

## 16.5: Protein Synthesis and the Genetic Code

### Learning Objectives

- describe the characteristics of the genetic code.
- describe how a protein is synthesized from mRNA.

One of the definitions of a gene is as follows: a segment of deoxyribonucleic acid (DNA) carrying the code for a specific polypeptide. Each molecule of messenger RNA (mRNA) is a transcribed copy of a gene that is used by a cell for synthesizing a polypeptide chain. If a protein contains two or more different polypeptide chains, each chain is coded by a different gene. We turn now to the question of how the sequence of nucleotides in a molecule of ribonucleic acid (RNA) is translated into an amino acid sequence.

How can a molecule containing just 4 different nucleotides specify the sequence of the 20 amino acids that occur in proteins? If each nucleotide coded for 1 amino acid, then obviously the nucleic acids could code for only 4 amino acids. What if amino acids were coded for by groups of 2 nucleotides? There are  $4^2$ , or 16, different combinations of 2 nucleotides (AA, AU, AC, AG, UU, and so on). Such a code is more extensive but still not adequate to code for 20 amino acids. However, if the nucleotides are arranged in groups of 3, the number of different possible combinations is  $4^3$ , or 64. Here we have a code that is extensive enough to direct the synthesis of the primary structure of a protein molecule.



**Video:** NDSU Virtual Cell Animations project animation "Translation". For more information, see [VCell, NDSU](http://vcell.ndsu.nodak.edu) (opens in new window) [vcell.ndsu.nodak.edu]

The genetic code can therefore be described as the identification of each group of three nucleotides and its particular amino acid. The sequence of these triplet groups in the mRNA dictates the sequence of the amino acids in the protein. Each individual three-nucleotide coding unit, as we have seen, is called a *codon*.

Protein synthesis is accomplished by orderly interactions between mRNA and the other ribonucleic acids (transfer RNA [tRNA] and ribosomal RNA [rRNA]), the ribosome, and more than 100 enzymes. The mRNA formed in the nucleus during transcription is transported across the nuclear membrane into the cytoplasm to the ribosomes—carrying with it the genetic instructions. The process in which the information encoded in the mRNA is used to direct the sequencing of amino acids and thus ultimately to synthesize a protein is referred to as translation.

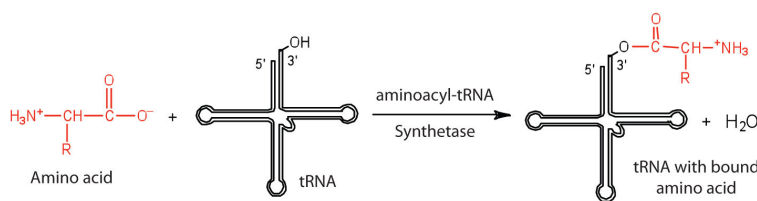


Figure 16.5.1: Binding of an Amino Acid to Its tRNA

Before an amino acid can be incorporated into a polypeptide chain, it must be attached to its unique tRNA. This crucial process requires an enzyme known as aminoacyl-tRNA synthetase (Figure 16.5.1). There is a specific aminoacyl-tRNA synthetase for each amino acid. This high degree of specificity is vital to the incorporation of the correct amino acid into a protein. After the amino acid molecule has been bound to its tRNA carrier, protein synthesis can take place. Figure 16.5.2 depicts a schematic stepwise representation of this all-important process.

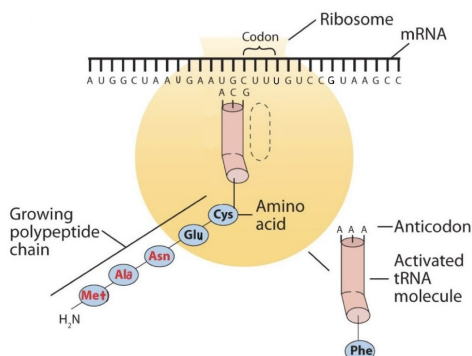


Figure 16.5.2a: The Elongation Steps in Protein Synthesis - Protein synthesis is already in progress at the ribosome. The growing polypeptide chain is attached to the tRNA that brought in the previous amino acid (in this illustration, cys.)

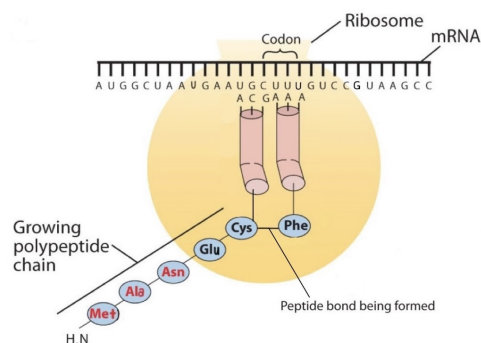


Figure 16.5.2b: The Elongation Steps in Protein Synthesis - An activated tRNA, which has the anticodon AAA, binds to the ribosome next to the previous bound tRNA and interacts with the mRNA molecule through basepairing of the codon and anticodon. The amino acid Phe is being incorporated into the polypeptide chain by the formation of a peptide linkage between the carboxyl group of Cys and the amino acid group of the Phe. This reaction is catalyzed by the enzyme peptidyl transferase, a component of the ribosome.

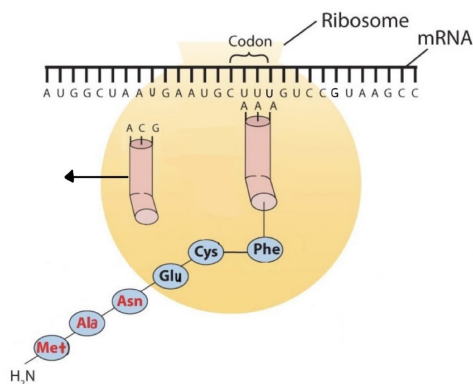


Figure 16.5.2c: The Elongation Steps in Protein Synthesis - The Cys-Phe linkage is now complete, and the growing polypeptide chain remains attached to the tRNA for Phe.

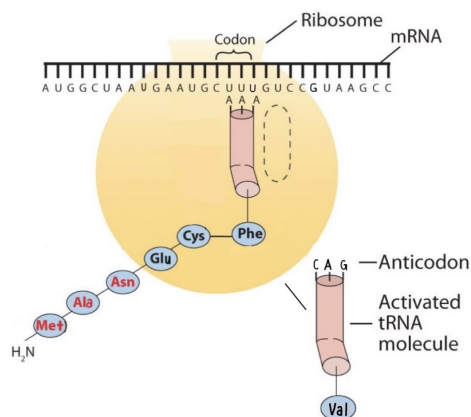


Figure 16.5.2d: The Elongation Steps in Protein Synthesis - The ribosome moves to the right along the mRNA strand. This shift brings the next codon, GUC, into its correct position on the surface of the ribosome. Note that an activated tRNA molecule, containing the next amino acid to be attached to the chain is moving to the ribosome. Steps (b)-(d) will be repeated until the ribosome reaches a stop codon.

Early experimenters were faced with the task of determining which of the 64 possible codons stood for each of the 20 amino acids. The cracking of the genetic code was the joint accomplishment of several well-known geneticists—notably Har Khorana, Marshall Nirenberg, Philip Leder, and Severo Ochoa—from 1961 to 1964. The genetic dictionary they compiled, summarized in Figure 16.5.3 shows that 61 codons code for amino acids, and 3 codons serve as signals for the termination of polypeptide synthesis (much like the period at the end of a sentence). Notice that only methionine (AUG) and tryptophan (UGG) have single codons. All other amino acids have two or more codons.

		Second base				
		U	C	A	G	
First base	U	Phe	Ser	Tyr	Cys	U
		Phe	Ser	Tyr	Cys	C
		Leu	Ser	Stop	Stop	A
		Leu	Ser	Stop	Trp	G
	C	Leu	Pro	His	Arg	U
		Leu	Pro	His	Arg	C
		Leu	Pro	Gln	Arg	A
		Leu	Pro	Gln	Arg	G
	A	Ile	Thr	Asn	Ser	U
		Ile	Thr	Asn	Ser	C
		Ile	Thr	Lys	Arg	A
		Met	Thr	Lys	Arg	G
	G	Val	Ala	Asp	Gly	U
		Val	Ala	Asp	Gly	C
		Val	Ala	Glu	Gly	A
		Val	Ala	Glu	Gly	G

Figure 16.5.3: The Genetic Code

#### ✓ Example 16.5.1: Using the Genetic Code

A portion of an mRNA molecule has the sequence 5'-AUGCCACGAGUUGAC-3'. What amino acid sequence does this code for?

#### Solution

Use Figure 16.5.3 to determine what amino acid each set of three nucleotides (codon) codes for. Remember that the sequence is read starting from the 5' end and that a protein is synthesized starting with the N-terminal amino acid. The sequence 5'-AUGCCACGAGUUGAC-3' codes for met-pro-arg-val-asp.

### ? Exercise 16.5.4

A portion of an RNA molecule has the sequence 5'-AUGCUGAAUUGCGUAGGA-3'. What amino acid sequence does this code for?

Further experimentation threw much light on the nature of the genetic code, as follows:

1. The code is virtually universal; animal, plant, and bacterial cells use the same codons to specify each amino acid (with a few exceptions).
2. The code is “degenerate”; in all but two cases (methionine and tryptophan), more than one triplet codes for a given amino acid.
3. The first two bases of each codon are most significant; the third base often varies. This suggests that a change in the third base by a mutation may still permit the correct incorporation of a given amino acid into a protein. The third base is sometimes called the “wobble” base.
4. The code is continuous and nonoverlapping; there are *no* nucleotides between codons, and adjacent codons do not overlap.
5. The three termination codons are read by special proteins called release factors, which signal the end of the translation process.
6. The codon AUG codes for methionine and is also the initiation codon. Thus methionine is the first amino acid in each newly synthesized polypeptide. This first amino acid is usually removed enzymatically before the polypeptide chain is completed; the vast majority of polypeptides do not begin with methionine.

### Summary

In translation, the information in mRNA directs the order of amino acids in protein synthesis. A set of three nucleotides (codon) codes for a specific amino acid.

---

This page titled [16.5: Protein Synthesis and the Genetic Code](#) is shared under a [CC BY-NC-SA 4.0](#) license and was authored, remixed, and/or curated by [Eden Francis](#) via [source content](#) that was edited to the style and standards of the LibreTexts platform.

- [19.4: Protein Synthesis and the Genetic Code](#) by Anonymous is licensed [CC BY-NC-SA 3.0](#). Original source: <https://2012books.lardbucket.org/books/introduction-to-chemistry-general-organic-and-biological>.