



College of pharmacy

Biochemistry I third stage

Dr. Maytham Ahmed

Lecture 3

RNA Structure, Transcription, and Translation

RNA Structure and Transcription

Transcription is the process during which a **DNA strand** serves as a **template** for the synthesis of **RNA**. **Transcription** produces **messenger RNAs** that are **translated** into sequences of **amino acids** (polypeptide chains or proteins), and **ribosomal RNAs**, **transfer RNAs**, molecules that perform specialized structural, catalytic, and regulatory functions and are **not translated**, that is, they are **noncoding RNAs** (**ncRNAs**). A central feature of transcription is that it is highly **selective**.

RNA Structure

There are three major types of RNA that participate in the process of protein synthesis:

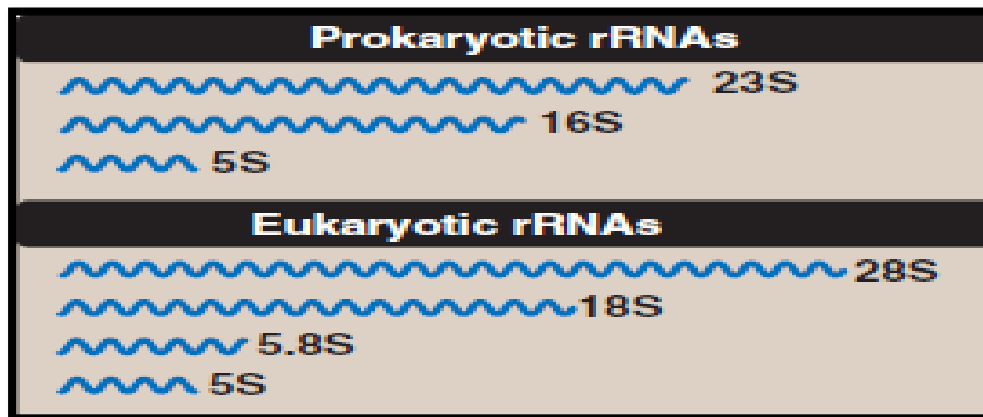
1. ribosomal RNA (rRNA)
2. transfer RNA (tRNA)
3. messenger RNA (mRNA)

The three major types of RNA also **differ** from each other in **size**, **function**, and **special structural modifications**. Like DNA, these three types of RNA are unbranched polymeric molecules composed of **nucleoside monophosphates** joined together by phosphodiester bonds. RNA differ from DNA in several ways, for example, they are considerably **smaller** than DNA, contain **ribose** instead of **deoxyribose** and **uracil** instead of **thymine**, and exist as **single** strands.

1. Ribosomal RNA

rRNAs are found in association with several proteins as components of the ribosomes. **Ribosomes** serve as the sites for **protein synthesis**. rRNAs make up about **80%** of the total RNA in the cell. There are **three** distinct size species of rRNA (23S, 16S, and 5S) in **prokaryotic** cells. In the **eukaryotic** cytosol, there are

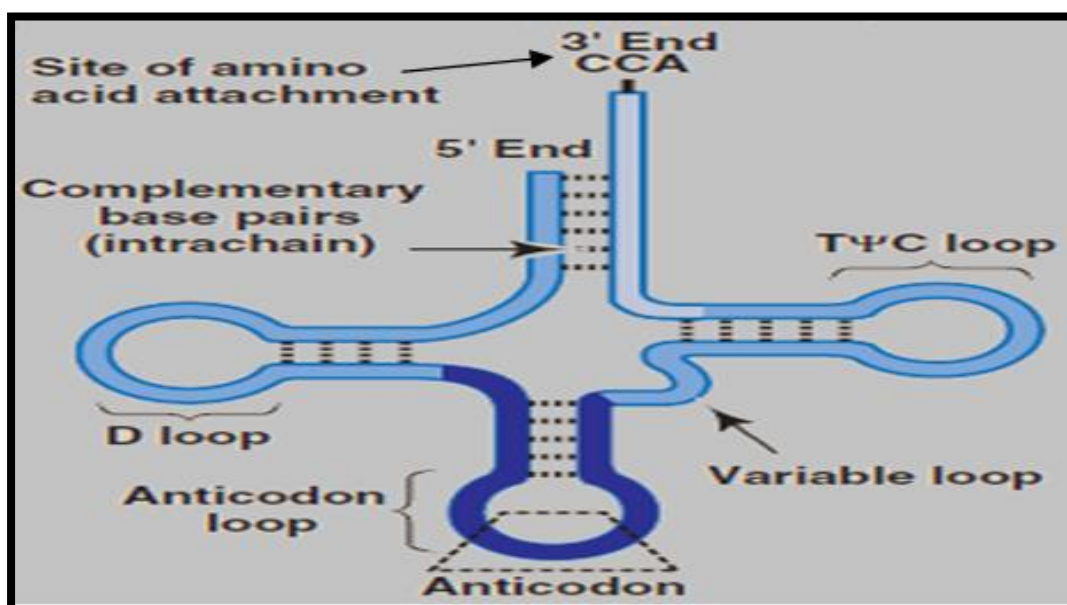
four rRNA species (28S, 18S, 5.8S, and 5S). “S” is the Svedberg unit, which is related to the molecular weight and shape of the compound.



Prokaryotic and eukaryotic rRNAs

2. Transfer RNA

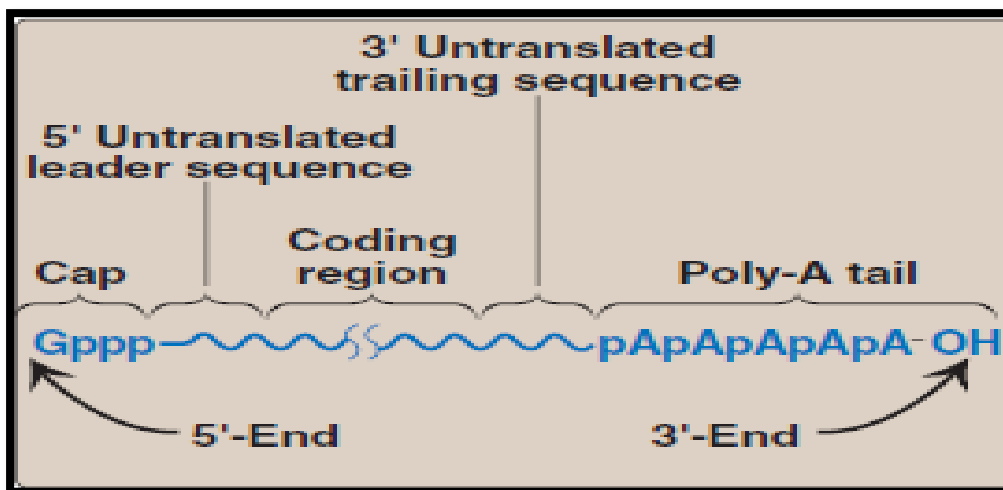
tRNAs are the **smallest** of the three major types of RNA molecules. There is at least one **specific** type of tRNA molecule for each of the **20** amino acids commonly found in proteins. Together, tRNAs make up about **15%** of the total RNA in the cell. Each tRNA serves as an “**adaptor**” molecule that **carries** its **specific amino acid** covalently attached to its 3'-end to the site of protein synthesis. There it recognizes the **genetic code sequence on an mRNA**, which specifies the addition of its amino acid to the growing peptide chain.



tRNA structure

3. Messenger RNA

mRNA comprises only about 5% of the RNA in the cell. The mRNA carries genetic information from the nuclear DNA to the cytosol, where it is used as the template for protein synthesis. If the mRNA carries information from more than one gene, it is said to be polycistronic. Polycistronic mRNA is characteristic of prokaryotes. If the mRNA carries information from just one gene, it is said to be monocistronic and is characteristic of eukaryotes. In addition to the protein coding regions that can be translated, mRNA contains untranslated regions at its 5'- and 3'-ends. Special structural characteristics of eukaryotic (but not prokaryotic) mRNA include a long sequence of adenine nucleotides (a “poly-A tail”) on the 3'-end of the RNA chain, plus a “cap” on the 5'-end consisting of a molecule of 7-methylguanosine attached “backward” (5'→5') through a triphosphate linkage.



Eukaryotic mRNA structure

Transcription of eukaryotic genes

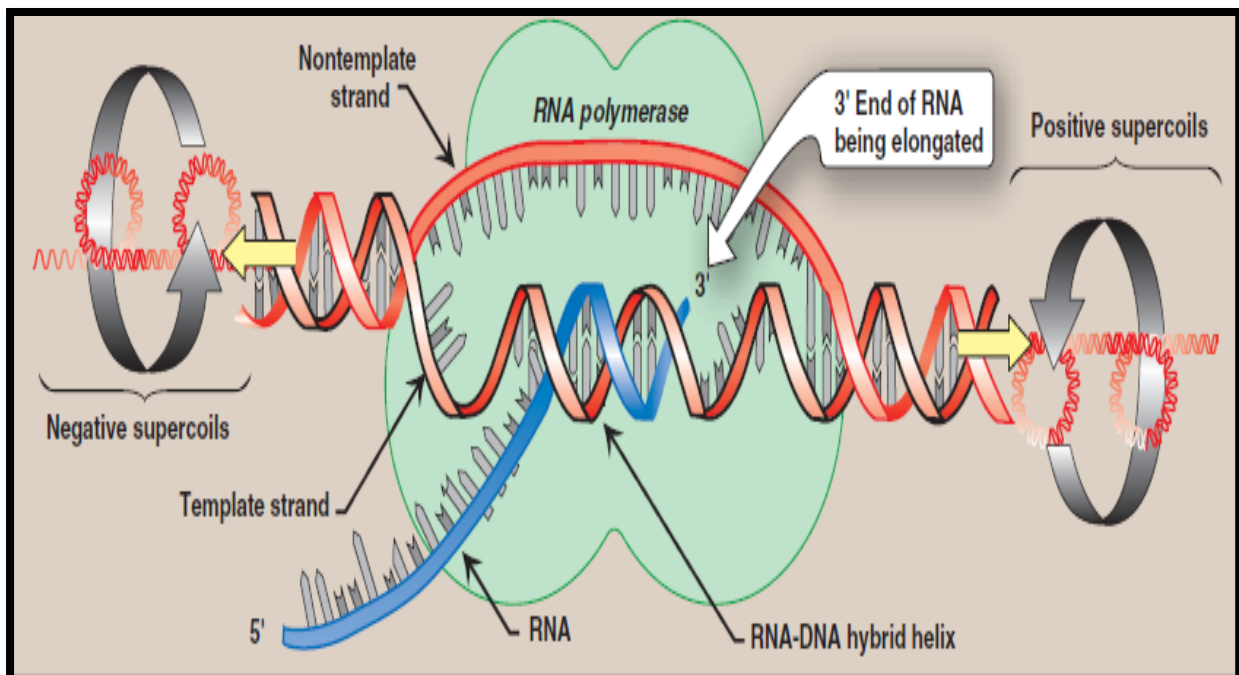
Eukaryotic transcription involves separate polymerases for the synthesis of rRNA, tRNA, and mRNA. In addition, a large number of proteins called transcription factors (TFs) are involved. TFs bind to distinct sites on the DNA either within the core promoter region, close (proximal) to it, or some distance away (distal). They are required both for the assembly of a transcription complex at the promoter and the determination of which genes are to be transcribed. [Note: Each eukaryotic

RNA polymerase has its own **promoters** and **TFs**] For TFs to recognize and bind to their specific DNA sequences.

RNA polymerase is a multisubunit enzyme that recognizes a nucleotide sequence (the **promoter region**) at the **beginning** of a length of DNA that is to be transcribed. It next makes a complementary RNA copy of the DNA template strand, and then recognizes the **end** of the DNA sequence to be transcribed (the **termination region**).

As with replication, **transcription** is always in the **5'→3'** direction. In contrast to **DNA polymerase**, **RNA polymerase** does **not** require a **primer**. RNA is synthesized from its 5'-end to its 3'-end, antiparallel to its DNA template strand. Unwinding generates supercoils in the DNA that can be relieved by **DNA topoisomerases**.

The template is copied as it is in DNA synthesis, in which a **G** on the DNA specifies a **C** in the RNA, a **C** specifies a **G**, a **T** specifies an **A**, but an **A** specifies a **U** instead of a **T**. The **RNA**, then, is **complementary** to the **DNA template** strand and **identical to the coding strand**, with **U** replacing **T**. Within the DNA molecule, regions of both strands can serve as templates for transcription.



Transcription

Translation

Translation is the pathway of **protein synthesis** because the “**language**” of the **nucleotide** sequence on the mRNA is translated into the “**language**” of an **amino acid** sequence. The process of **translation** requires a **genetic code**, through which the **information** contained in the nucleic acid sequence is **expressed** to produce a specific sequence of amino acids. Any **alteration** in the nucleic acid sequence may result in an incorrect amino acid being inserted into the polypeptide chain, potentially causing **disease** or even **death** of the organism.

Genetic code

The genetic code is a **dictionary** that identifies the correspondence between a **sequence of nucleotide bases** and a **sequence of amino acids**. Each individual “**word**” in the code is **composed of three nucleotide bases**. These **genetic words** are called **codons**.

Codons

Codons are presented in the mRNA language of **adenine (A)**, **guanine (G)**, **cytosine (C)**, and **uracil (U)**. Their nucleotide sequences are always written from the **5'-end** to the **3'-end**. The four nucleotide bases are used to produce the three base codons. There are, therefore, **64 different** combinations of bases, taken three at a time (a triplet code). The **initiation** (start) codon for translation is **AUG** (**methionine**). **Sixty one** of the **64** codons code for the **20 common amino acids**. **Three** of the codons, **UAG**, **UGA**, and **UAA**, do not code for amino acids, but rather are **termination** codons. When one of these codons appears in an mRNA sequence, synthesis of the polypeptide coded for by that mRNA **stops**.

5' - BASE	MIDDLE BASE				3' - BASE
	U	C	A	G	
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

1 These four rows show 16 amino acids whose codons begin (5') with A.

3 These four, separated rows show 16 amino acids whose codons end (3') with G.

2 This column shows 16 amino acids whose codons have the middle base U.

4 The codon 5'-AUG-3' designates methionine (Met).

Codons

Characteristics of the genetic code include the following:

- 1. Specificity:** The genetic code is specific, that is, a particular codon always codes for the **same** amino acid.
- 2. Universality:** The genetic code is virtually universal, that is, its specificity has been conserved from very early stages of evolution, with only slight differences in the manner in which the code is translated.
- 3. Degeneracy:** The genetic code is degenerate. Although each codon corresponds to a single amino acid, a given amino acid **may have more than** one triplet coding for it. For example, arginine is specified by six different codons.
- 4. Non-overlapping and comma less:** that is, the code is read from a fixed starting point as a continuous sequence of bases, taken three at a time. For example, AGCUGGAUACAU is read as AGC/UGG/AUA/CAU without any “punctuation” between the codons.

Components required for translation

A large number of components are required for the synthesis of a protein. These include:

1. Amino acids

All the amino acids that eventually appear in the finished protein must be present at the time of protein synthesis.

2. Transfer RNA

At least one specific type of **tRNA** is required for each amino acid. In humans, there are at least **50** species of **tRNA**. Because there are only **20** different **amino acids** commonly carried by tRNA, some amino acids have **more than one** specific **tRNA** molecule. This is particularly true of those amino acids that are coded for by several codons.

3. Aminoacyl-tRNA synthetases

This family of enzymes is required for **attachment** of amino acids to their corresponding tRNAs.

4. Messenger RNA

The specific mRNA required as a template for the synthesis of the desired polypeptide chain must be present.

5. Functionally competent ribosomes

Ribosomes are large complexes of **protein** and **ribosomal RNA**. Ribosomes are the macromolecular complexes in which the synthesis of proteins occurs.

6. Protein factors

Initiation, elongation, and termination (or release) factors are required for peptide synthesis.

7. Sources of energy (ATP and GTP)

Steps of protein synthesis in prokaryotic

