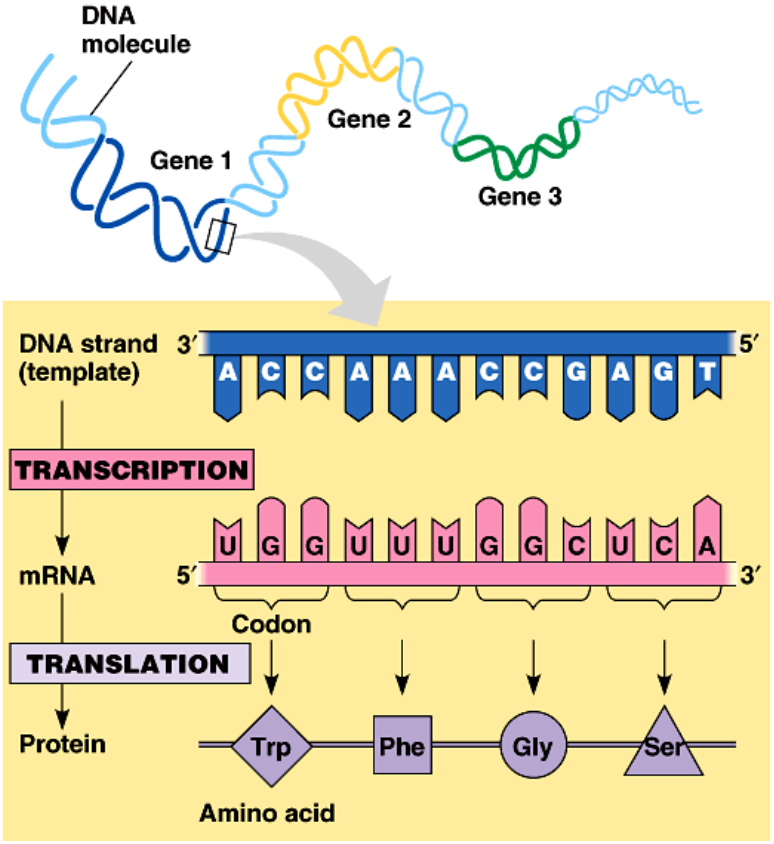


Fig. 17.3

- During translation, the codons are read in the 5’-3’ direction along the mRNA.
- Each codon specifies which one of the 20 amino acids will be incorporated at the corresponding position along a polypeptide.
- Because codons are base triplets, the number of nucleotides making up a genetic message must be three times the number of amino acids making up the protein product.
  - It would take at least 300 nucleotides to code for a polypeptide that is 100 amino acids long. (average human protein length is ~485 amino acids)

- The task of matching each codon to its amino acid counterpart began in the early 1960s.
- Marshall Nirenberg determined the first match, that UUU coded for the amino acid phenylalanine.
  - He created an artificial mRNA molecule entirely of uracil and added it to a test tube mixture of amino acids, ribosomes, and other components for protein synthesis.
  - This “poly(U)” translated into a polypeptide containing a single amino acid, phenyalanine, in a long chain.

- By the mid-1960s the entire code was deciphered.
  - 61 of 64 triplets code for amino acids.
  - The codon AUG not only codes for the amino acid methionine but also indicates the start of translation.
  - Three codons do not indicate amino acids but signal the termination of translation.

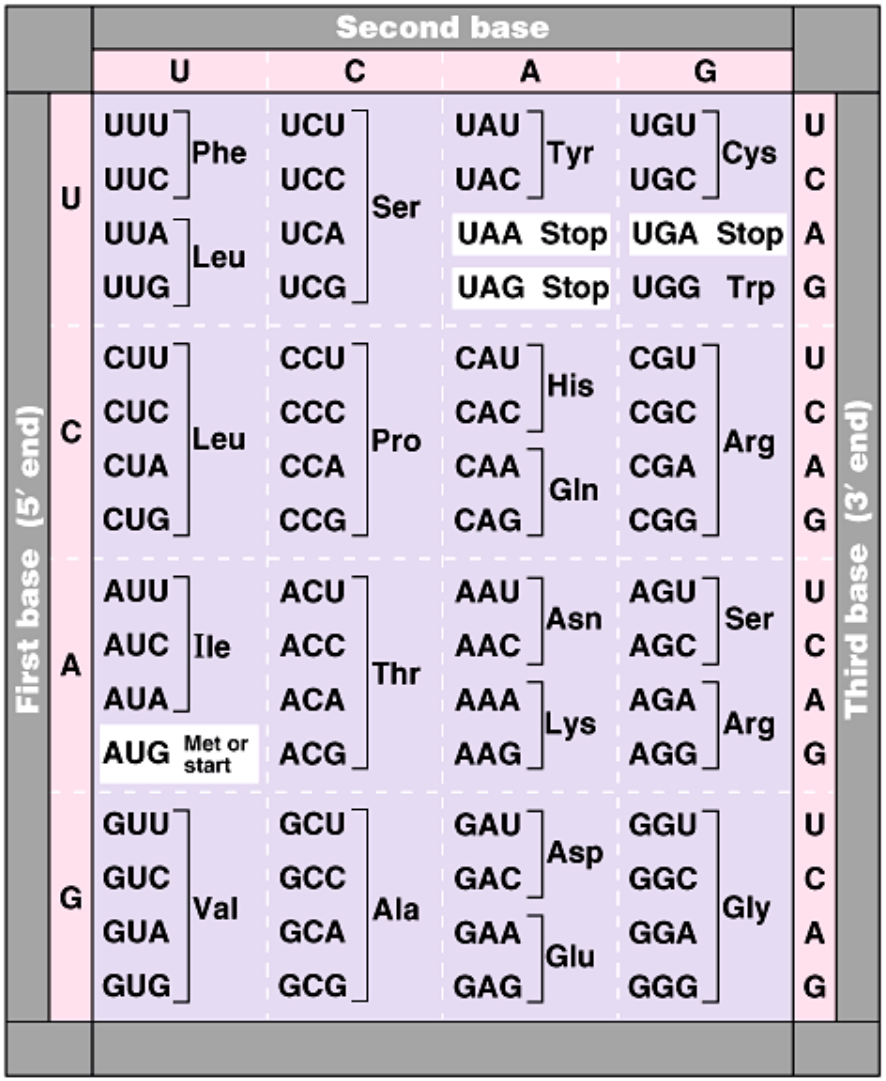


Fig. 17.4

- The genetic code is *redundant* but not *ambiguous*.
  - There are typically several different codons that would indicate a specific amino acid.
  - However, any one codon indicates only one amino acid.
    - [If you have a specific codon, you can be sure of the corresponding amino acid, but if you know only the amino acid, there may be several possible codons.]
    - Both GAA and GAG specify glutamate, but no other amino acid.
  - Codons synonymous for the same amino acid often differ only in the third codon position.

- To extract the message from the genetic code requires specifying the correct starting point.
  - This establishes the **reading frame** and subsequent codons are read in groups of three nucleotides.
  - The cell's protein-synthesizing machinery reads the message as a series of nonoverlapping three-letter words.
- In summary, genetic information is encoded as a sequence of nonoverlapping base triplets, or codons, each of which is translated into a specific amino acid during protein synthesis.

The 20 amino acids

Abbreviation	Amino acid	Abbreviation	Amino acid
Ala	Alanine	Leu	Leucine
Arg	Arginine	Lys	Lysine
Asp	Aspartic acid	Met	Methionine
Asn	Asparagine	Phe	Phenylalanine
Cys	Cysteine	Pro	Proline
Gln	Glutamine	Ser	Serine
Glu	Glutamic acid	Thr	Threonine
Gly	Glycine	Trp	Tryptophan
His	Histidine	Tyr	Tyrosine
Ile	Isoleucine	Val	Valine

Table 1. Dietary Requirements for Amino Acids in Humans	
Essential	**Nonessential**
Histidine	Alanine
Isoleucine	Arginine
Leucine	Asparagine
Lysine	Aspartate
Methionine	Cysteine
Phenylalanine	Glutamate
Threonine	Glutamine
Tryptophan	Glycine
Valine	Proline
	Serine
	Tyrosine

4. The genetic code must have evolved very early in the history of life

- The genetic code is nearly universal, shared by organisms from the simplest bacteria to the most complex plants and animals.
- In laboratory experiments, genes can be transcribed and translated after they are transplanted from one species to another.
  - This tobacco plant is expressing a firefly gene.

Fig. 17.5

- This has permitted bacteria to be programmed to synthesize certain human proteins after insertion of the appropriate human genes.
- This and other similar applications are exciting developments in biotechnology.
- Exceptions to the universality of the genetic code exist in translation systems where a few codons differ from standard ones.
  - These occur in certain single-celled eukaryotes like *Paramecium*.
  - Other examples include translation in certain mitochondria and chloroplasts.

- The near universality of the genetic code must have been operating very early in the history of life.
- A shared genetic vocabulary is a reminder of the kinship that bonds all life on Earth.

CHAPTER 17
FROM GENE TO PROTEIN

Section B: The Synthesis and Processing of RNA

1. **Transcription is the DNA-directed synthesis of RNA: *a closer look***
2. **Eukaryotic cells modify RNA after transcription**

1. Transcription is the DNA-directed synthesis of RNA: *a closer look*

- Messenger RNA is transcribed from the template strand of a gene.
- **RNA polymerase** separates the DNA strands at the appropriate point and bonds the RNA nucleotides as they base-pair along the DNA template.
- Like DNA polymerases, RNA polymerases can add nucleotides only to the 3’ end of the growing polymer.
  - Genes are read 3’-5’, creating a 5’-3’ RNA molecule.

- Specific sequences of nucleotides along the DNA mark where gene transcription begins and ends.
  - RNA polymerase attaches and initiates transcription at the **promotor**, “upstream” of the information contained in the gene, the **transcription unit**.
  - The **terminator** signals the end of transcription.
- Bacteria have a single type of RNA polymerase that synthesizes all RNA molecules.
- In contrast, eukaryotes have three RNA polymerases (I, II, and III) in their nuclei.
  - RNA polymerase II is used for mRNA synthesis.

- Transcription can be separated into three stages: initiation, elongation, and termination.

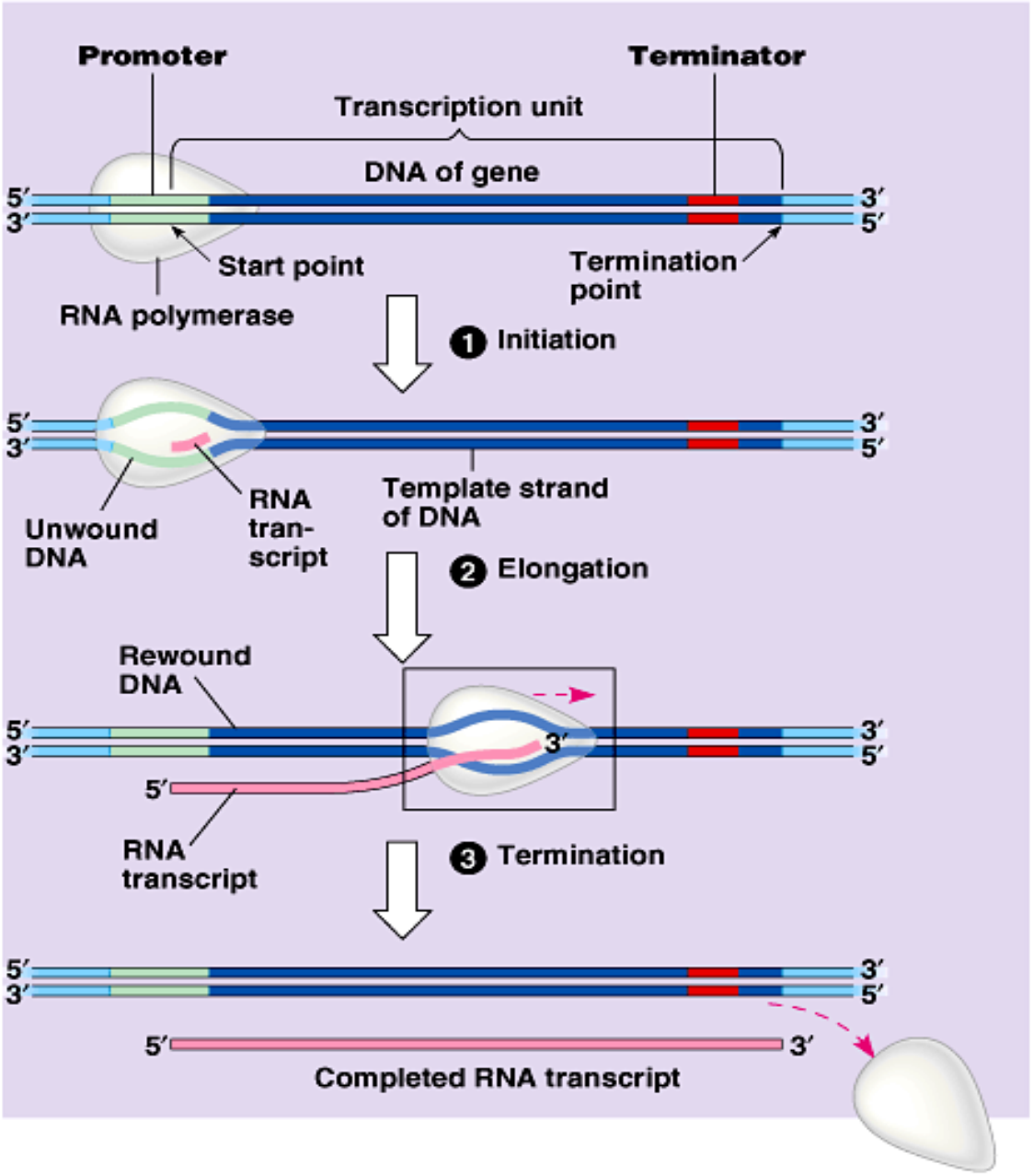


Fig. 17.6a

- The presence of a promotor sequence determines which strand of the DNA helix is the template.
  - Within the promotor is the starting point for the transcription of a gene.
  - The promotor also includes a binding site for RNA polymerase several dozen nucleotides upstream of the start point.
  - In prokaryotes, RNA polymerase can recognize and bind directly to the promotor region.

- In eukaryotes, proteins called **transcription factors** recognize the promotor region, especially a **TATA box**, and bind to the promotor.
- After they have bound to the promotor, RNA polymerase binds to transcription factors to create a **transcription initiation complex**.
- RNA polymerase then starts transcription.

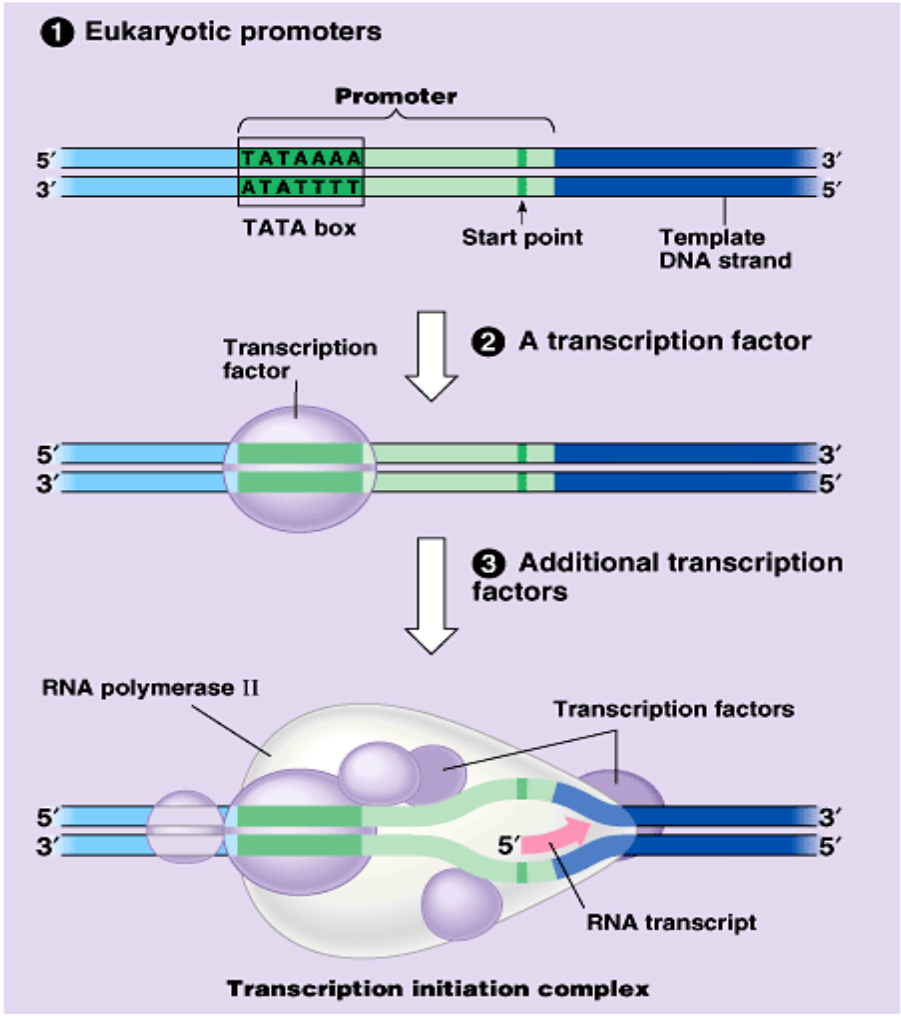


Fig. 17.7

- As RNA polymerase moves along the DNA, it untwists the double helix, 10 to 20 bases at time.
- The enzyme adds nucleotides to the 3’ end of the growing strand.
- Behind the point of RNA synthesis, the double helix re-forms and the RNA molecule moves away.

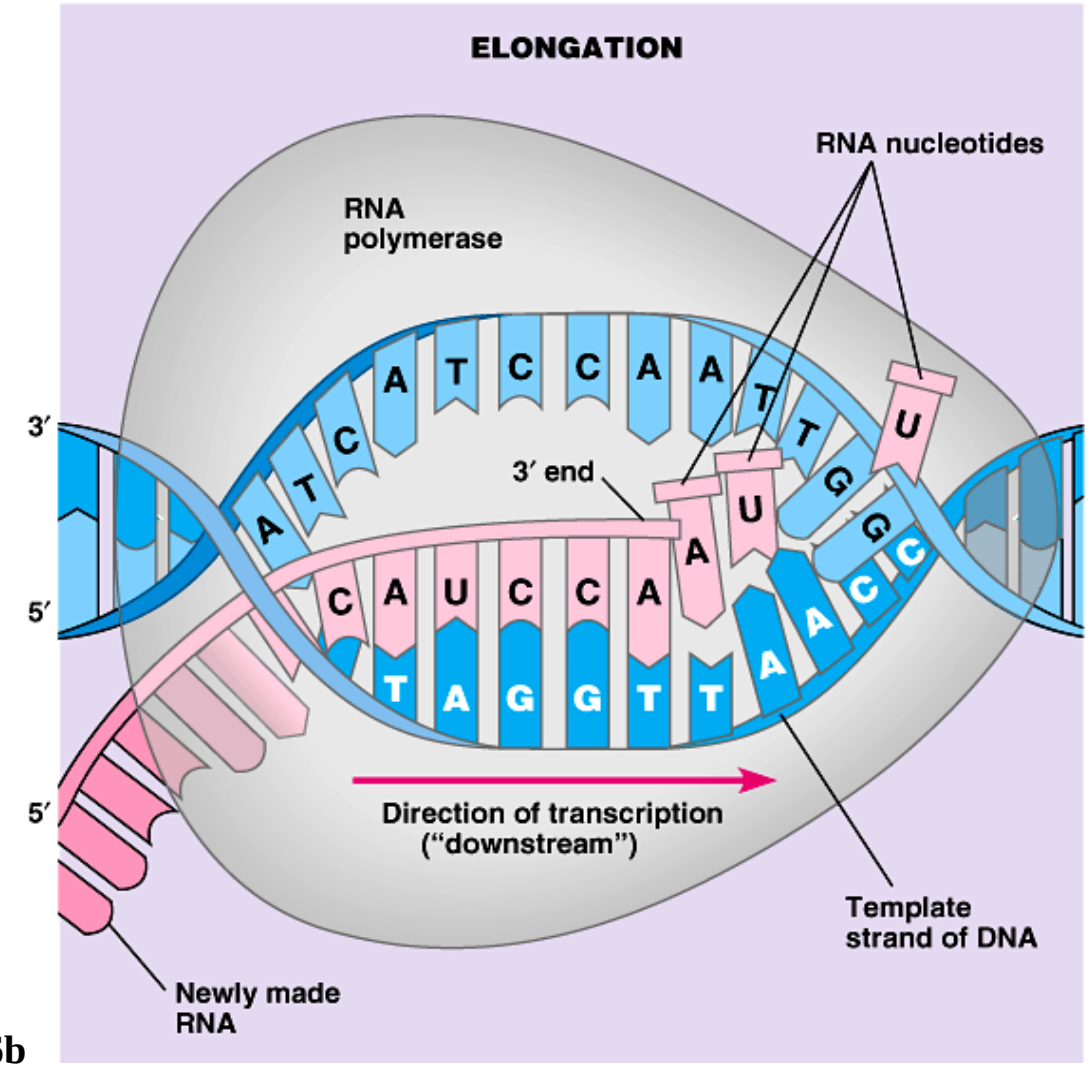


Fig. 17.6b

- A single gene can be transcribed simultaneously by several RNA polymerases at a time.
- A growing strand of RNA trails off from each polymerase.
  - The length of each new strand reflects how far along the template the enzyme has traveled from the start point.
- The congregation of many polymerase molecules simultaneously transcribing a single gene increases the amount of mRNA transcribed from it.
- This helps the cell make the encoded protein in large amounts.

- Transcription proceeds until after the RNA polymerase transcribes a terminator sequence in the DNA.
  - In prokaryotes, RNA polymerase stops transcription right at the end of the terminator.
    - Both the RNA and DNA is then released.
  - In eukaryotes, the polymerase continues for hundreds of nucleotides past the terminator sequence, AAUAAA.
    - At a point about 10 to 35 nucleotides past this sequence, the pre-mRNA is cut from the enzyme.

2. Eukaryotic cells modify RNA after transcription

- Enzymes in the eukaryotic nucleus modify pre-mRNA before the genetic messages are dispatched to the cytoplasm.
- At the 5’ end of the pre-mRNA molecule, a modified form of guanine is added, the **5’ cap**.
  - This helps protect mRNA from hydrolytic enzymes.
  - It also functions as an “attach here” signal for ribosomes.

- At the 3’ end, an enzyme adds 50 to 250 adenine nucleotides, the **poly(A) tail**.
  - In addition to inhibiting hydrolysis and facilitating ribosome attachment, the poly(A) tail also seems to facilitate the export of mRNA from the nucleus.
- The mRNA molecule also includes nontranslated leader and trailer segments.

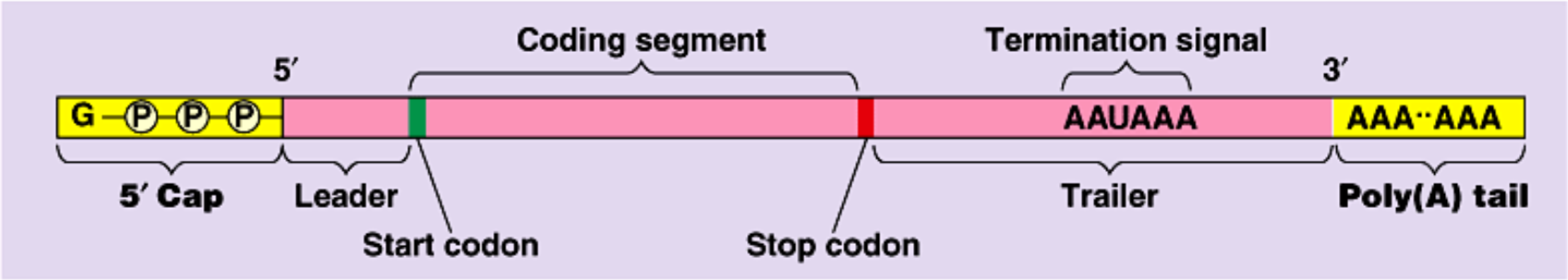


Fig. 17.8

- The most remarkable stage of RNA processing occurs during the removal of a large portion of the RNA molecule during **RNA splicing**.
- Most eukaryotic genes and their RNA transcripts have long noncoding stretches of nucleotides.
  - Noncoding segments, **introns**, lie between coding regions.
  - The final mRNA transcript includes coding regions, **exons**, that are translated into amino acid sequences, plus the leader and trailer sequences.

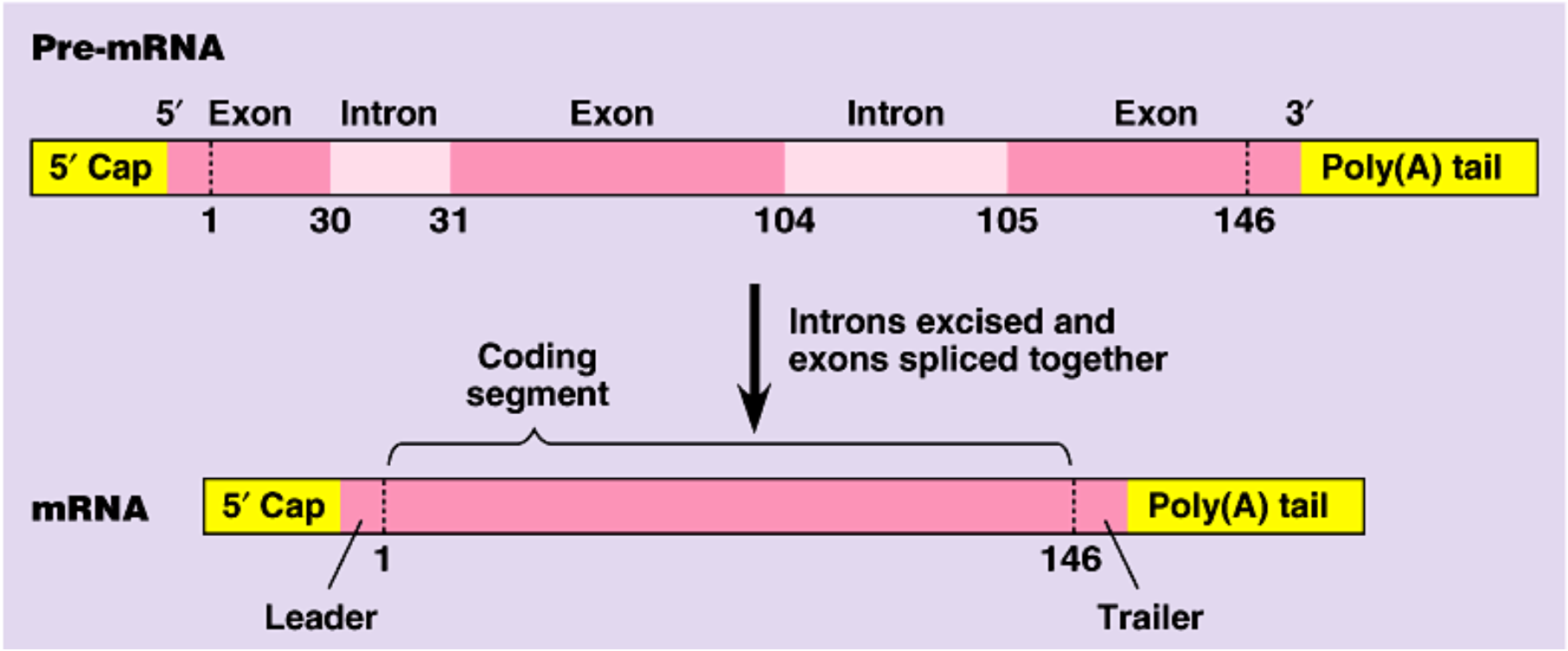


Fig. 17.9

- RNA splicing removes introns and joins exons to create an mRNA molecule with a continuous coding sequence.

- This splicing is accomplished by a **spliceosome**.
  - spliceosomes consist of a variety of proteins and several *small nuclear ribonucleoproteins* (*snRNPs*).
  - Each snRNP has several protein molecules and a *small nuclear RNA molecule* (*snRNA*).
    - Each is about 150 nucleotides long.

(1) Pre-mRNA combines with snRNPs and other proteins to form a spliceosome.

(2) Within the spliceosome, snRNA base-pairs with nucleotides at the ends of the intron.

(3) The RNA transcript is cut to release the intron, and the exons are spliced together; the spliceosome then comes apart, releasing mRNA, which now contains only exons.

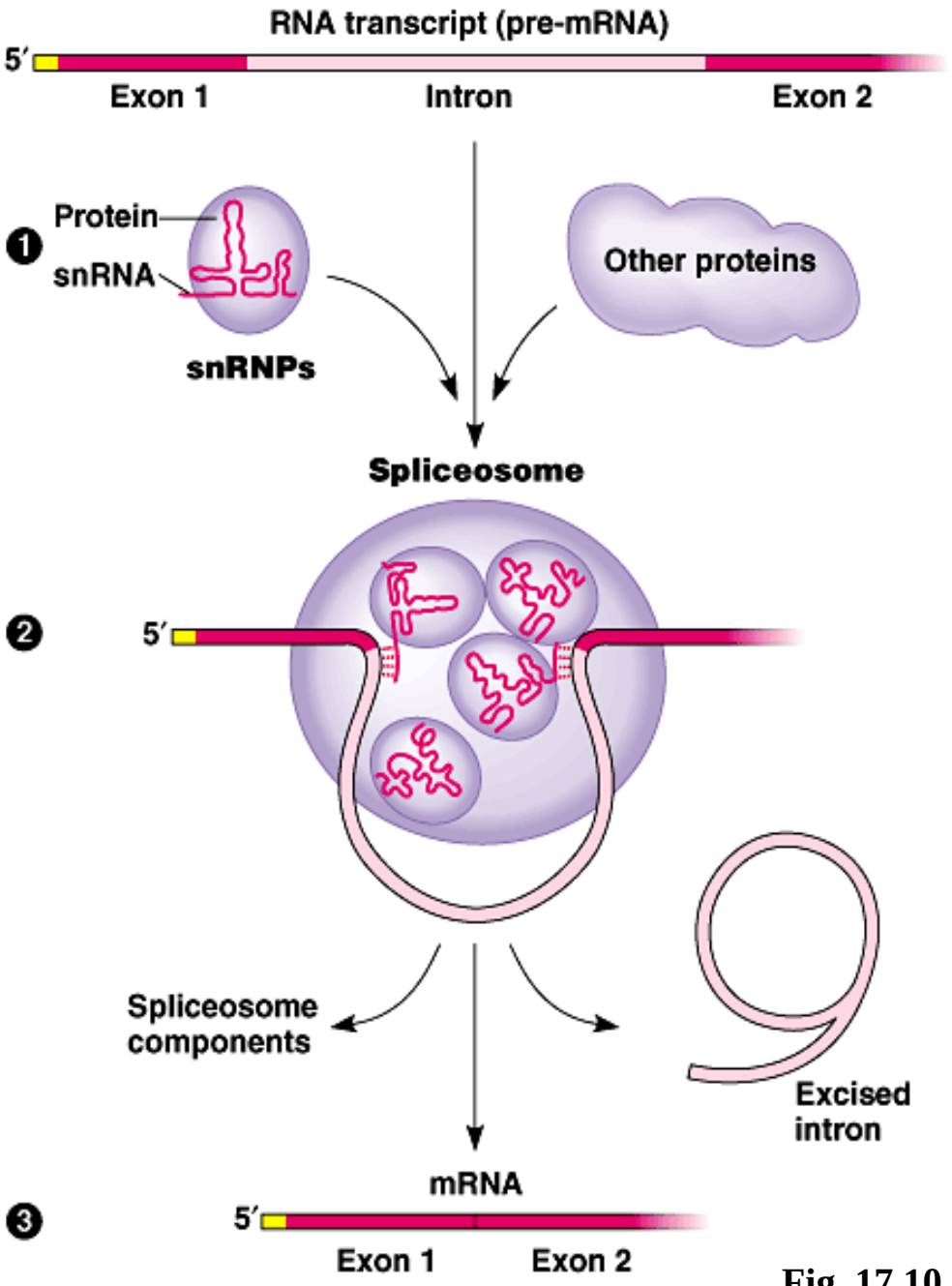


Fig. 17.10

- In this process, the snRNA acts as a **ribozyme,** an RNA molecule that functions as an enzyme.
- Like pre-mRNA, other kinds of primary transcripts may also be spliced, but by diverse mechanisms that do not involve spliceosomes.
- In a few cases, intron RNA can catalyze its own excision without proteins or extra RNA molecules.
- The discovery of ribozymes rendered obsolete the statement, “All biological catalysts are proteins.”

- RNA splicing appears to have several functions.
  - First, at least some introns contain sequences that control gene activity in some way.
  - Splicing itself may regulate the passage of mRNA from the nucleus to the cytoplasm.
  - One clear benefit of split genes is to enable a one gene to encode for more than one polypeptide.
- **Alternative RNA splicing** gives rise to two or more different polypeptides, depending on which segments are treated as exons.
  - Early results of the Human Genome Project indicate that this phenomenon may be common in humans.

- Split genes may also facilitate the evolution of new proteins.
- Proteins often have a modular architecture with discrete structural and functional regions called **domains**.
- In many cases, different exons code for different domains of a protein.

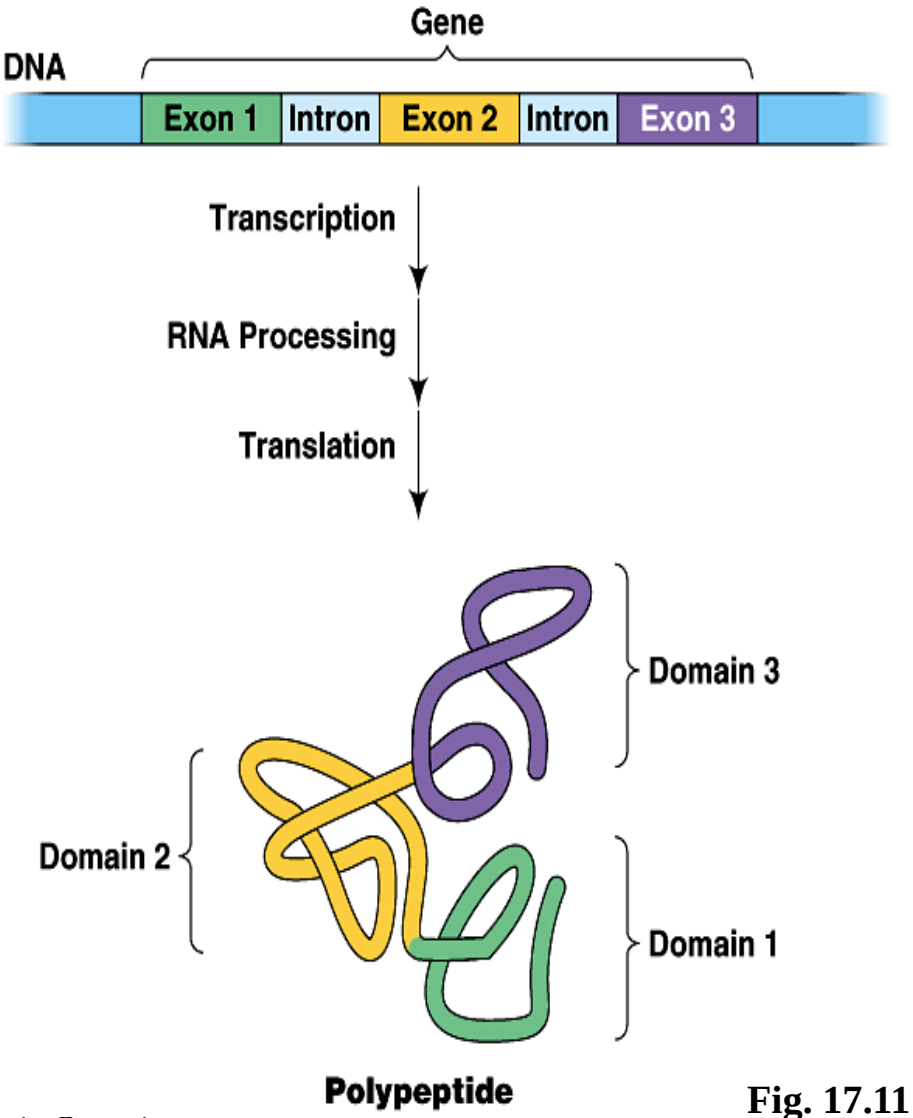


Fig. 17.11

- The presence of introns increases the probability of potentially beneficial crossing over between genes.
  - Introns increase the opportunity for recombination between two alleles of a gene.
    - This raises the probability that a crossover will switch one version of an exon for another version found on the homologous chromosome.
  - There may also be occasional mixing and matching of exons between completely different genes.
  - Either way, exon shuffling could lead to new proteins through novel combinations of functions.

CHAPTER 17
FROM GENE TO PROTEIN

Section C: The Synthesis of Protein

1. Translation is the RNA-directed synthesis of a polypeptide: *a closer look*
2. Signal peptides target some eukaryotic polypeptides to specific destinations in the cell
3. RNA plays multiple roles in the cell: *a review*
4. Comparing protein synthesis in prokaryotes and eukaryotes: *a review*
5. Point mutations can affect protein structure and function
6. What is a gene? *revisiting the question*

1. Translation is the RNA-directed synthesis of a polypeptide: a closer look

- In the process of translation, a cell interprets a series of codons along a mRNA molecule.
- **Transfer RNA** (**tRNA**) transfers amino acids from the cytoplasm's pool to a ribosome.
- The ribosome adds each amino acid carried by tRNA to the growing end of the polypeptide chain.

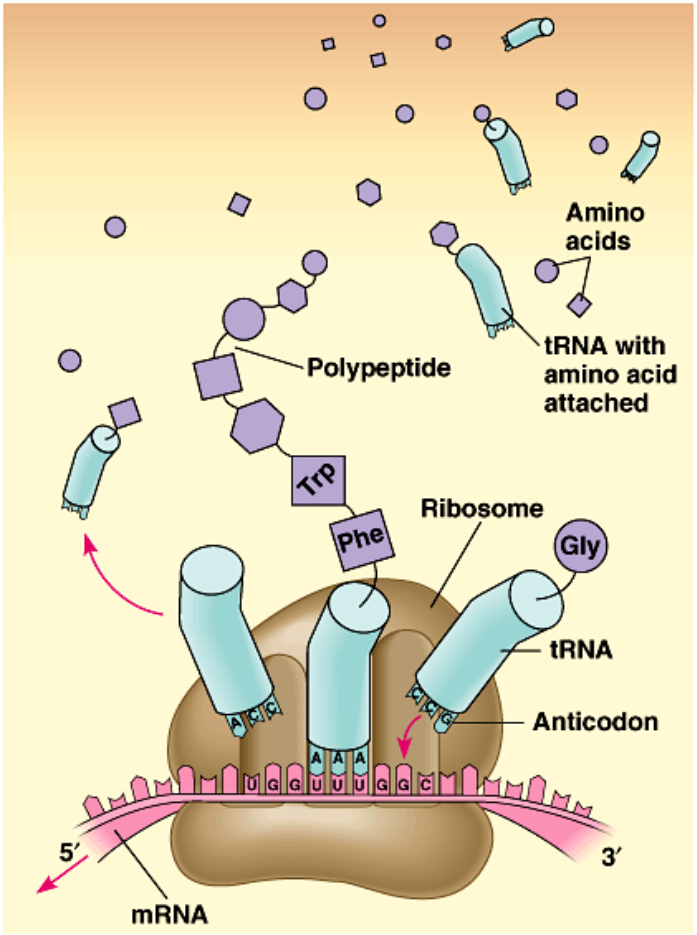


Fig. 17.12

- During translation, each type of tRNA links a mRNA codon with the appropriate amino acid.
- Each tRNA arriving at the ribosome carries a specific amino acid at one end and has a specific nucleotide triplet, an **anticodon**, at the other.
- The anticodon base-pairs with a complementary codon on mRNA.
  - If the codon on mRNA is UUU, a tRNA with an AAA anticodon and carrying phenyalanine will bind to it.
- Codon by codon, tRNAs deposit amino acids in the prescribed order and the ribosome joins them into a polypeptide chain.

- Like other types of RNA, tRNA molecules are transcribed from DNA templates in the nucleus.
- Once it reaches the cytoplasm, each tRNA is used repeatedly
  - to pick up its designated amino acid in the cytosol,
  - to deposit the amino acid at the ribosome, and
  - to return to the cytosol to pick up another copy of that amino acid.

- A tRNA molecule consists of a strand of about 80 nucleotides that folds back on itself to form a three-dimensional structure.
  - It includes a loop containing the anticodon and an attachment site at the 3’ end for an amino acid.

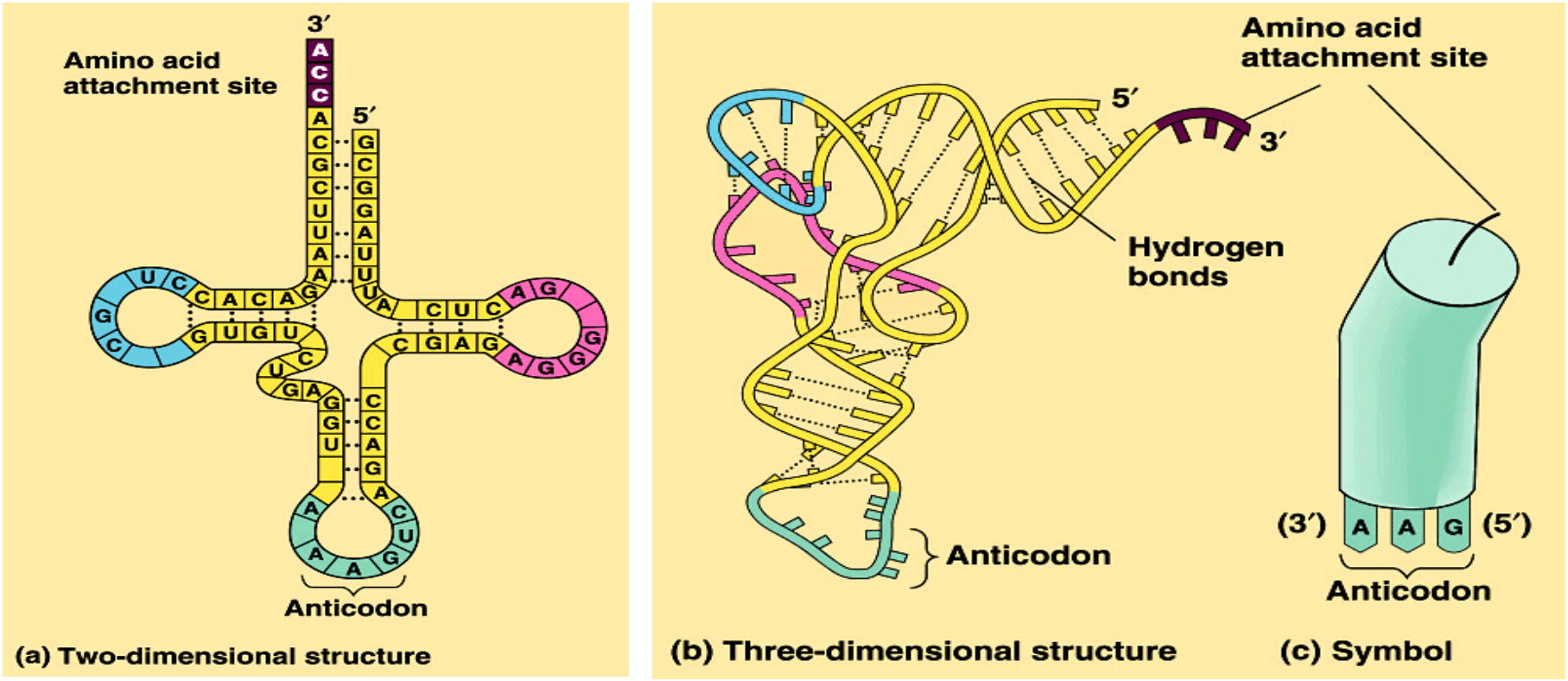


Fig. 17.13

- If each anticodon had to be a perfect match to each codon, we would expect to find 61 types of tRNA, but the actual number is about 45.
- The anticodons of some tRNAs recognize more than one codon.
- This is possible because the rules for base pairing between the third base of the codon and anticodon are relaxed (called **wobble**).
  - At the wobble position, U on the anticodon can bind with A or G in the third position of a codon.
  - Some tRNA anticodons include a modified form of adenine, inosine, which can hydrogen bond with U, C, or A on the codon.

- Each amino acid is joined to the correct tRNA by **aminoacyl-tRNA synthetase.**
- The 20 different synthetases match the 20 different amino acids.
  - Each has active sites for only a specific tRNA and amino acid combination.
  - The synthetase catalyzes a covalent bond between them, forming aminoacyl-tRNA or activated amino acid.

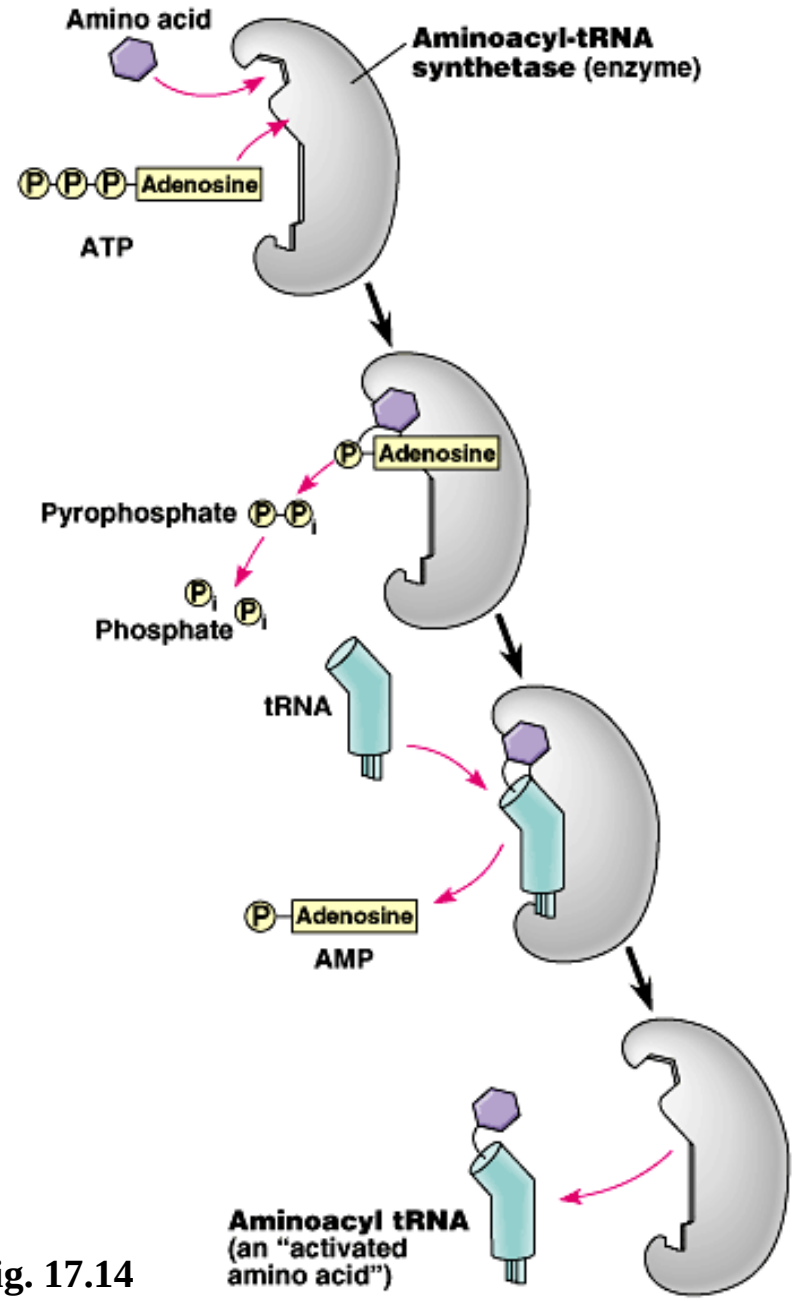
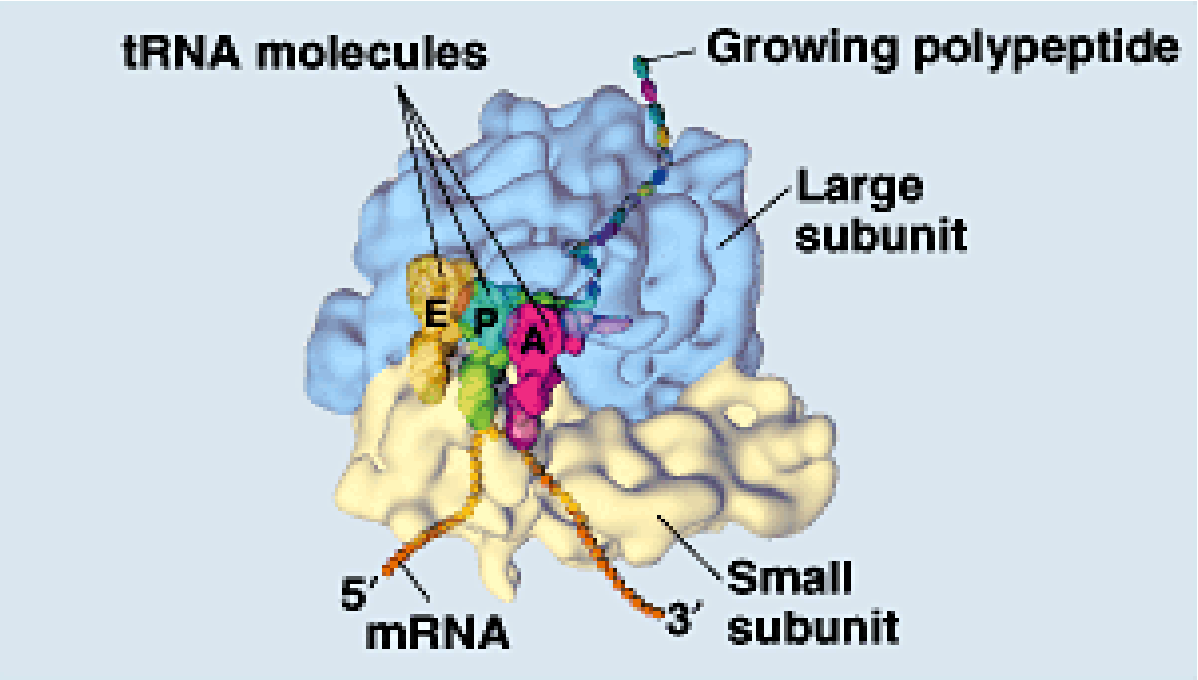


Fig. 17.14

- Ribosomes facilitate the specific coupling of the tRNA anticodons with mRNA codons.
  - Each ribosome has a large and a small subunit.
  - These are composed of proteins and ribosomal RNA (rRNA), the most abundant RNA in the cell.

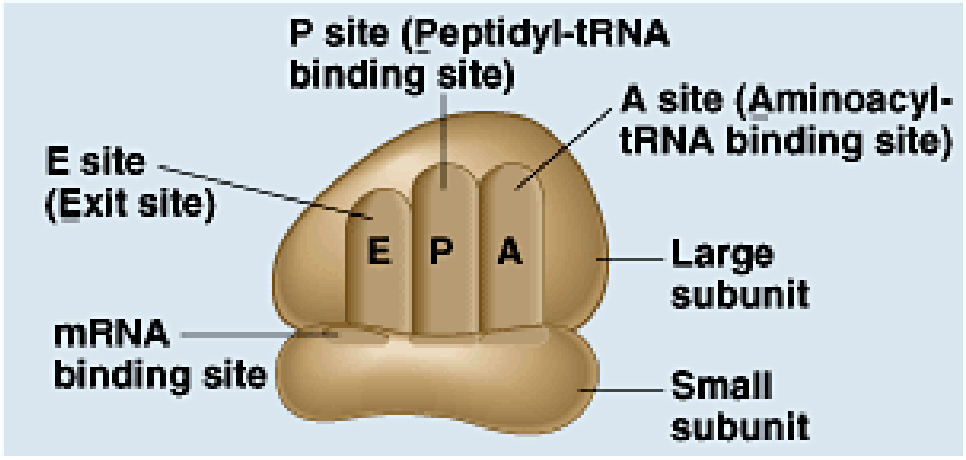


(a) Computer model of functioning ribosome

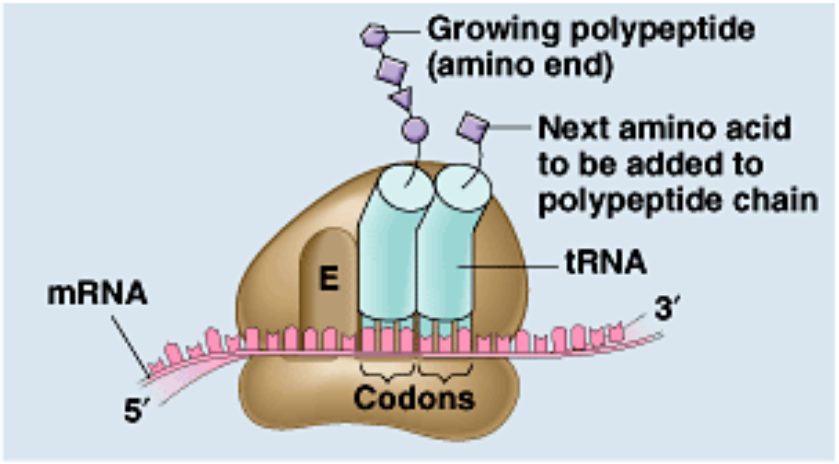
Fig. 17.15a

- After rRNA genes are transcribed to rRNA in the nucleus, the rRNA and proteins form the subunits in the nucleolus.
- The subunits exit the nucleus via nuclear pores.
- The large and small subunits join to form a functional ribosome only when they attach to an mRNA molecule.
- While very similar in structure and function, prokaryotic and eukaryotic ribosomes have enough differences that certain antibiotic drugs (like tetracycline) can paralyze prokaryotic ribosomes without inhibiting eukaryotic ribosomes.

- Each ribosome has a binding site for mRNA and three binding sites for tRNA molecules.
  - The **P site** holds the tRNA carrying the growing polypeptide chain.
  - The **A site** carries the tRNA with the next amino acid.
  - Discharged tRNAs leave the ribosome at the **E site**.



(b) Schematic model showing binding sites



(c) Schematic model with mRNA and tRNA

Fig. 17.15b &c

- Recent advances in our understanding of the structure of the ribosome strongly supports the hypothesis that rRNA, not protein, carries out the ribosome's functions.
  - RNA is the main constituent at the interphase between the two subunits and of the A and P sites.
  - It is the catalyst for peptide bond formation

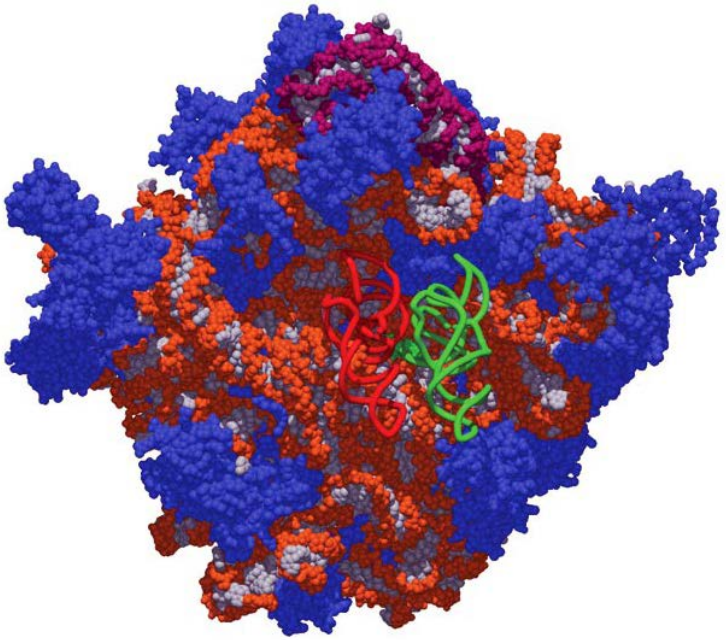


Fig. 17.16

- Translation can be divided into three stages:
  - initiation
  - elongation
  - termination
- All three phase require protein “factors” that aid in the translation process.
- Both initiation and chain elongation require energy provided by the hydrolysis of GTP.

- **Initiation** brings together mRNA, a tRNA with the first amino acid, and the two ribosomal subunits.
  - First, a small ribosomal subunit binds with mRNA and a special initiator tRNA, which carries methionine and attaches to the start codon.
  - *Initiation factors* bring in the large subunit to assemble the complete ribosome.

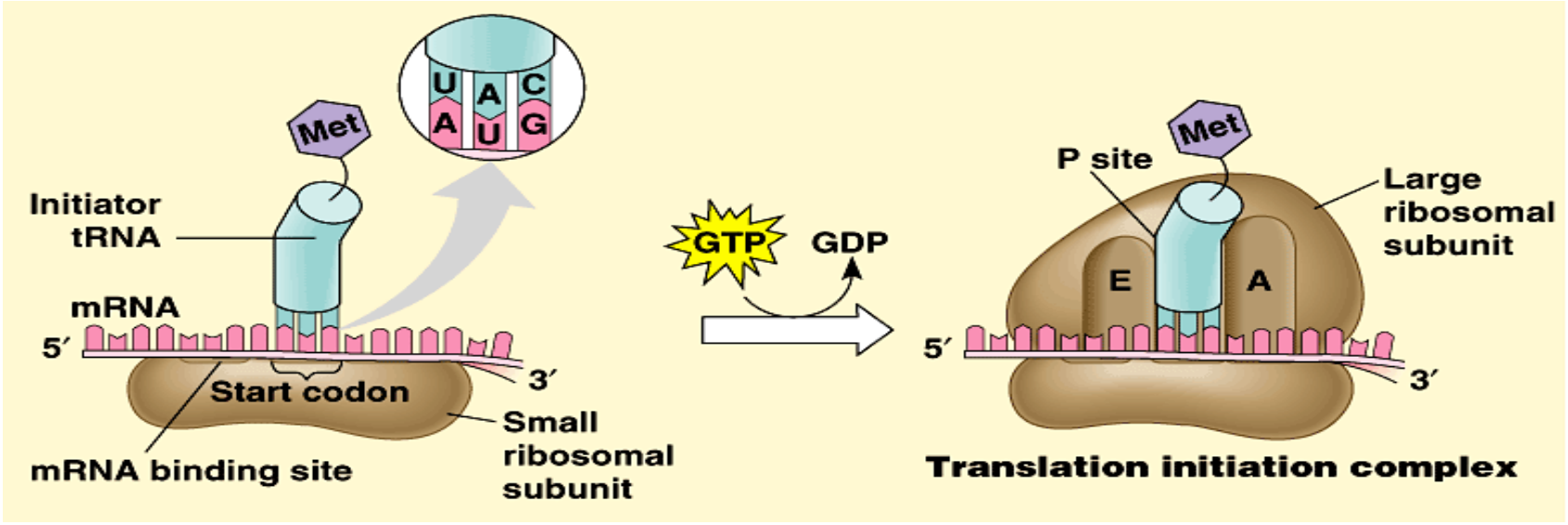


Fig. 17.17

- **Elongation** consists of a series of three step cycles as each amino acid is added to the proceeding one.
  - This step requires the hydrolysis of two GTP.

- The steps involved in elongation continue codon by codon to add amino acids until the polypeptide chain is completed.

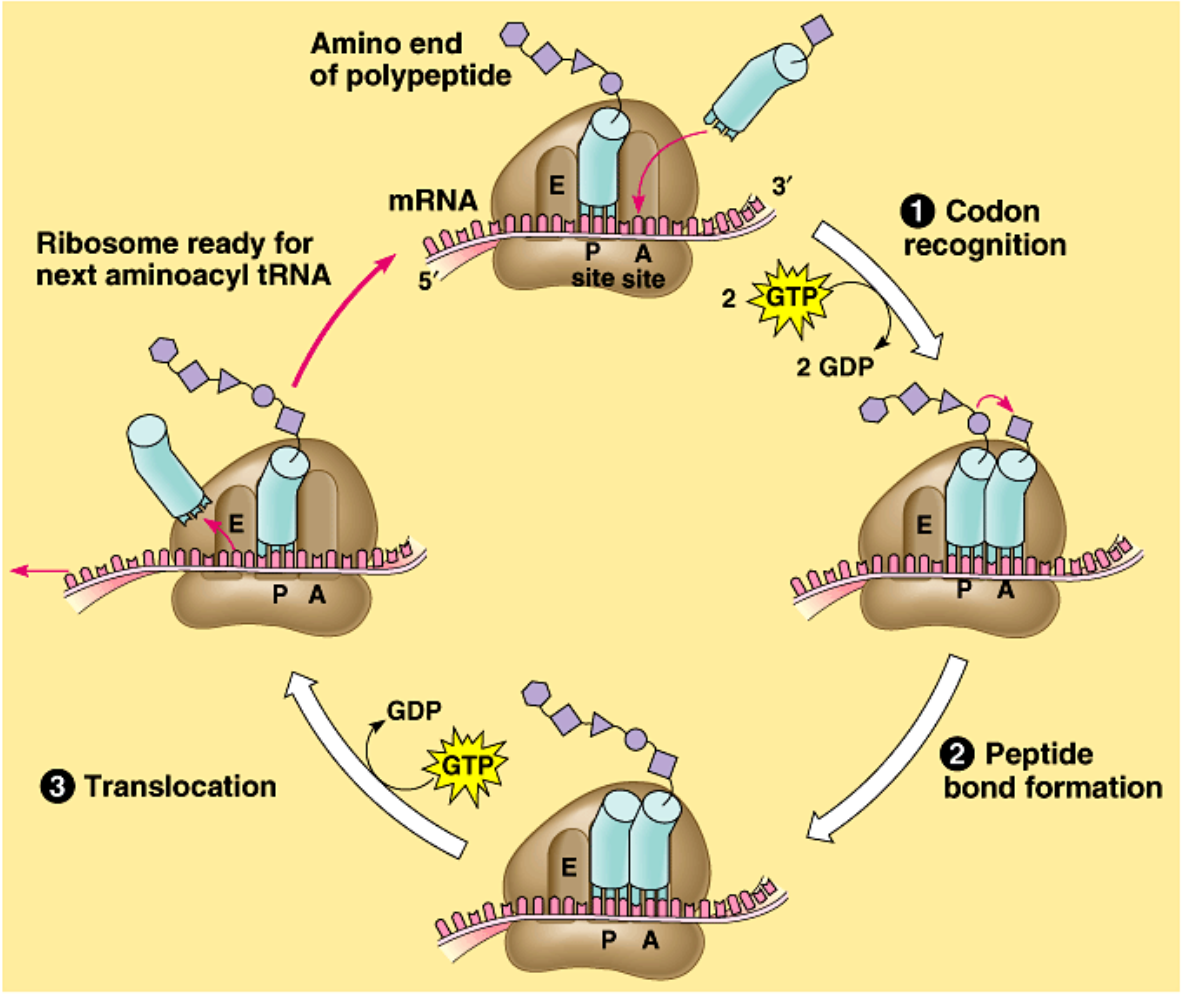


Fig. 17.18

- **Termination** occurs when one of the three stop codons reaches the A site.
- A *release factor* binds to the stop codon freeing the polypeptide and the translation complex disassembles.

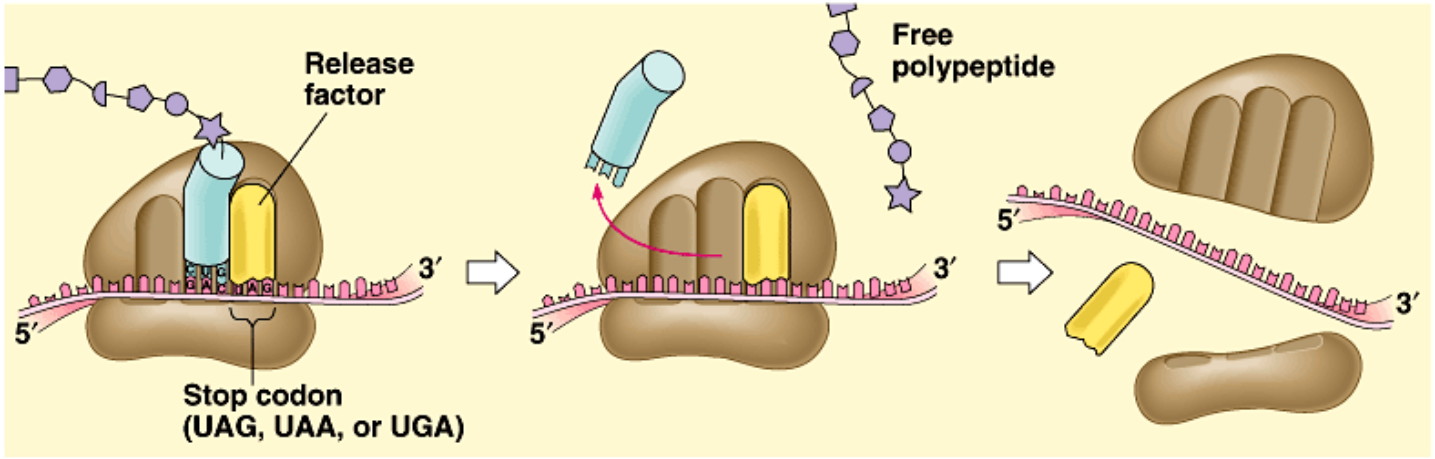
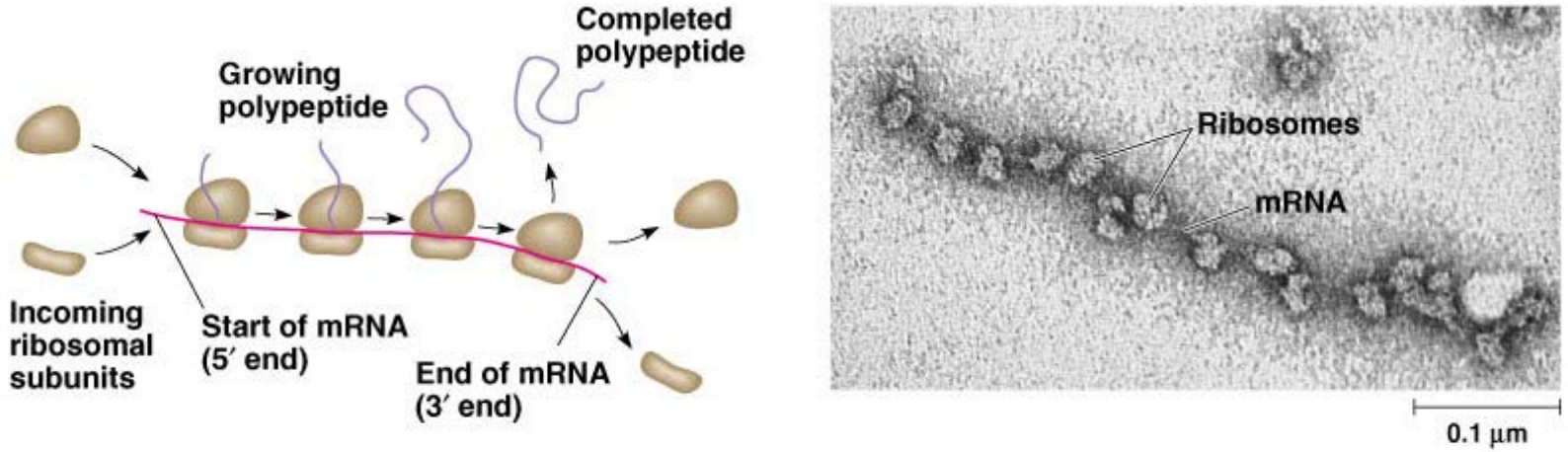


Fig. 17.19

- Typically a single mRNA is used to make many copies of a polypeptide simultaneously.
- Multiple ribosomes, **polyribosomes**, may trail along the same mRNA.
- A ribosome requires less than a minute to translate an average-sized mRNA into a polypeptide.



(a) An mRNA molecule is generally translated simultaneously by several ribosomes in clusters called polyribosomes.

(b) This micrograph shows a large polyribosome in a prokaryotic cell (TEM).

Fig. 17.20

- During and after synthesis, a polypeptide coils and folds to its three-dimensional shape spontaneously.
  - The primary structure, the order of amino acids, determines the secondary and tertiary structure.
- Chaperone proteins may aid correct folding.
- In addition, proteins may require *posttranslational modifications* before doing their particular job.
  - This may require additions like sugars, lipids, or phosphate groups to amino acids.
  - Enzymes may remove some amino acids or cleave whole polypeptide chains.
  - Two or more polypeptides may join to form a protein.

2. Signal peptides target some eukaryotic polypeptides to specific destinations in the cell

- Two populations of ribosomes, free and bound, are active participants in protein synthesis.
- Free ribosomes are suspended in the cytosol and synthesize proteins that reside in the cytosol.
- Bound ribosomes are attached to the cytosolic side of the endoplasmic reticulum.
  - They synthesize proteins of the endomembrane system as well as proteins secreted from the cell.

- While bound and free ribosomes are identical in structure, their location depends on the type of protein that they are synthesizing.
- Translation in all ribosomes begins in the cytosol, but a polypeptide destined for the endomembrane system or for export has a specific **signal peptide** region at or near the leading end.
- Secretory proteins are released entirely into the cisternal space, but membrane proteins remain partially embedded in the ER membrane.

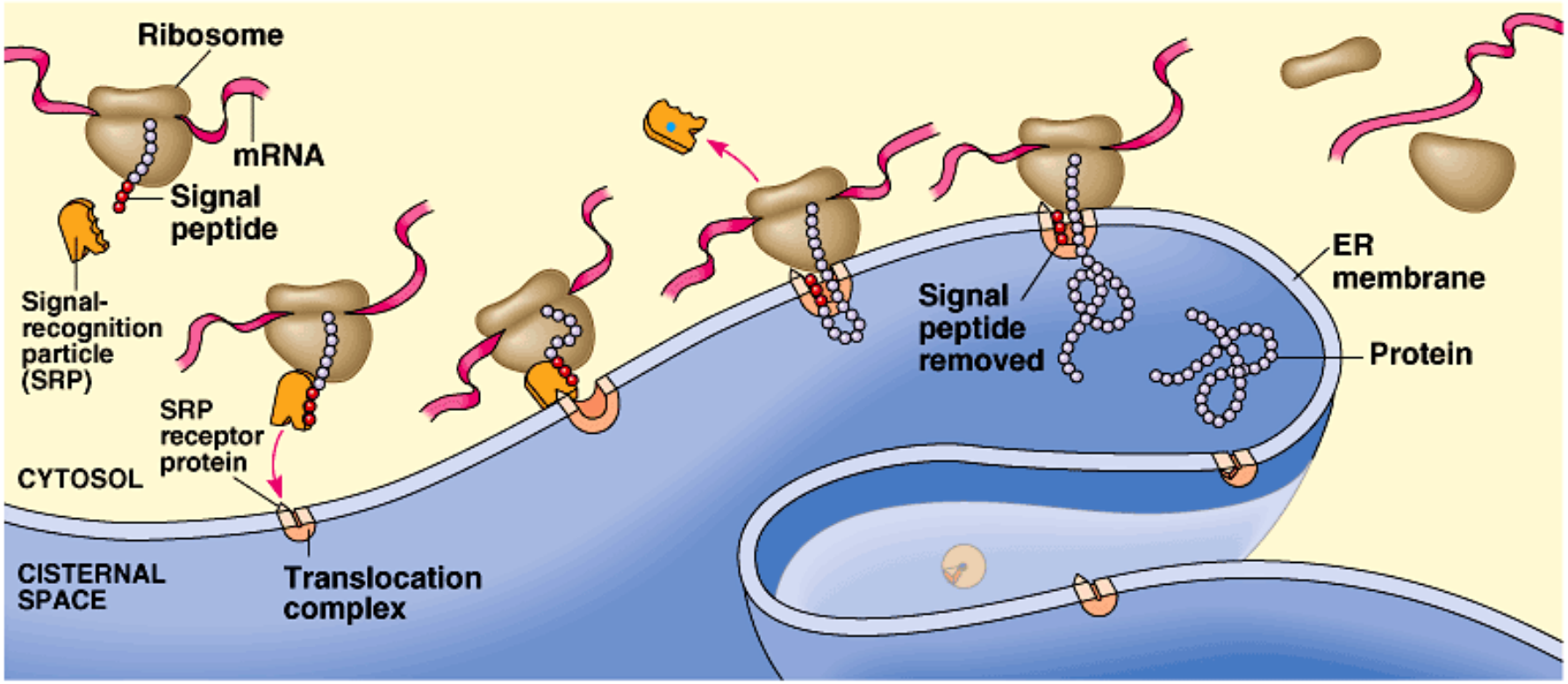


Fig. 17.21

- Other kinds of signal peptides are used to target polypeptides to mitochondria, chloroplasts, the nucleus, and other organelles that are not part of the endomembrane system.

- While the mechanisms of translocation vary, each of these polypeptides has a “postal” code that ensures its delivery to the correct cellular location.

3. RNA plays multiple roles in the cell: *a review*

- The cellular machinery of protein synthesis and ER targeting is dominated by various kinds of RNA.
  - The diverse functions of RNA are based, in part, on its ability to form hydrogen bonds with other nucleic acid molecules (DNA or RNA).
  - It can also assume a specific three-dimensional shape by forming hydrogen bonds between bases in different parts of its polynucleotide chain.
- DNA may be the genetic material of all living cells today, but RNA is much more versatile.

- The diverse functions of RNA range from structural to informational to catalytic.

Table 17.1 Types of RNA in a Eukaryotic Cell

Type of RNA	Functions
Messenger RNA (mRNA)	Carries information specifying amino acid sequences of proteins from DNA to ribosomes.
Transfer RNA (tRNA)	Plays catalytic (ribozyme) roles and structural roles in ribosomes.
Ribosomal RNA (rRNA)	Plays structural and catalytic (ribozyme) roles in ribosomes.
Primary transcript	Serves as a precursor to mRNA, rRNA, or tRNA and may be processed by splicing or cleavage . In eukaryotes, pre-mRNA commonly contains introns, noncoding segments that are spliced out as the primary transcript is processed. Some intron RNA acts as a ribozyme, catalyzing its own splicing.
Small nuclear RNA (snRNA)	Plays structural and catalytic roles in spliceosomes, the complexes of protein and RNA that splice pre-mRNA in the eukaryotic nucleus.
SRP RNA	Is a component of the signal-recognition particle (SRP), the protein-RNA complex that recognizes the signal peptides of polypeptides targeted to the ER.

4. Comparing protein synthesis in prokaryotes and eukaryotes: *a review*

- Although bacteria and eukaryotes carry out transcription and translation in very similar ways, they do have differences in cellular machinery and
  in details of the processes.
  - Eukaryotic RNA polymerases differ from those of prokaryotes and require transcription factors.
  - They differ in how transcription is terminated.
  - Their ribosomes are also different.

- In one big differences, prokaryotes can transcribe and translate the same gene simultaneously.
- The new protein quickly diffuses to its operating site.

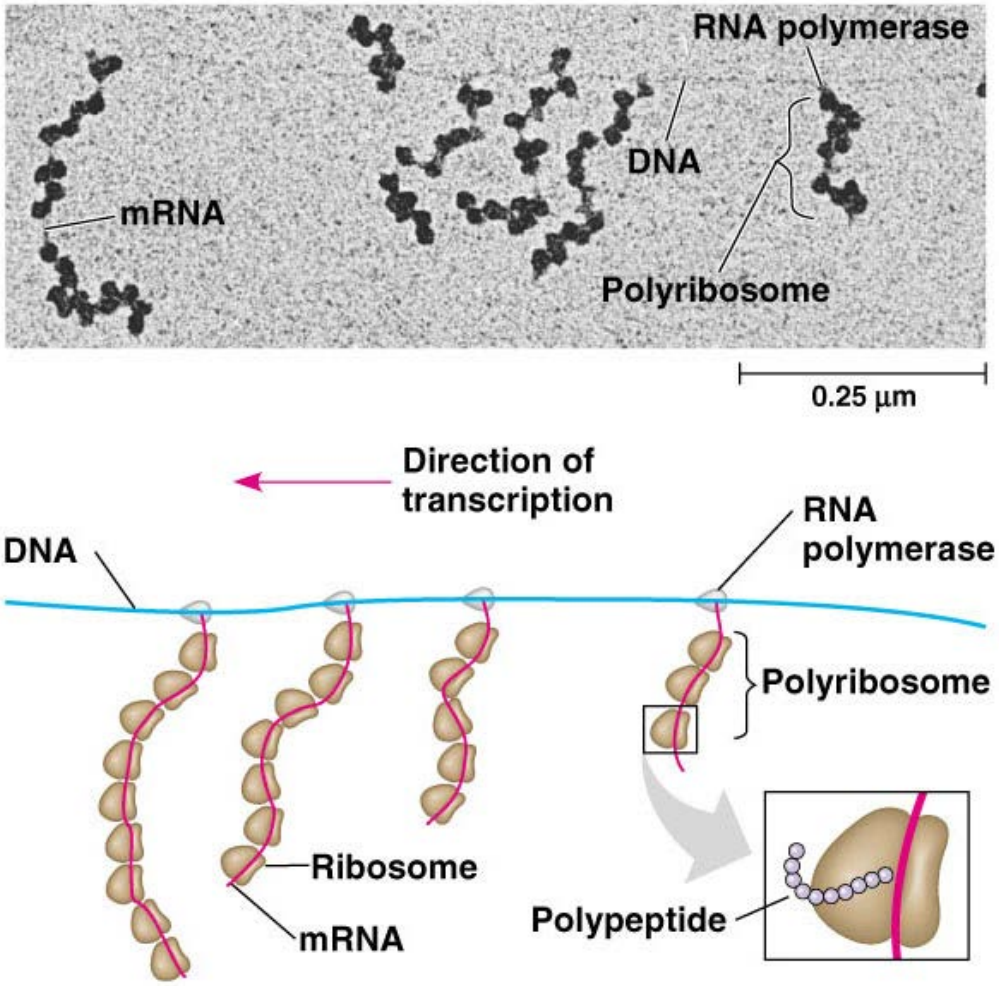


Fig. 17.22

- In eukaryotes, the nuclear envelope segregates transcription from translation.
- In addition, extensive RNA processing is inserted between these processes.
  - This provides additional steps whose regulation helps coordinate the elaborate activities of a eukaryotic cell.
- In addition, eukaryotic cells have complicated mechanisms for targeting proteins to the appropriate organelle.

5. Point mutations can affect protein structure and function

- **Mutations** are changes in the genetic material of a cell (or virus).
- These include large-scale mutations in which long segments of DNA are affected (for example, translocations, duplications, and inversions).
- A chemical change in just one base pair of a gene causes a **point mutation**.
- If these occur in gametes or cells producing gametes, they may be transmitted to future generations.

- For example, sickle-cell disease is caused by a mutation of a single base pair in the gene that codes for one of the polypeptides of hemoglobin.
  - A change in a single nucleotide from T to A in the DNA template leads to an abnormal protein.

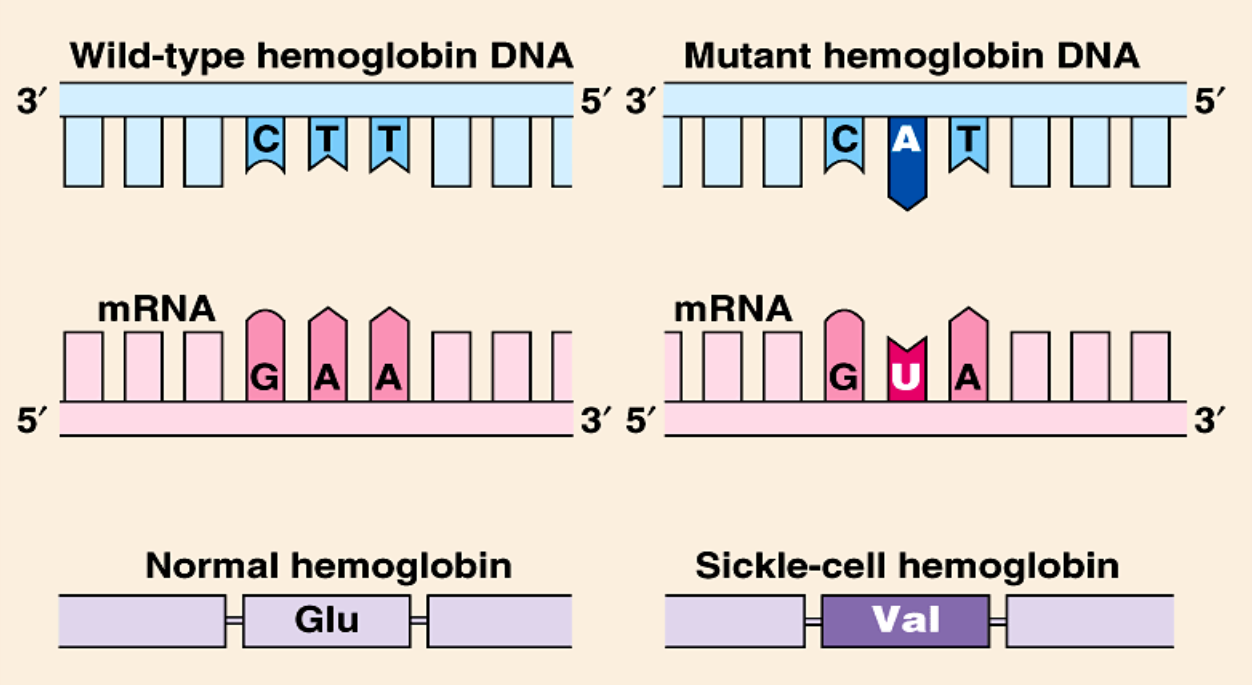


Fig. 17.23

- A point mutation that results in replacement of a pair of complimentary nucleotides with another nucleotide pair is called a **base-pair substitution**.
- Some base-pair substitutions have little or no impact on protein function.
  - In *silent mutations*, alterations of nucleotides still indicate the same amino acids because of redundancy in the genetic code.
  - Other changes lead to switches from one amino acid to another with similar properties.
  - Still other mutations may occur in a region where the exact amino acid sequence is not essential for function.

- Other base-pair substitutions cause a readily detectable change in a protein.
  - These are usually detrimental but can occasionally lead to an improved protein or one with novel capabilities.
  - Changes in amino acids at crucial sites, especially active sites, are likely to impact function.
- **Missense mutations** are those that still code for an amino acid but change the indicated amino acid.
- **Nonsense mutations** change an amino acid codon into a stop codon, nearly always leading to a nonfunctional protein.

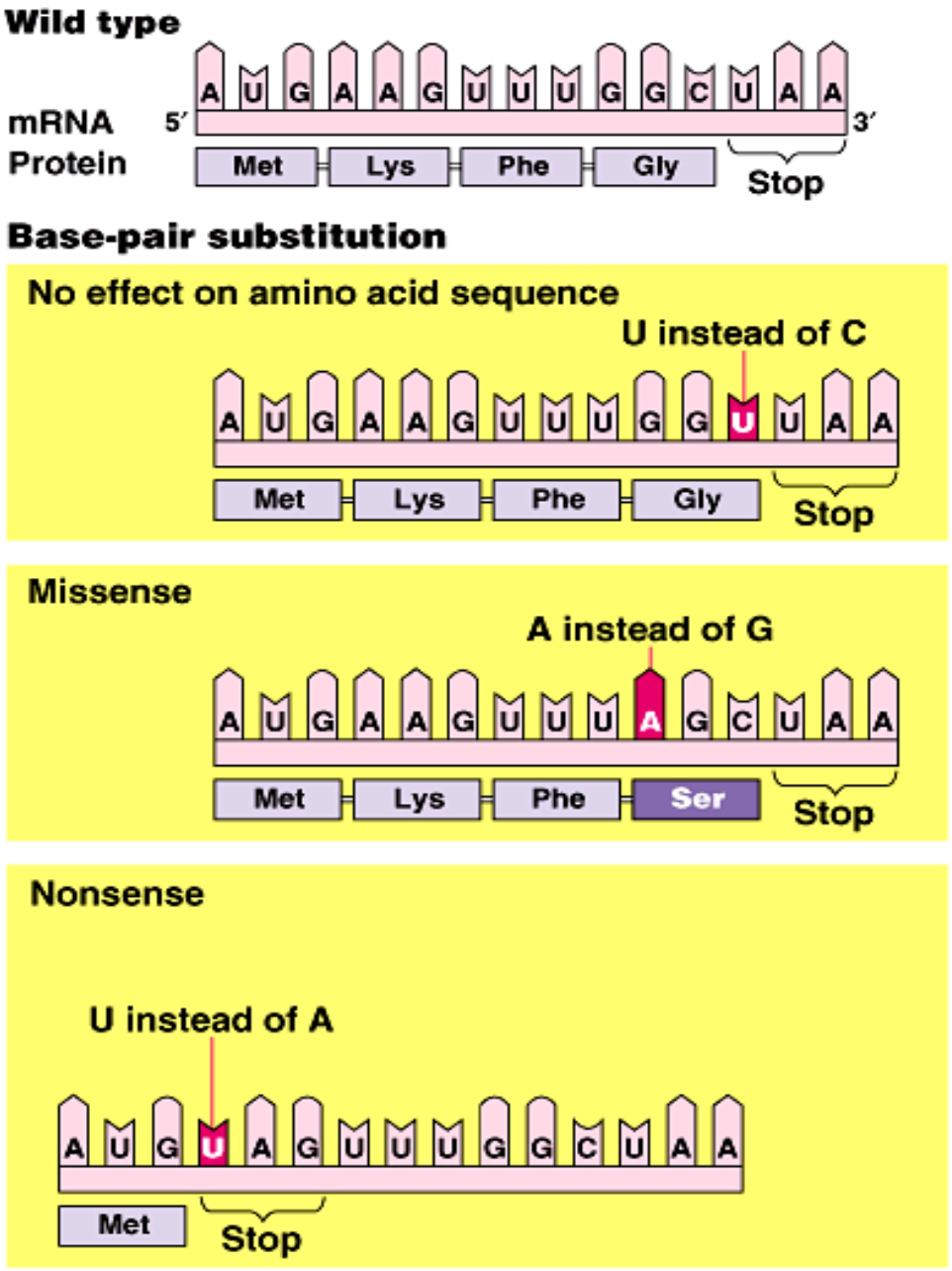


Fig. 17.24

- **Insertions** and **deletions** are additions or losses of nucleotide pairs in a gene.
  - These have a disastrous effect on the resulting protein more often than substitutions do.
- Unless these mutations occur in multiples of three, they cause a **frameshift mutation**.
  - All the nucleotides downstream of the deletion or insertion will be improperly grouped into codons.
  - The result will be extensive missense, ending sooner or later in nonsense - premature termination.

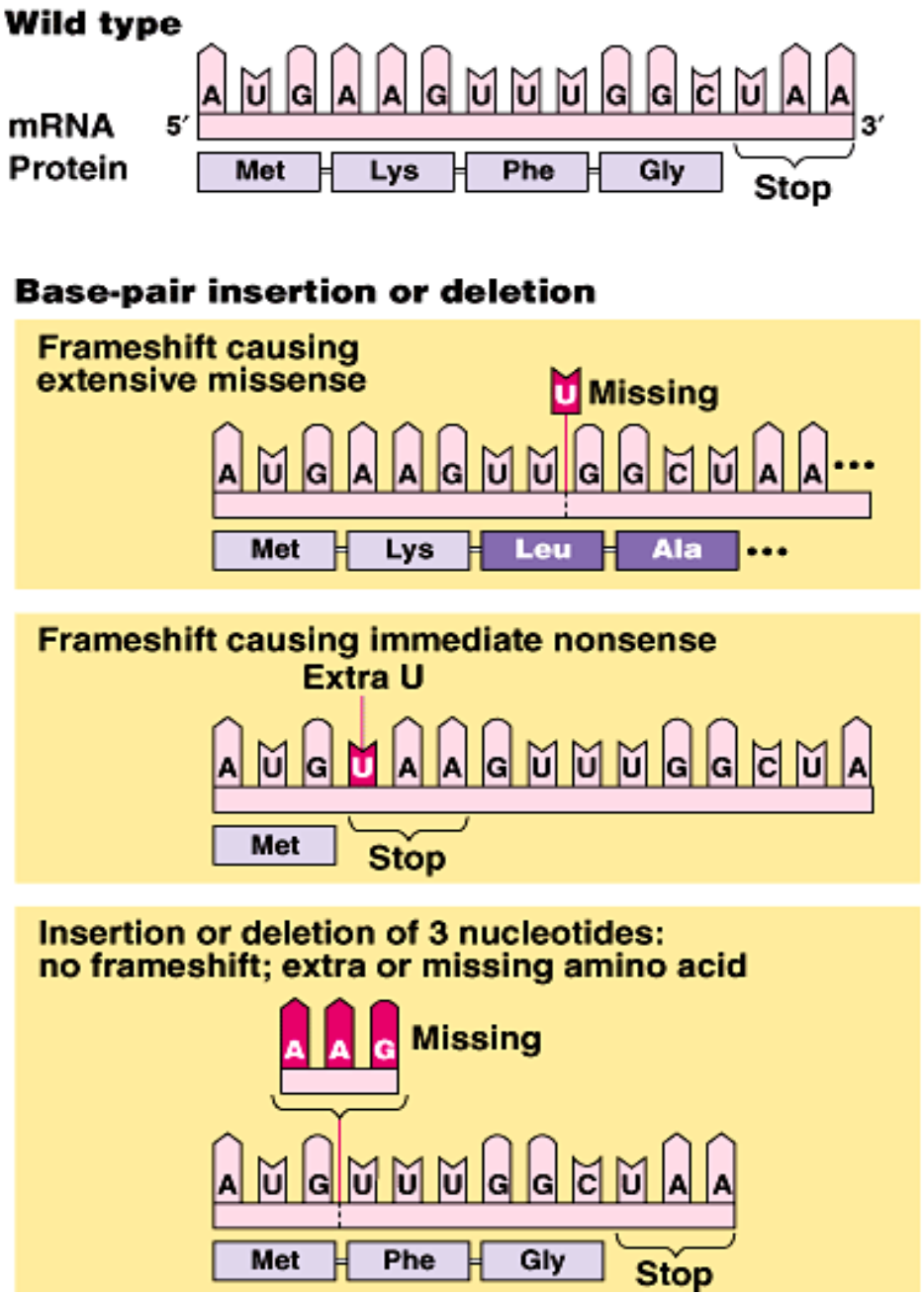


Fig. 17.24

- Mutations can occur in a number of ways.
  - Errors can occur during DNA replication, DNA repair, or DNA recombination.
  - These can lead to base-pair substitutions, insertions, or deletions, as well as mutations affecting longer stretches of DNA.
  - These are called *spontaneous mutations*.

- **Mutagens** are chemical or physical agents that interact with DNA to cause mutations.
- Physical agents include high-energy radiation like X-rays and ultraviolet light.
- Chemical mutagens may operate in several ways.
  - Some chemicals are base analogues that may be substituted into DNA, but that pair incorrectly during DNA replication.
  - Other mutagens interfere with DNA replication by inserting into DNA and distorting the double helix.
  - Still others cause chemical changes in bases that change their pairing properties.

- Researchers have developed various methods to test the mutagenic activity of different chemicals.
  - These tests are often used as a preliminary screen of chemicals to identify those that may cause cancer.
  - This make sense because most carcinogens are mutagenic and most mutagens are carcinogenic.

6. What is a gene? *revisiting the question*

- The Mendelian concept of a gene views it as a discrete unit of inheritance that affects phenotype.
- Morgan and his colleagues assigned genes to specific loci on chromosomes.
- We can also view a gene as a specific nucleotide sequence along a region of a DNA molecule.
- We can define a gene functionally as a DNA sequence that codes for a specific polypeptide chain.

- Transcription, RNA processing, and translation are the processes that link DNA sequences to the synthesis of a specific polypeptide chain.

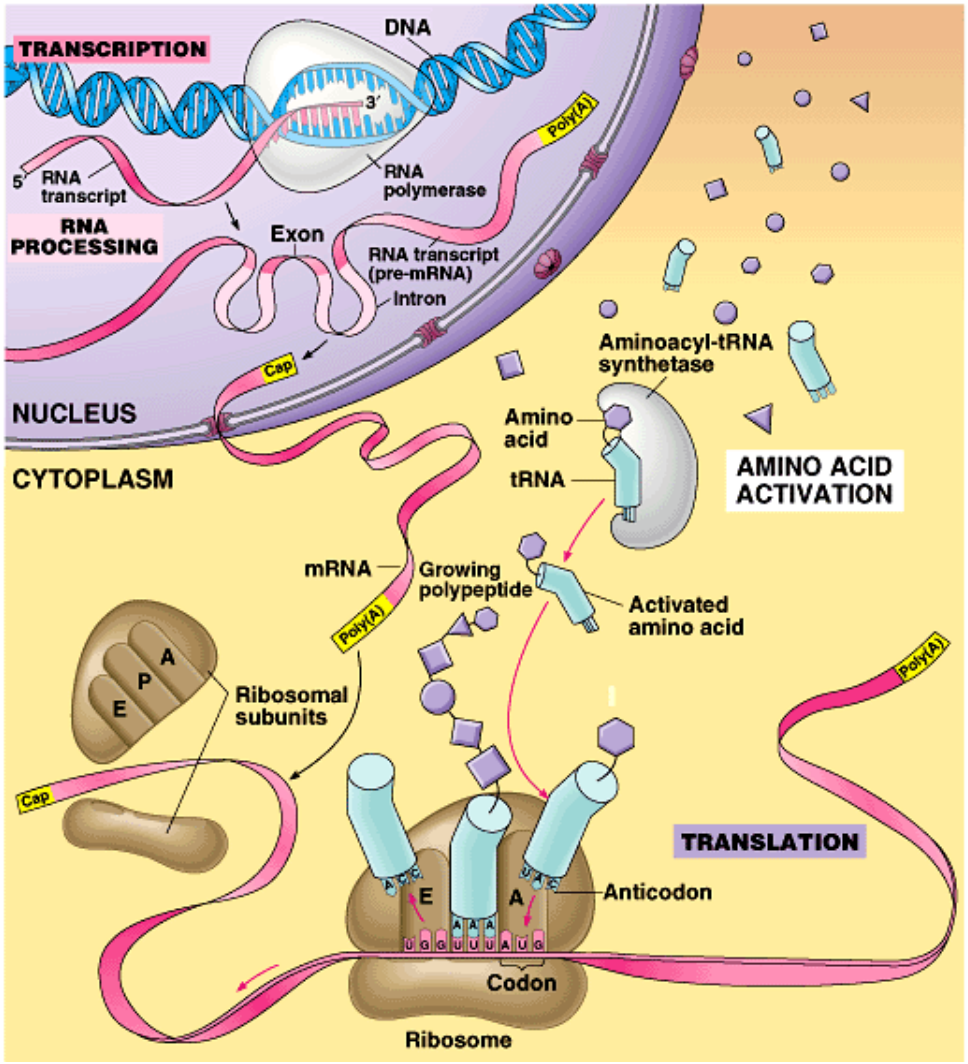


Fig. 17.25

- Even the one gene-one polypeptide definition must be refined and applied selectively.
  - Most eukaryotic genes contain large introns that have no corresponding segments in polypeptides.
  - Promotors and other regulatory regions of DNA are not transcribed either, but they must be present for transcription to occur.
  - Our definition must also include the various types of RNA that are not translated into polypeptides.
- *A gene is a region of DNA whose final product is either a polypeptide or an RNA molecule.*