Protein Synthesis

Inherited instructions in DNA direct protein synthesis. Thus, proteins are the links between genotype and phenotype, since proteins are directly involved in the expression of specific phenotypic traits.

Proteins are responsible for controlling many different **processes** in our bodies. As enzymes they break things down, put things back together, and catalyze chemical reactions. They make pigments, they form antigens and antibodies, give structure to cells and membranes, transport substances in cells and across membranes, and they perform hundreds of other functions.

A. EVIDENCE THAT GENES SPECIFY PROTEINS

Archibald Garrod was the first person to propose the relationship between genes and proteins (1909).

- He suggested that genes dictate phenotypes through enzymes that catalyze reactions.
- As a physician, Garrod was familiar with inherited disease which he called "inborn errors in metabolism." He hypothesized that such diseases reflected the patient's inability to make particular enzymes.

Garrod's hypothesis was confirmed decades later by research which determined that the function of a gene was to direct production of a specific enzyme.

- Biochemists accumulated evidence that cells synthesize and degrade organic compounds via metabolic pathways, with each step in the sequence catalyzed by a specific enzyme.
- Geneticists George **Beadle** and Edward **Tatum** were able to demonstrate the relationship between genes and enzymes by studying mutants of a bread mold, *Neurospora crassa.*
  - Wild-type *Neurospora* in laboratory colonies can survive on minimal medium. All other molecules needed by the mold are produced by its own metabolic pathways from this minimal nutrient source.
  - Beadle and Tatum searched for mutants that could not survive on minimal medium because they lacked the ability to synthesize essential molecules.
  - Mutants were identified by transferring fragments of growing fungi (in complete medium) to vials containing minimal medium. Fragments that didn't grow were identified as mutants.
  - Beadle and Tatum then identified specific metabolic defects (from mutations) by transferring fragment of mutants growing on complete growth medium to vials containing minimal medium each supplemented with only <u>one</u> additional nutrient.

- For example, if a mutant grew on minimal medium supplemented with only arginine, it could be concluded that the mutant was defective in the arginine synthesis pathway.

- Assuming that each mutant was defective in a single gene, they formulated the **one gene – one enzyme** hypothesis, which states that the function of a gene is to dictate the production of a specific enzyme.
- Beadle and Tatum's one gene – one enzyme hypothesis has been slightly modified:
  - While most enzymes are proteins, many proteins are not enzymes. However, proteins that are not enzymes are still gene products.
  - Also, many proteins are comprised of two or more polypeptide chains, each chain specified by a different gene.
- As a result of this new information, Beadle and Tatum's hypothesis has been restated as ***one gene – one polypeptide***.

B. AN OVERVIEW OF PROTEIN SYNTHESIS

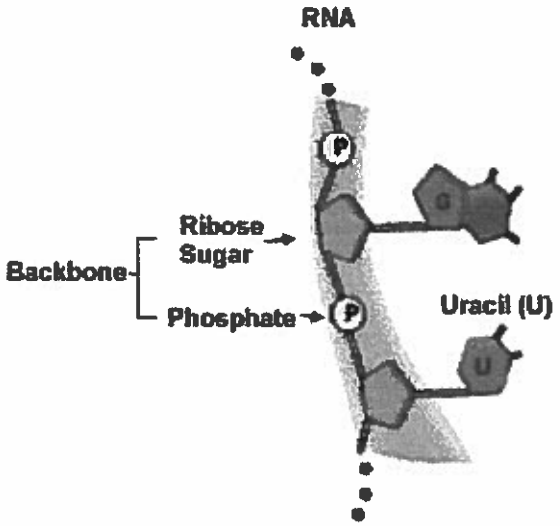
Ribonucleic acid (RNA) links DNA's genetic instruction for making proteins to the process of protein synthesis. It copies or transcribes the message from DNA and then translates that message into a protein.

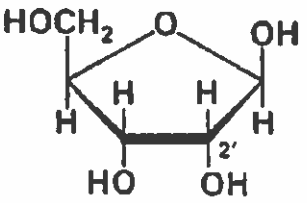
I. *Comparing RNA and DNA.*

⇨ RNA, like DNA, is a **nucleic acid** or polymer of nucleotides.

⇨ RNA structure differs from DNA in the following ways:

1. The five-carbon sugar in RNA nucleotides is ***ribose*** rather than deoxyribose.
2. RNA is **single** stranded.
3. The nitrogeneous base ***uracil*** is found in place of thymine. Therefore uracil on a RNA molecule would pair with adenine on a DNA molecule.





Ribose

II. How does information flow from gene to protein?

The linear sequence of nucleotides in DNA ultimately determines the linear sequence of amino acids in a protein.

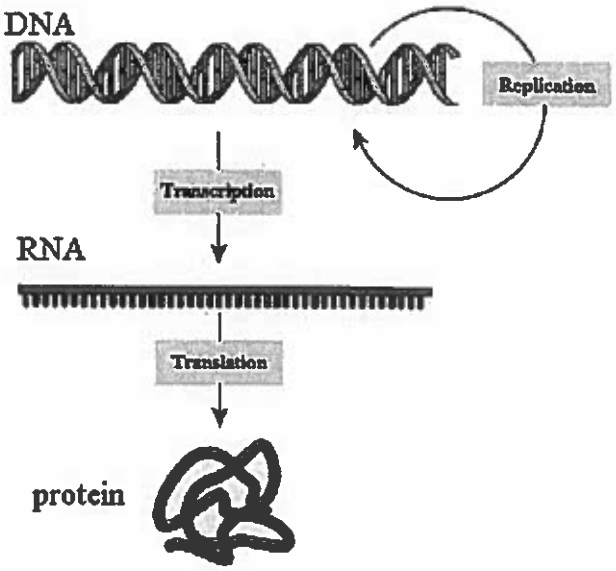
- ⇨ Nucleic acids are made of 4 types of nucleotides which differ in their nitrogenous bases. Hundreds or thousands of nucleotides, long, each gene has a specific linear sequence of the 4 possible bases.
- ⇨ Proteins are made of twenty types of **amino acids** (a.a) linked in a particular linear sequence (the protein's primary structure).
- ⇨ Information flows from gene to protein through two major processes, transcription and translation.

Transcription: The synthesis of RNA using DNA as a template.

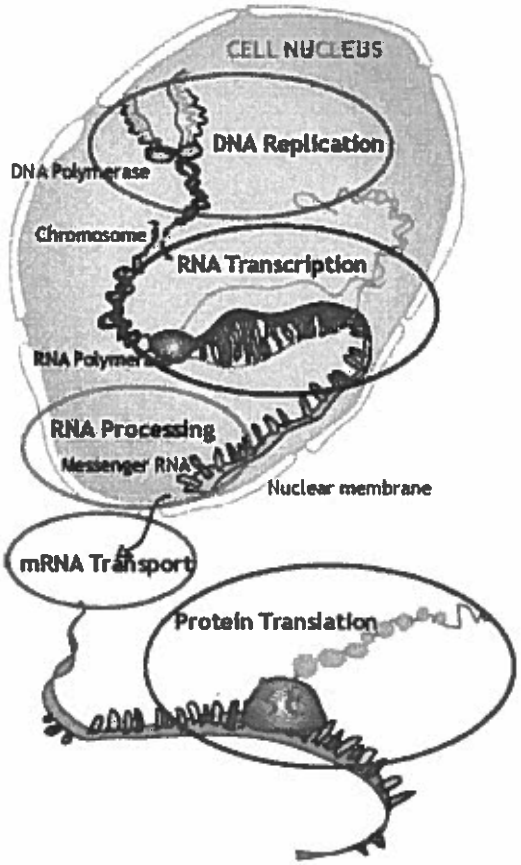
- A gene's unique nucleotide sequence is transcribed from DNA to a complementary nucleotide sequence in **messenger RNA** (mRNA).
- The resulting mRNA carries this transcript of protein-building instructions to the cell's protein–synthesizing machinery.

Translation: Synthesis of a polypeptide, which occurs under the direction of messenger RNA (mRNA).

- During this process, the linear sequence of bases in mRNA is translated into the linear sequence of amino acids in a polypeptide.
- Translation occurs on ribosomes which facilitate the orderly linking of amino acids into polypeptide chains.



Protein Synthesis: Helpful Background Information

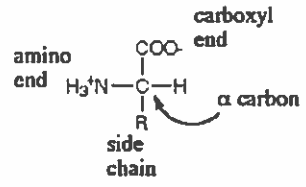


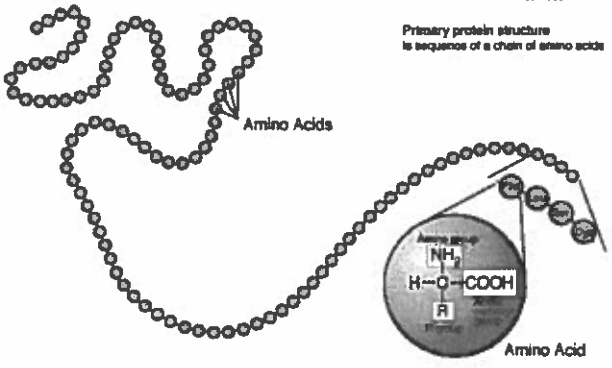
Ever wonder how a section of DNA instructs a cell how to make a protein? Actually, that's DNA's main purpose: to make proteins within the cell. Cells use the information encoded in their genes (which are kind of "protein libraries") as the "blueprint" for making proteins. Each gene in the DNA encodes information about how to make an individual protein. These proteins, which include enzymes, structural proteins, signaling proteins, transport proteins, receptor proteins, gene regulatory proteins, control the activities of the cell. Different cells have different activities. By controlling protein synthesis within each cell, the genes that make up DNA control the life of the entire organism.

Although the outcome of protein synthesis can be involved and quite complex, its purpose is rather straightforward. The purpose of protein synthesis is simply to create a polypeptide -- a protein made out of a chain of amino acids. Many ribosomes can be working on a single strand of mRNA at once. Protein synthesis isn't a slow process, either. A protein chain 400 amino acids long can be assembled in 20 seconds!

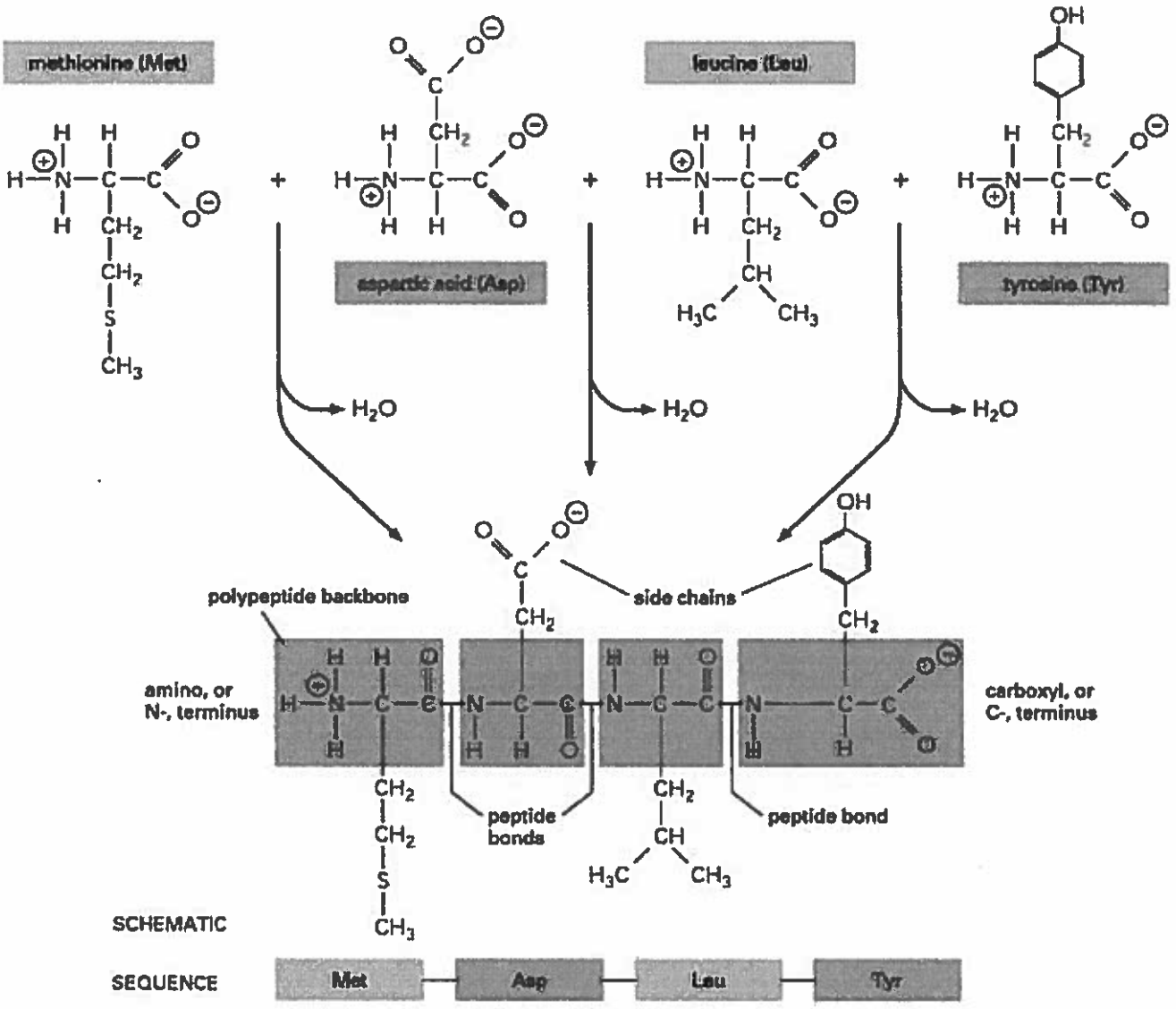
Proteins are made up of amino acids. So, amino acids are the building blocks of proteins. There are 20 common amino acids found in proteins. These molecules differ in their unique side chain (R group), which can be used to classify the molecules into functional types. All amino acids have the same general chemical formula. They contain a central or alpha carbon to which are bonded 4 groups:

- A hydrogen
- An amino group
- A carboxyl group
- A unique side chain (aka "R" group)





MAKING THE PROTEIN



Website Resources: Check these OUT!!

www.scilinks.org

Login as a **Guest, Keyword: HX209**

Topic is the Genetic Code – towards bottom of page lots of links to simulations and or diagrams of DNA/RNA/Proteins.

C. THE GENETIC CODE

There is not a one-to-one correspondence between the nitrogeneous bases and the amino acids they specify, since there are only 4 nucleotides and 20 amino acids. Researches have verified that the flow of information from a gene to a protein is based on a **triplet code**.

- Triplets of nucleotides are the smallest units of uniform length to allow translation into all 20 amino acids. These 3-nucleotide "words" are called **codons**. *(See text page 535 and Edge page 113)*

Codon: A 3-nucleotide sequence in mRNA that specifies which amino acid will be added to a growing polypeptide or that signals termination; the basic unit of the genetic code.

Genes are not directly translated into amino acids, but first are transcribed as codons into mRNA.

- For each gene, only one of the two DNA strands (the **coding** or **template** strand) is transcribed.

An mRNA is **complementary** to the DNA template from which it is **transcribed**.

- For example, if the triplet nucleotide sequence on the coding DNA strand is CCG; GGC, the codon for glycine, will be the complementary mRNA transcript.
- Remember, that according to the base-pairing rules, uracil (U) in RNA is used in place of thymine (T); uracil thus base pairs with adenine (A).

During **translation**, the linear sequence of codons along mRNA is translated into the linear sequence of amino acids in a polypeptide.

- Each mRNA codon specifies which one of 20 amino acids will be incorporated into the corresponding position in a polypeptide.

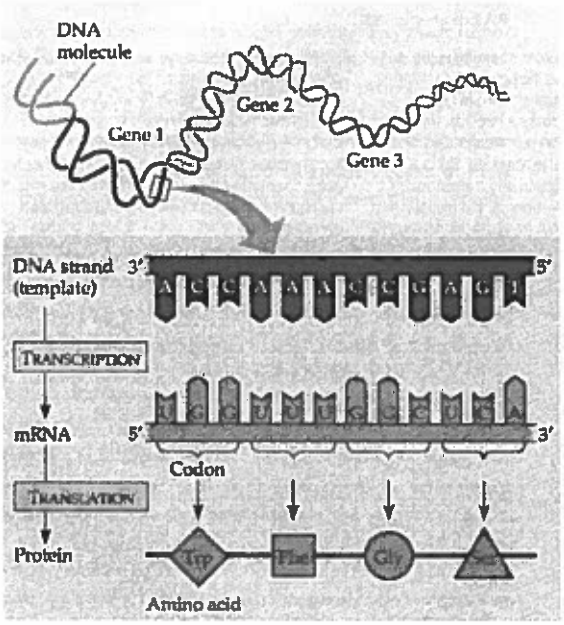


FIGURE 17.3 • The triplet code. For each gene, one of the two strands of DNA functions as a template for transcription—the synthesis of an mRNA molecule of complementary sequence. The same base-pairing rules that apply to DNA synthesis also guide transcription, but the base uracil (U) takes the place of thymine (T) in RNA. During translation, the genetic message (mRNA) is read as a sequence of base triplets, analogous to three-letter code words. Each of these triplets, called a codon (bracketed in the figure), specifies the amino acid to be added at the corresponding position along a growing polypeptide chain. The gene, its mRNA transcript, and the polypeptide product are all much longer than the segments shown here.

Campbell, 5th Edition, pg. 298.

Characteristics of the Genetic Code

1. All 64 possible codons have a specific meaning.
   - Three of the 64 codons act the same way that a period does at the end of a sentence. **UAA, UAG** and **UGA** on mRNA cause the assembly process to stop and the newly formed polypeptide to be released (referred to as **terminator codons**).
   - The remaining 61 codons each designate a specific amino acid. Obviously, several codons must code for the same amino acid but each codon only corresponds to one amino acid.
2. Codons which correspond to the same amino acid are very similar, usually sharing the same first two nucleotides.
3. The codon **AUG** has a dual function. It codes for an amino acid, but also signals where a translation sequence should start (**initiator codon**).

D. A CLOSER LOOK AT TRANSCRIPTION

The Making of Messenger RNA

- The copy of the genetic blueprint made during **transcription** is called **mRNA** (messenger ribonucleic acid).
- The mRNA carries the information from the nucleus to the **ribosomes**, where it directs the manufacture of a polypeptide (a long chain of amino acids).
- **ANALOGY:** Consider a reference library where books can be used but not borrowed. These books could represent the DNA. When specific information is needed, photocopies can be made and taken out. In the cell, the equivalent of the photocopying process makes the messenger RNA as a copy of the DNA. Just as one normally photocopies only a few pages of a book, the mRNA copy represents only a small segment of the genetic information contained in the DNA.
- Transcription of **messenger RNA** (mRNA) from the coding DNA strand is catalyzed by enzymes called **RNA polymerases**.
  - RNA polymerases **separate** the two DNA strands and **link RNA nucleotides** as they base pair along the DNA template.
  - Specific DNA nucleotide sequences mark where transcription of a gene begins (***initiation***) and ends (***termination***). Initiation and termination sequences plus the nucleotides in between are called a **transcription unit.**
- Transcription occurs in **three key steps:**
  1. *polymerase binding and initiation*
  2. *elongation*
  3. *termination*

RNA Polymerase Binding and Initiation of Transcription

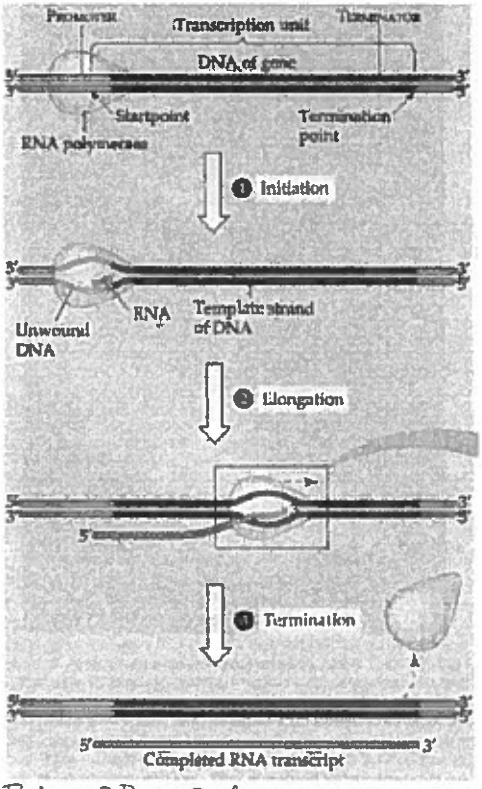
- RNA polymerases bind to DNA at regions called **promoters** (area where ***initiator codons*** are located).
- When active RNA polymerase binds to a promoter, the enzyme separates the two DNA strands at the initiation site, and transcription begins.

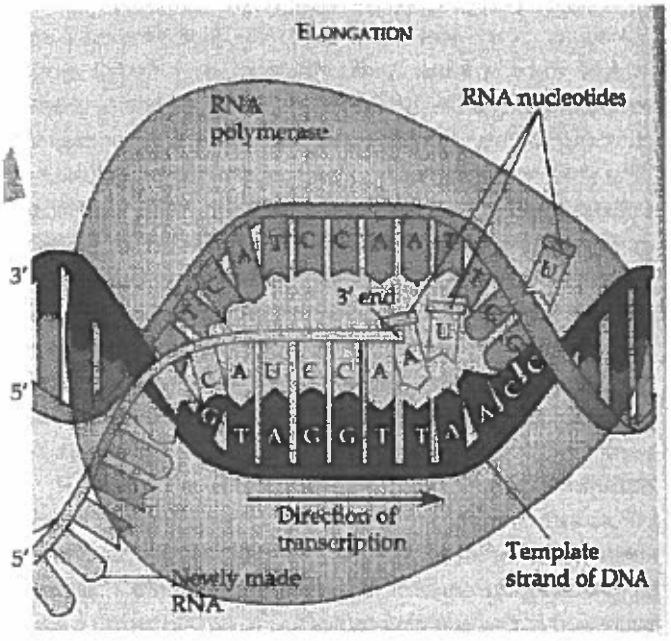
Elongation of the RNA Strand

- Once transcription begins, RNA polymerase moves along DNA and performs two primary functions:
    - It untwists and opens a short segment of DNA exposing about ten nucleotide bases; one of the exposed DNA strands is the **template** for base pairing with RNA nucleotides.
    - It links incoming RNA nucleotides to the elongating strand.
- During transcription, mRNA grows about 30-60 nucleotides per second. As the mRNA elongates:
    - It peels away from its DNA template.
    - The non-coding strand of DNA re-forms a DNA-DNA double helix by pairing with the coding strand of DNA.

Termination of Transcription

- Transcription proceeds until RNA polymerase reaches a **termination codon** on the DNA.
- **Terminator sequence:** DNA sequence (codon) that signals RNA polymerase to stop transcription and to release the RNA molecule and DNA template.





Campbell, 5th Edition, pg. 300.

E. A CLOSER LOOK AT TRANSLATION

During translation, **proteins** are synthesized according to a genetic message in the form of sequential codons along mRNA.

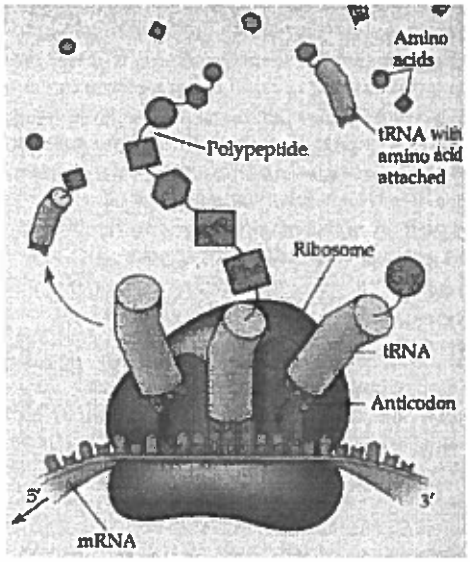


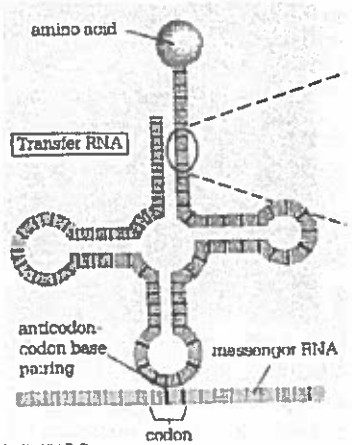
FIGURE 17.11 • Translation: the basic concept. As a molecule of mRNA slides through a ribosome, codons are translated into amino acids, one by one. The interpreters are tRNA molecules, each type with a specific anticodon at one end and a certain amino acid at the other end. A tRNA adds its amino acid cargo to a growing polypeptide chain when the anticodon bonds to a complementary codon on the mRNA. The figures that follow show some of the details of translation in the prokaryotic cell.

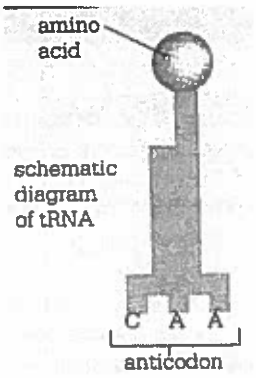
Campbell, 5th Edition, pg. 304.

Transfer RNA

Transfer **RNA** (tRNA) is the **interpreter** between the two forms of information – base sequences in mRNA and amino acid sequence in polypeptides.

- tRNA aligns the appropriate amino acids to form a new polypeptide. To perform this function, tRNA must:
  - Transfer amino acids from the cytoplasm's amino acid pool to a **ribosome**.
  - Recognize the correct codons in mRNA.
- Molecules of tRNA are specific for only one particular amino acid. Each type of tRNA associates a distinct mRNA codon with one of the 20 amino acids used to make proteins.
  - One end of a tRNA molecule attaches to a specific amino acid.
  - The other end attaches to an mRNA codon by base pairing with its **anticodon** (a nucleotide triplet in tRNA that base pairs with a complementary nucleotide triplet (codon) in mRNA).





Galbraith, Biology Directions, pg. 531.

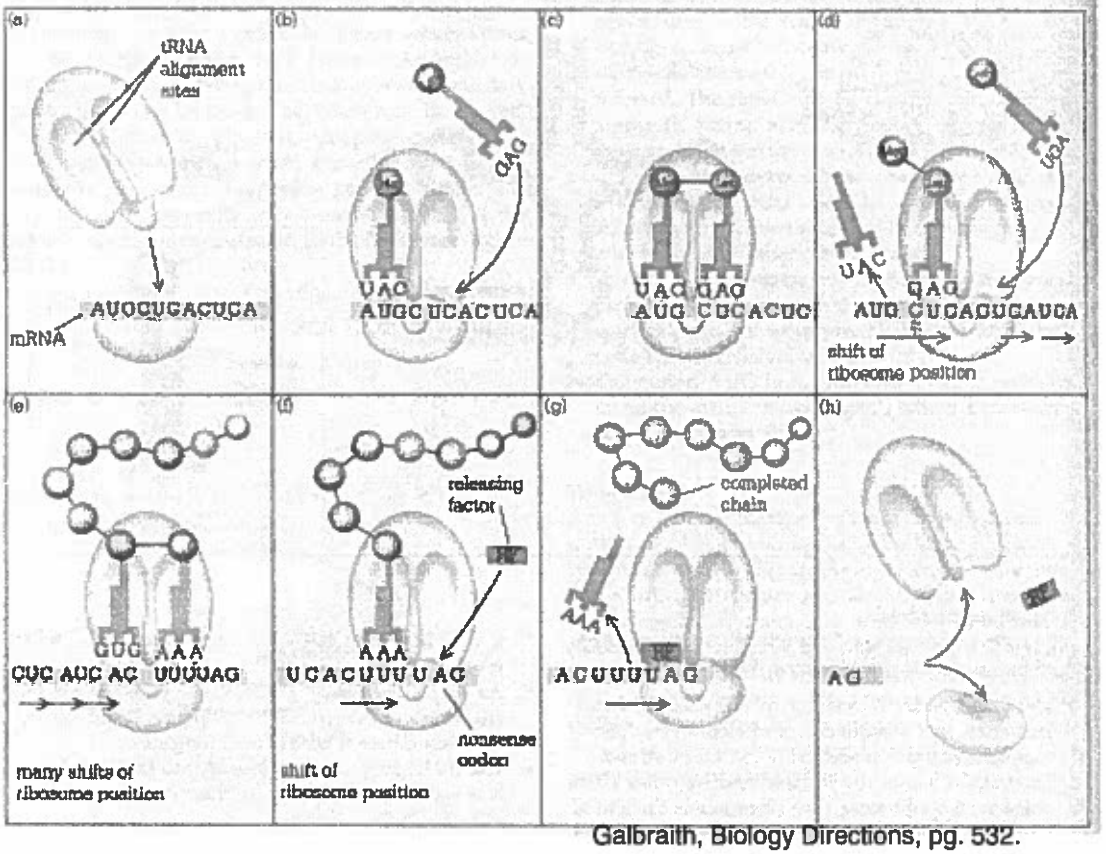
- tRNAs decode the genetic message, codon by codon. For example,
  - The mRNA codon UUU is translated as the amino acid phenylalanine.
  - The tRNA that transfers phenylalanine to the ribosome has an anticodon of AAA.
  - When the codon UUU is presented for translation, phenylalanine will be added to the growing polypeptide.
  - As tRNAs deposit amino acids in the correct order, ribosomal enzymes link then into a chain.

POLYPEPTIDE SYNTHESIS:
THE STEPS

The process of polypeptide (protein) synthesis can be summarized in a series of general steps.

1. **mRNA** is formed in the nucleus (**transcription**). It is a sequence of nucleotides which is complementary to a section of the DNA strand.
2. The mRNA leaves the nucleus and becomes associated with a **ribosome**. The ribosome is the site of polypeptide synthesis.
3. In the cytoplasm there is a pool of amino acids. Each kind of amino acid is linked to one specific **tRNA** molecule.
4. Each tRNA molecule carries its amino acid to the ribosome. The anticodon of three specific nucleotides on the tRNA determines where on the mRNA it will fit. If the mRNA codon reads AAG, what must the anticodon read?
5. As the amino acids are brought together on the ribosome in sequence, a peptide bond forms between adjacent amino acid molecules.
6. The tRNA molecule is released into the cytoplasm. It is now free to pick up another of the same kind of amino acid.
7. When the mRNA codes reads "stop" (**terminator codons** – UAG, UAA or UGA), the polypeptide is released. The mRNA is also released from the ribosome.

Translation occurs on ribosomes in the cytoplasm of the cell. In (a), the two pieces of ribosome come together around a "start" codon on the mRNA. As the ribosome moves along the mRNA, the tRNAs are sequentially brought in according to the mRNA codons, and the amino acids are bonded together to form the specific protein. When the protein is complete, a "nonsense" or "stop" codon releases the mRNA from the ribosome.



Galbraith, Biology Directions, pg. 532.

F. MUTATIONS AND THEIR EFFECTS ON PROTEINS

Mutations are changes in the genetic material of a cell. Previously, we considered large-scale mutations, chromosomal rearrangements that affect long segments of DNA. Now, that we have discussed the genetic code and its translation we will look at **point mutations**, which are chemical changes in just one or a few base pairs in a single gene.

> **Point Mutation** – a mutation limited to about one or two nucleotides in a single gene.

If a point mutation occurs in a gamete, or in a cell that gives rise to gametes, it may be transmitted to offspring and to a succession of future generations.

- If the mutation has an adverse effect on the phenotype of a human or other animal, the mutant condition is referred to as a **genetic disorder**, or **hereditary disease**.

Types of Mutations

There are two categories of point mutations:

1. Base-pair Substitutions
   - The replacement of one base pair with another
   - Occurs when a nucleotide and its partner form the complementary DNA strand are replaced with another pair of nucleotides according to base-pairing rules.
   - Depending on how base-pair substitutions are translated, they can result in little or no change in the protein encoded by mutated gene.
     - Redundancy in the genetic code is why some substitution mutations have no effect. A base pair change may simply transform one codon into another that codes for the same amino acid.
     - Even if the substitution alters an amino acid, the new amino acid may have similar properties to the one it replaces, or it may be in a part of the protein where the exact amino acid sequence is not essential to its activity.
   - Some base-pair substitutions result in readily detectable changes in proteins.
     - Alteration of a single amino acid in a crucial area of a protein will significantly alter protein activity.
     - On rare occasions, such a mutation will produce a protein that is improved or has capabilities that enhance success of the mutant organism and its descendants.
     - More often, such mutations produce a less active or inactive protein that impairs cell function.

       ⇨ Example: **Sickle-cell anemia.**

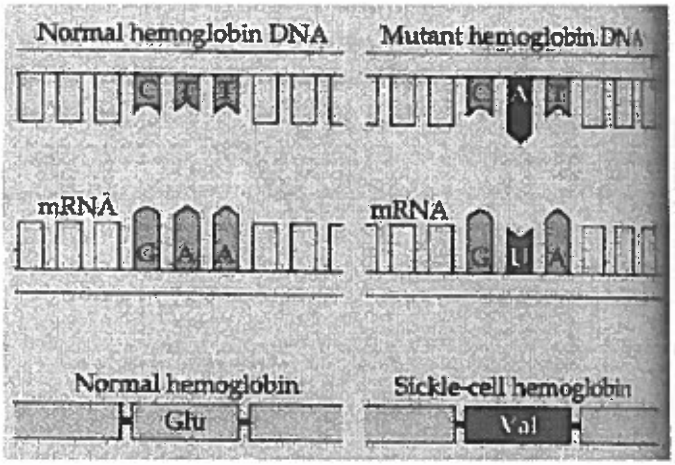


FIGURE 17.21 - The molecular basis of sickle-cell disease. The allele that causes sickle-cell disease differs from the normal allele by a change in a single base pair—a point mutation. The top row of this figure shows the template strand of the gene for one of the polypeptides making up the protein hemoglobin. Where the normal allele has the base thymine, the sickle-cell allele has adenine. This alters one of the codons in the mRNA transcribed from the gene, with the result that the amino acid valine appears in the polypeptide in place of the glutamic acid found in the normal polypeptide. In individuals who are homozygous for the mutant allele, the sickling of red blood cells caused by the altered hemoglobin produces the multiple symptoms associated with sickle-cell disease (see FIGURE 14.16).

Campbell, 5th Edition, pg. 312.

II. Base-pair Insertions or Deletions

- Usually have a greater negative effect on proteins than substitutions.
- **Base-pair insertion:** The insertion of one or more nucleotide pairs into a gene.
- **Base-pair deletion:** The deletion of one or more nucleotide pairs from a gene.
- Because mRNA is read as a series of triplets during translation, insertion or deletion of nucleotides may alter the reading frame (triplet grouping – codon) of the genetic message.
- This type of **frameshift mutation** will occur whenever the number of nucleotides inserted or deleted is not 3 or a multiple of 3.
- Frameshift will produce a nonfunctional protein unless the insertion or deletion is very near the end of the gene.

Mutagenesis

MUTAGENESIS: the creation of mutations.

- Mutations can occur as errors in DNA replication, repair or recombination that result in base-pair substitutions, insertions or deletions.
- Mutagenesis may be a naturally occurring event or may be caused by exposure to mutagens.

MUTAGEN: physical or chemical agents that interact with DNA to cause mutations.

- **Radiation** is the most common physical mutagen in nature and has been used in the laboratory to induce mutations.
- Several categories of **chemical** mutagens are known including *base analogues*, which are chemicals that mimic normal DNA bases, but base pair incorrectly.
- The **Ames test**, developed by Bruce Ames, is one of the most widely used tests for measuring the mutagenic strength of various chemicals. Since most mutagens are **carcinogenic**, this test is also used to screen for chemical carcinogens.

ONCOGENES: cancer causing genes.

REMEMBER – NOT ALL MUTATIONS ARE HARMFUL!

Can you think of any STS issues related to mutagens?

Making the Chromosome- Gene-Protein Connection

The Story of Sickle-Cell Anemia

Recent research in biology has been connecting known "disease genes" mapped within chromosomes to the altered proteins these genes code in the body. Hemoglobin is a prime example; it is the protein that carries oxygen and carbon dioxide in the red blood cell. The hemoglobin protein is made of four polypeptide chains: 2 ∝ chains (141 amino acids long) and 2 β chains (146 amino acids long), along with four iron-containing heme groups. In the genetic disease sickle-cell anemia, there is a mutation in the gene that codes the beta chain of hemoglobin. Within this gene (located on chromosome #11) one base in the DNA is replaced with another base and this mutation causes the normal amino acid #6 to be replaced by another amino acid.

Making a Normal Beta Chain of Hemoglobin

Here is the first part of the DNA sequence (McKusick, 1994) for the beta chain of normal hemoglobin. Fill in the complementary DNA strand, using the base-pairing rules for making DNA.

GTG CAC CTG ACT CCT GAG GAG

Now make the messenger RNA from this complementary strand of DNA, using the base-pairing rules for making RNA:

Now using the genetic code from your formula sheet, translate this messenger RNA into a sequence of amino acids:

Making Sickle Cell Hemoglobin

In sickle-cell anemia, there is a mutation at the 17th nucleotide of DNA in this gene; the nucleotide is changed from A to T so it reads.

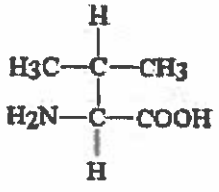
GTG CAC CTG ACT CCT GTG GAG

Fill in the complementary DNA strand:

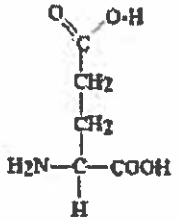
Now make the messenger RNA from this complementary mutated strand:

Now, using the genetic code, translate this new messenger RNA into a sequence of amino acids:

You can see that in normal hemoglobin, amino acid #6 is Glu and in sickle-cell hemoglobin amino acid #6 is Val. Observe the structural formulae for these two amino acids:



valine



glutamic acid

Although the altered hemoglobin has only one amino acid changed out of the total of 146, it's a crucial amino acid. When this new amino acid is at position #6 instead of the correct amino acid, the hemoglobin beta chain becomes more **hydrophobic**. As a result, when the hemoglobin chains fold into their three-dimensional shape and assemble together (2 normal alpha chains with 2 sickle-cell beta chains) into the quaternary hemoglobin molecule, the unit molecules tend to stick to each other, forming long insoluble fibers of hemoglobin. This altered hemoglobin deforms the normally rounded red blood cell into a distorted sickle shape (curved with pointed ends). The deformed cell is abnormally fragile, rupturing easily in tiny capillaries and clogging the vessels. As the red cells break down at an increased rate, the body experiences anemia. The anemic person becomes weak, dizzy, and short of breath during physical exertion. Also, as capillaries all over the body become blocked by clumped sickle cells, body organs are damaged, leading to pain and often premature death.

Genetics of Sickle-cell Anemia

Geneticists represent the single gene that codes for the beta chain of hemoglobin by using the letter H for the normal hemoglobin allele, and h for sickle-cell hemoglobin allele. Does this capital/lowercase letter symbolism convey that sickle-cell anemia (h) is an autosomal dominant or recessive disease?

_____.

Check Your Understanding

An adult with homozygous normal hemoglobin would have the genotype: _____.

An adult with sickle-cell anemia would have the genotype: _____.

An adult with normal hemoglobin, but heterozygous for the sickle-cell trait would have the genotype:

_____.

Genetic Crosses

Do a genetic cross, using the Punnett square, of a sickle cell individual with a person who is homozygous normal. List the genotypes and phenotypes of offspring. Next, do a genetic cross, using the Punnett square, of two normal heterozygotes (carriers) of the sickle-cell trait. List the genotypes and phenotypes of offspring.

The Heterozygote Advantage

One note on the subtlety of genetic disease: In parts of Africa where malaria is very prevalent and claims many lives, as much as 20% of the population may be carriers for the sickle cell gene. Although being homozygous for sickle-cell anemia leads to early death and lowered likelihood of reproduction to pass on the gene, the sickle cell heterozygotes (carriers) in high-malaria regions have improved survivability over the homozygous normal individuals. Why? It turns out that red blood cells of individuals heterozygous for sickle cell are less easily infected by the malaria parasite, thus improving the heterozygote's survival and ability to reproduce in that malaria infested environment. Hence the occurrence of the heterozygote is favored over the homozygous normal by selection pressures from the malaria parasite.

Genetic Testing and Counselling

It is now possible to test people for the presence of the sickle-cell gene. Testing can determine whether a person is homozygous normal or heterozygous normal (a carrier) for sickle-cell anemia. Sickle-cell anemia is known to be frequent among some African (and African-American) populations. Approximately, one in 10 African-Americans are carriers for the gene. But, the gene has also been found among those of Italian, Greek, Arabian, Indian and Turkish ancestry, with a carrier state prevalence ranging form 0 to 38% depending on the group sampled.