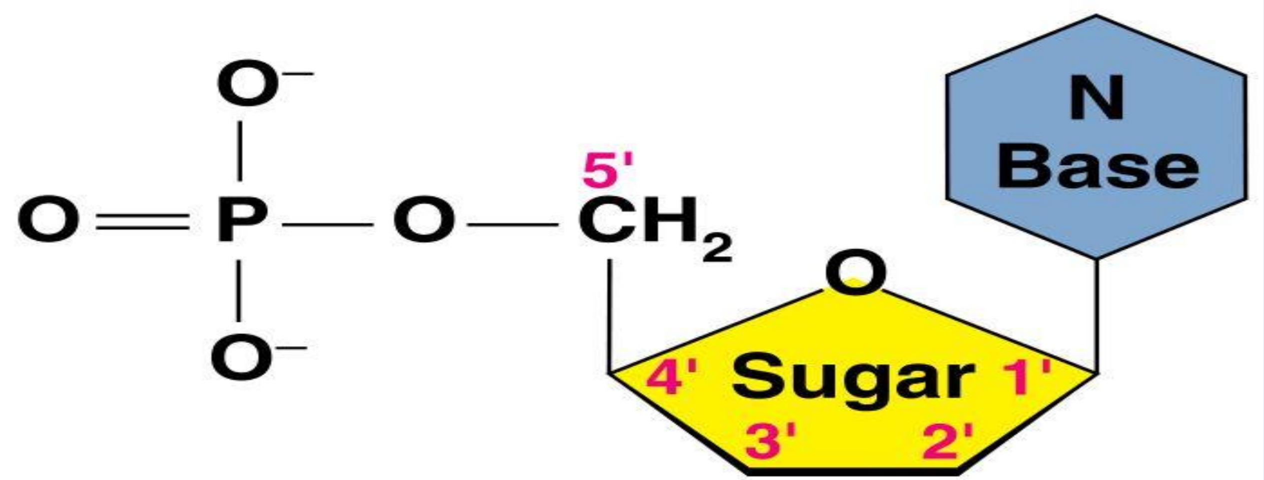
Nucleic Acids

Nucleic Acids

- **Nucleic acids** are molecules that store information for cellular growth and reproduction
- There are two types of nucleic acids:

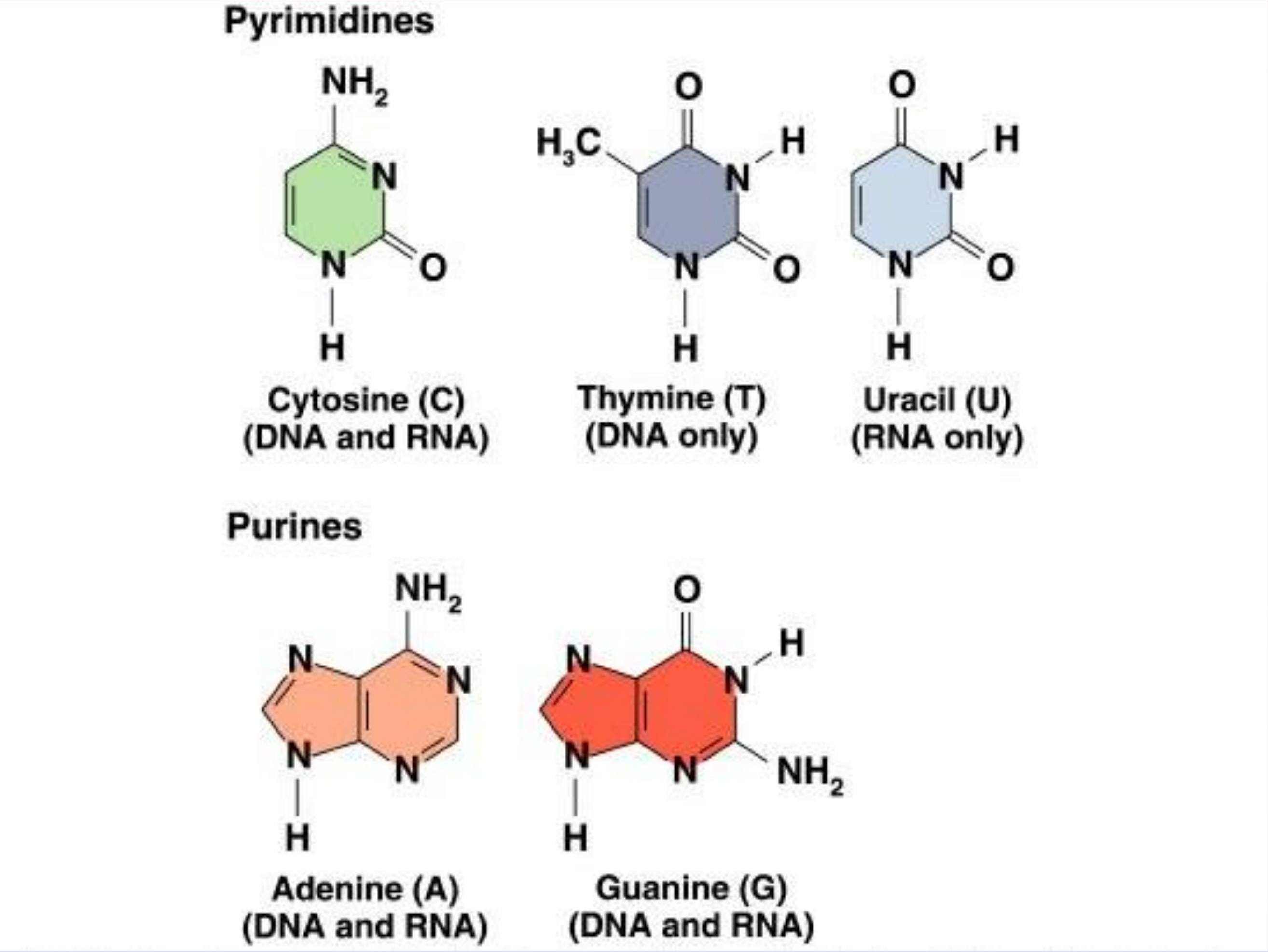
  - **deoxyribonucleic acid (DNA)** and **ribonucleic acid (RNA)**
- These are polymers consisting of long chains of monomers called nucleotides
- A **nucleotide** consists of a nitrogenous base, a pentose sugar and a phosphate group:

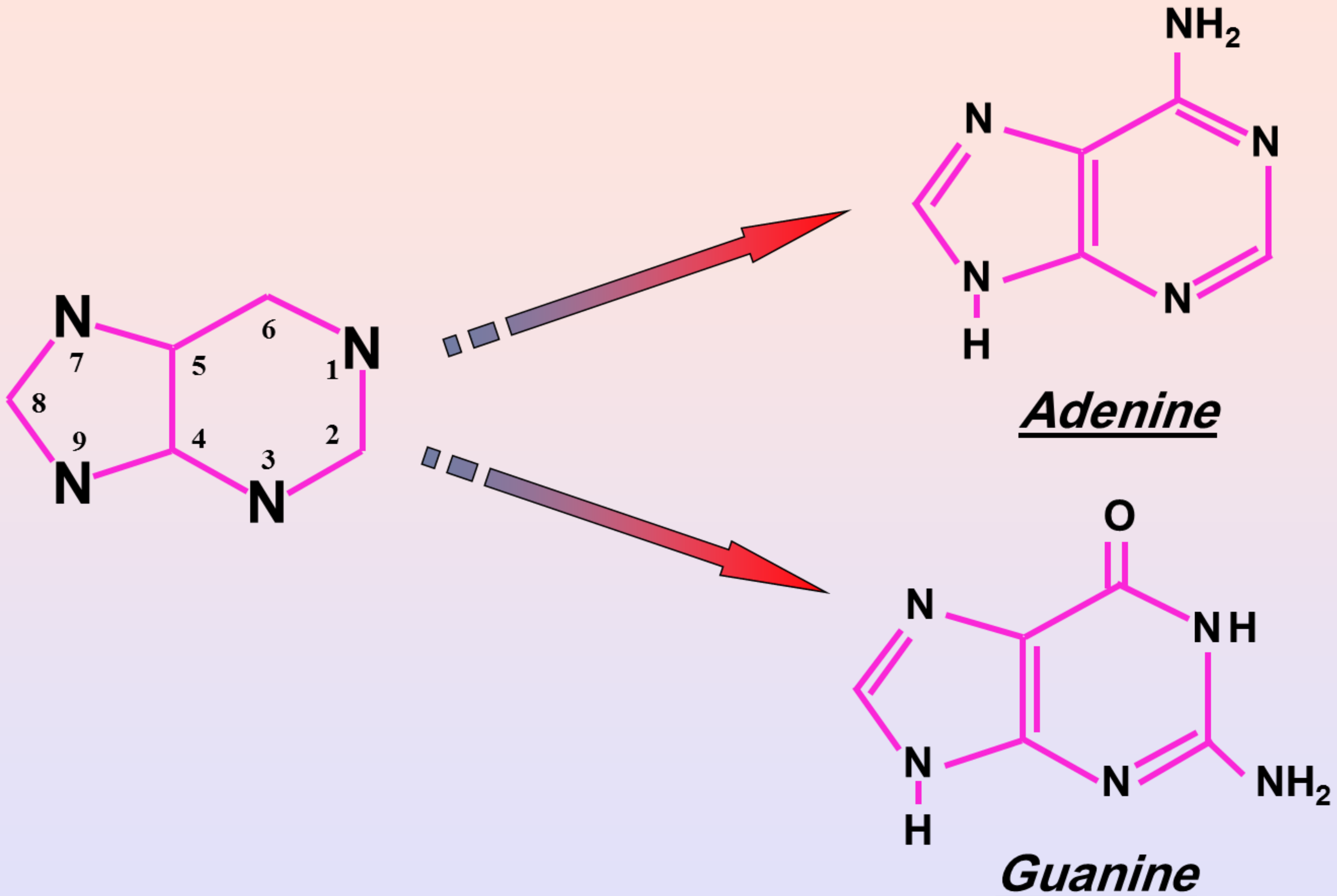


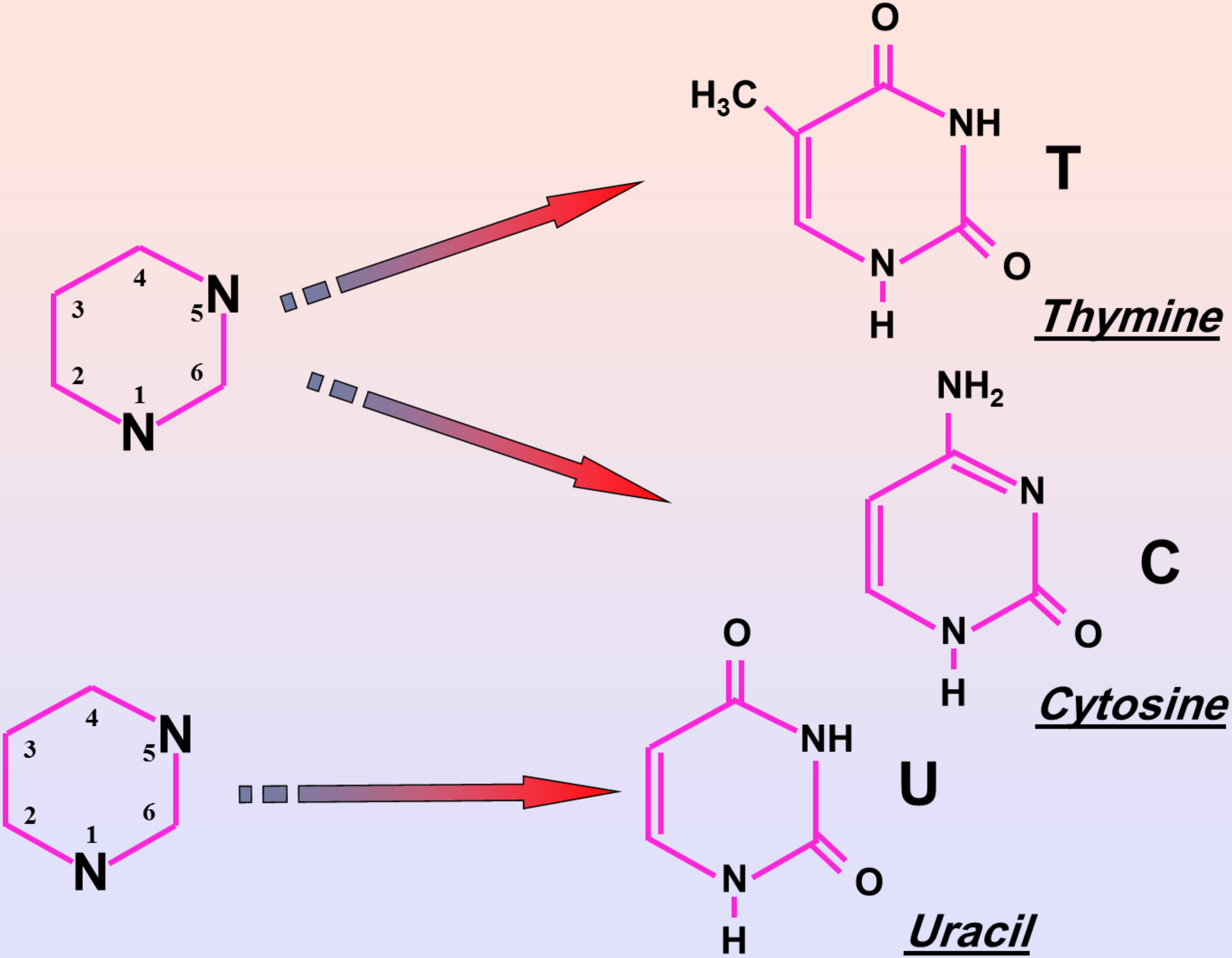
Timberlake, *General, Organic, and Biological Chemistry*. Copyright © Pearson Education Inc., publishing as Benjamin Cummings

Nitrogen Bases

- The **nitrogen bases** in nucleotides consist of two general types:
  - **purines**: adenine (A) and guanine (G)
  - **pyrimidines**: cytosine (C), thymine (T) and Uracil (U)



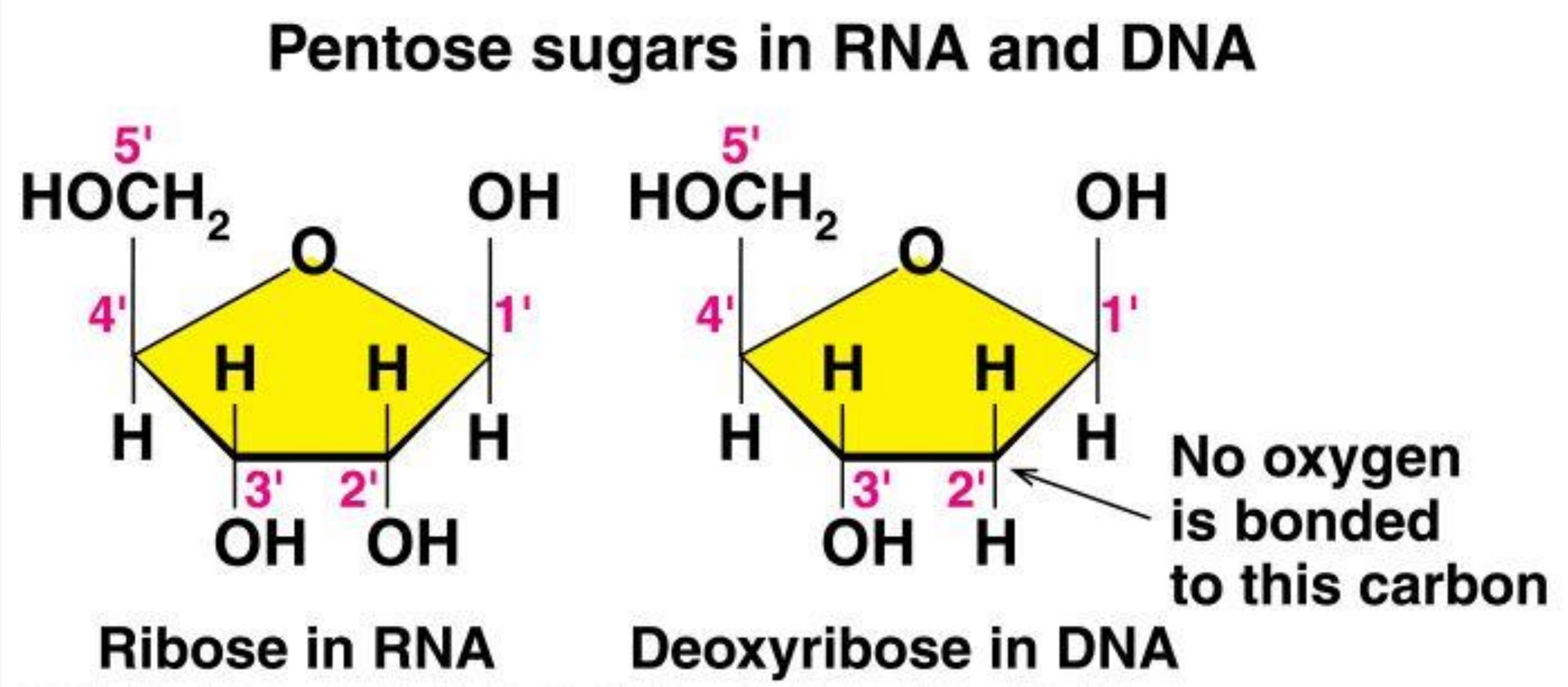

N
N
N
N
1
2
3
4
5
6
7
8
9
NH2
N
N
N
N
H
Adenine
O
N
NH
N
N
H
NH2
Guanine



4
3
5
N
2
6
1
N
4
3
5
N
2
6
1
N
O
H_3C
NH
T
N
O
H
Thymine
NH_2
N
C
N
O
H
Cytosine
O
NH
U
N
O
H
Uracil

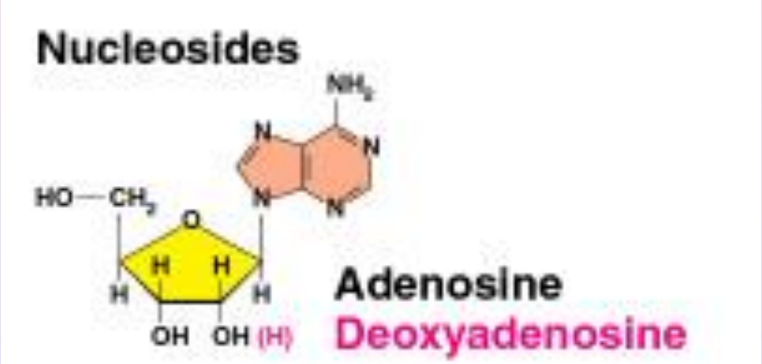
Pentose Sugars

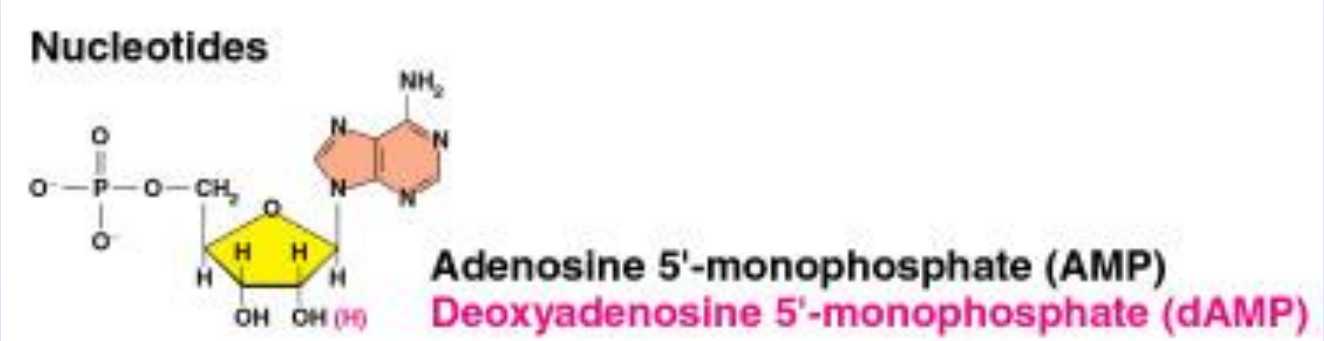
- There are two related **pentose sugars**:
  - RNA contains **ribose**
  - DNA contains **deoxyribose**
- The sugars have their carbon atoms numbered with primes to distinguish them from the nitrogen bases



Nucleosides and Nucleotides

- A nucleoside consists of a nitrogen base linked by a glycosidic bond to C1’ of a ribose or deoxyribose
- Nucleosides are named by changing the the nitrogen base ending to *-osine* for purines and *–idine* for pyrimidines
- A nucleotide is a nucleoside that forms a phosphate ester with the C5’ OH group of ribose or deoxyribose
- Nucleotides are named using the name of the nucleoside followed by *5’-monophosphate*



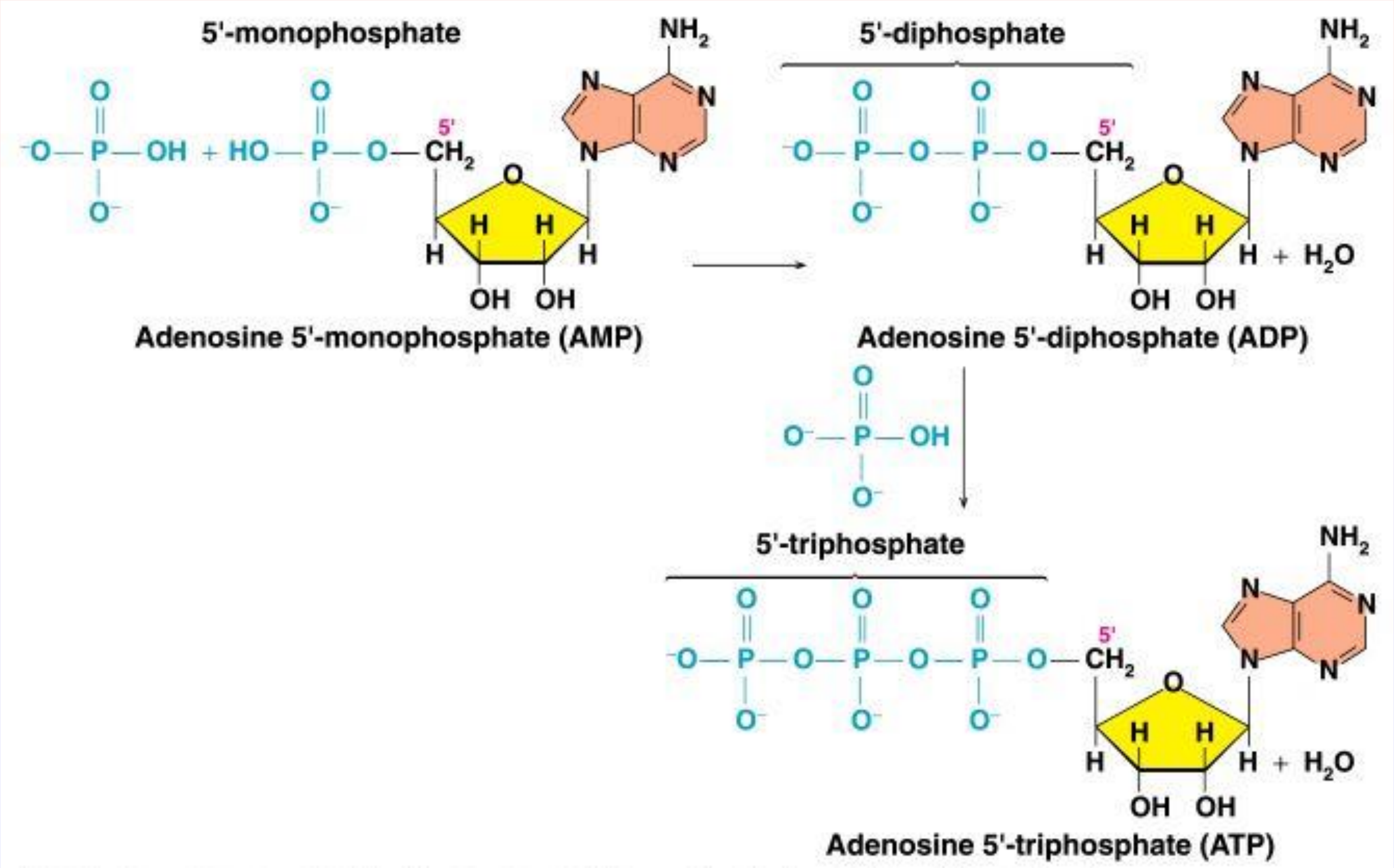


Names of Nucleosides and Nucleotides

Base	Nucleosides	Nucleotides
RNA		
Adenine (A)	Adenosine (A)	Adenosine 5'-monophosphate (AMP)
Guanine (G)	Guanosine (G)	Guanosine 5'-monophosphate (GMP)
Cytosine (C)	Cytidine (C)	Cytidine 5'-monophosphate (CMP)
Uracil (U)	Uridine (U)	Uridine 5'-monophosphate (UMP)
DNA		
Adenine (A)	Deoxyadenosine (A)	Deoxyadenosine 5'-monophosphate (dAMP)
Guanine (G)	Deoxyguanosine (G)	Deoxyguanosine 5'-monophosphate (dGMP)
Cytosine (C)	Deoxycytidine (C)	Deoxycytidine 5'-monophosphate (dCMP)
Thymine (T)	Deoxythymidine (T)	Deoxythymidine 5'-monophosphate (dTMP)

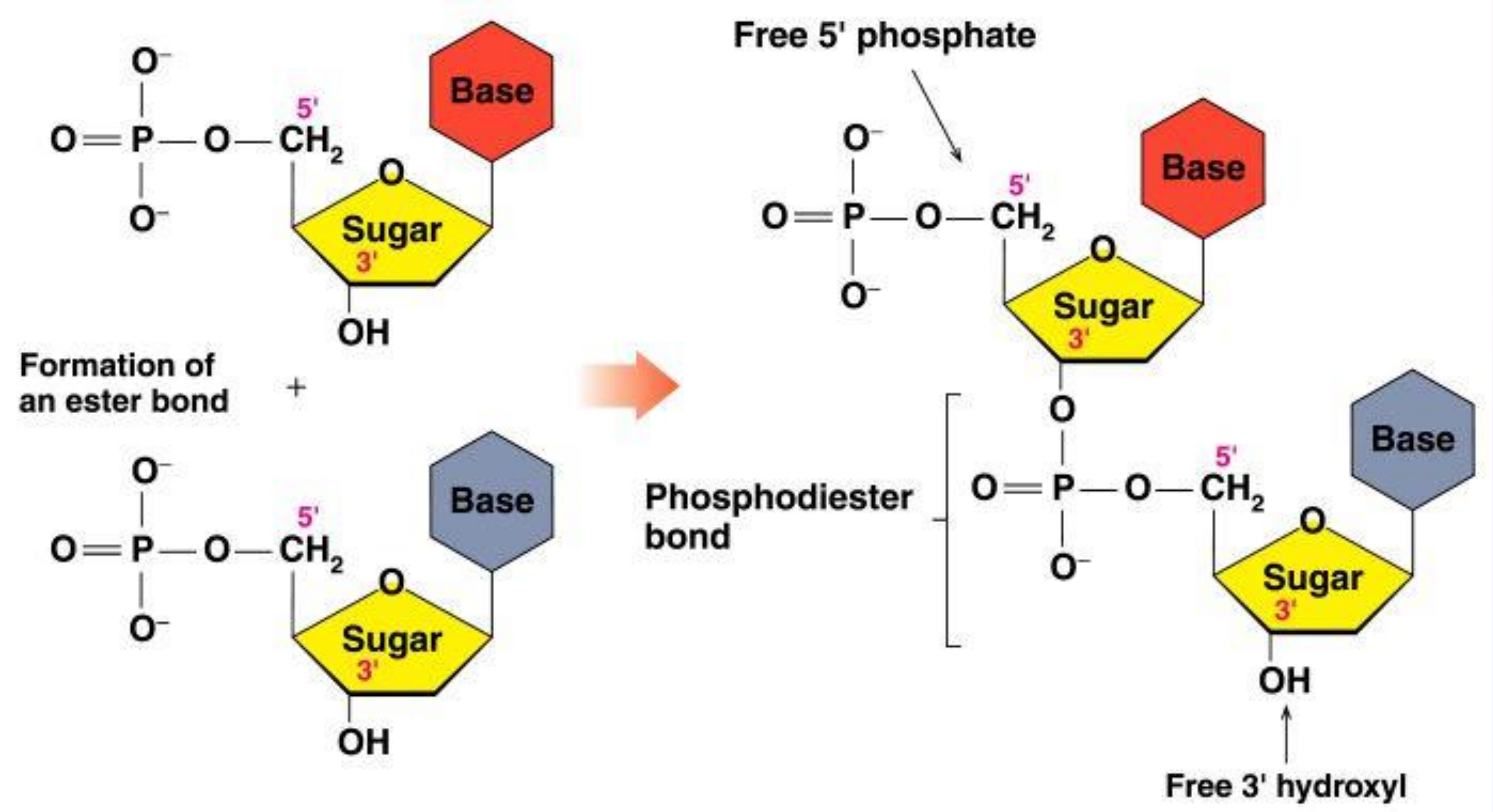
AMP, ADP and ATP

- Additional phosphate groups can be added to the nucleoside 5'-monophosphates to form **diphosphates** and **triphosphates**
- **ATP** is the major energy source for cellular activity



Primary Structure of Nucleic Acids

- The **primary structure** of a nucleic acid is the nucleotide sequence
- The nucleotides in nucleic acids are joined by phosphodiester bonds
- The 3'-OH group of the sugar in one nucleotide forms an ester bond to the phosphate group on the 5'-carbon of the sugar of the next nucleotide



base（purine、pyrimdine）+ribose（deoxyribos

N-glycosyl linkage

nucleoside+phosphate

phosphoester linkage

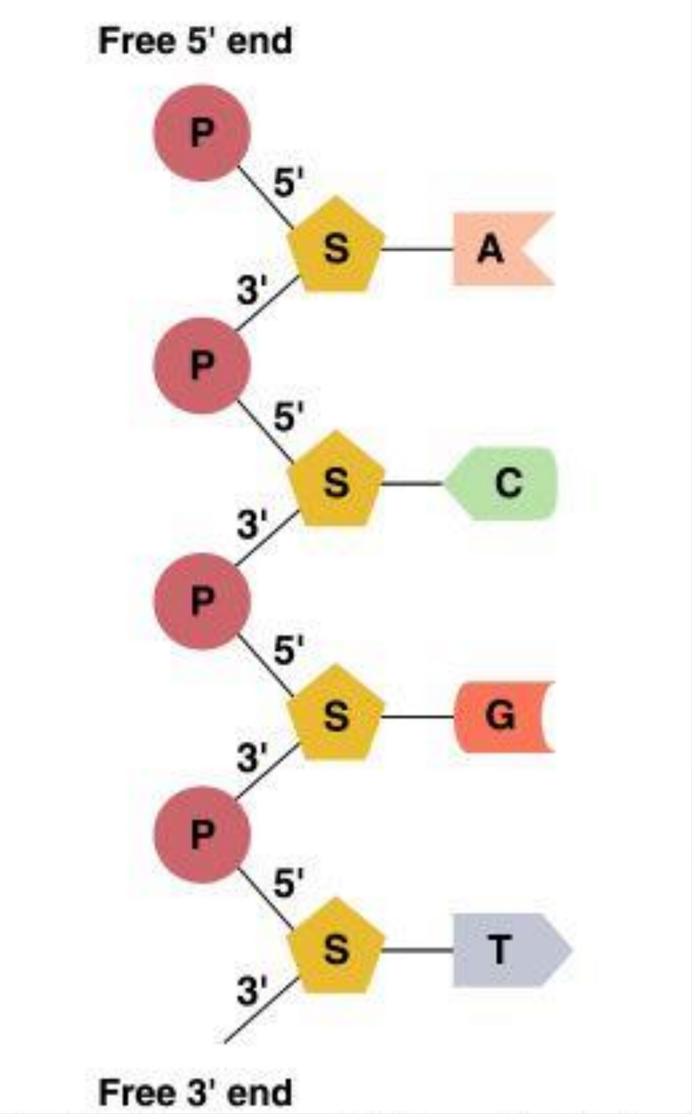
nucleotide

phosphodiester linkage

nucleic acid

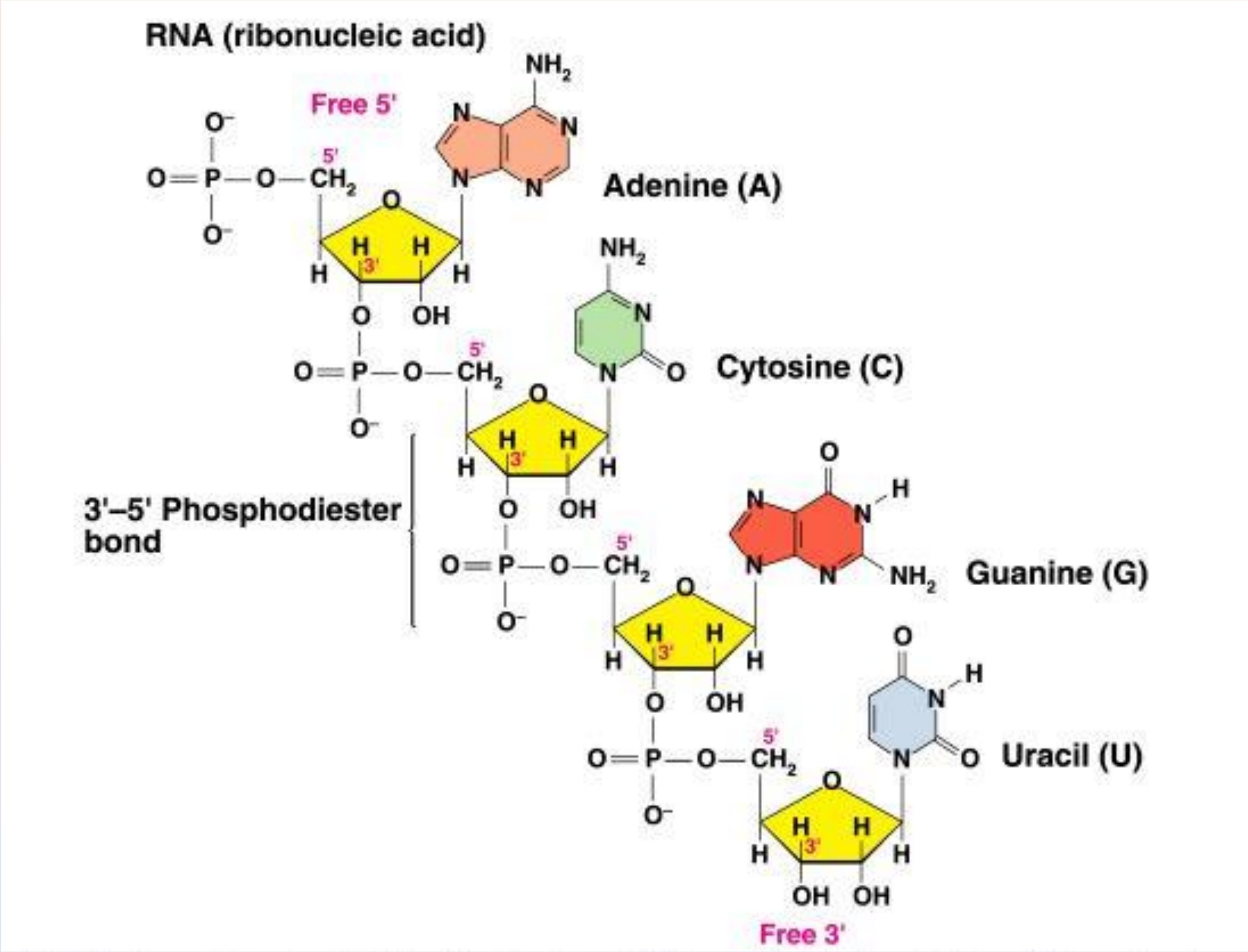
Reading Primary Structure

- A nucleic acid polymer has a free 5'-phosphate group at one end and a free 3'-OH group at the other end
- The sequence is read from the free 5'-end using the letters of the bases
- This example reads
  5'—A—C—G—T—3'



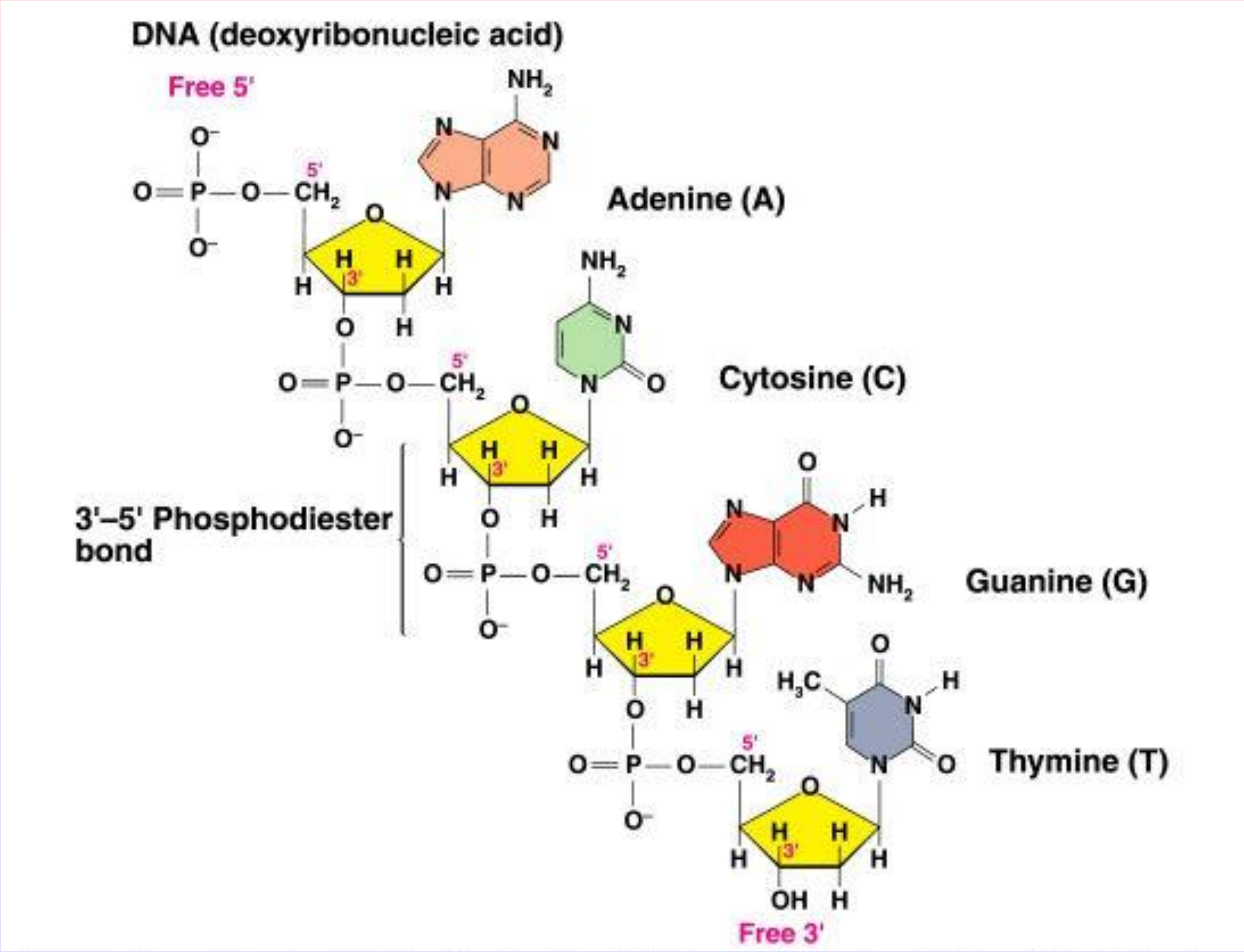
Example of RNA Primary Structure

- In RNA, A, C, G, and U are linked by 3'-5' ester bonds between ribose and phosphate



Example of DNA Primary Structure

- In DNA, A, C, G, and T are linked by 3'-5' ester bonds between deoxyribose and phosphate



Secondary Structure: DNA Double Helix

- In DNA there are two strands of nucleotides that wind together in a **double helix**

  - the strands run in opposite directions

  - the bases are are arranged in step-like pairs

  - the **base pairs** are held together by hydrogen bonding

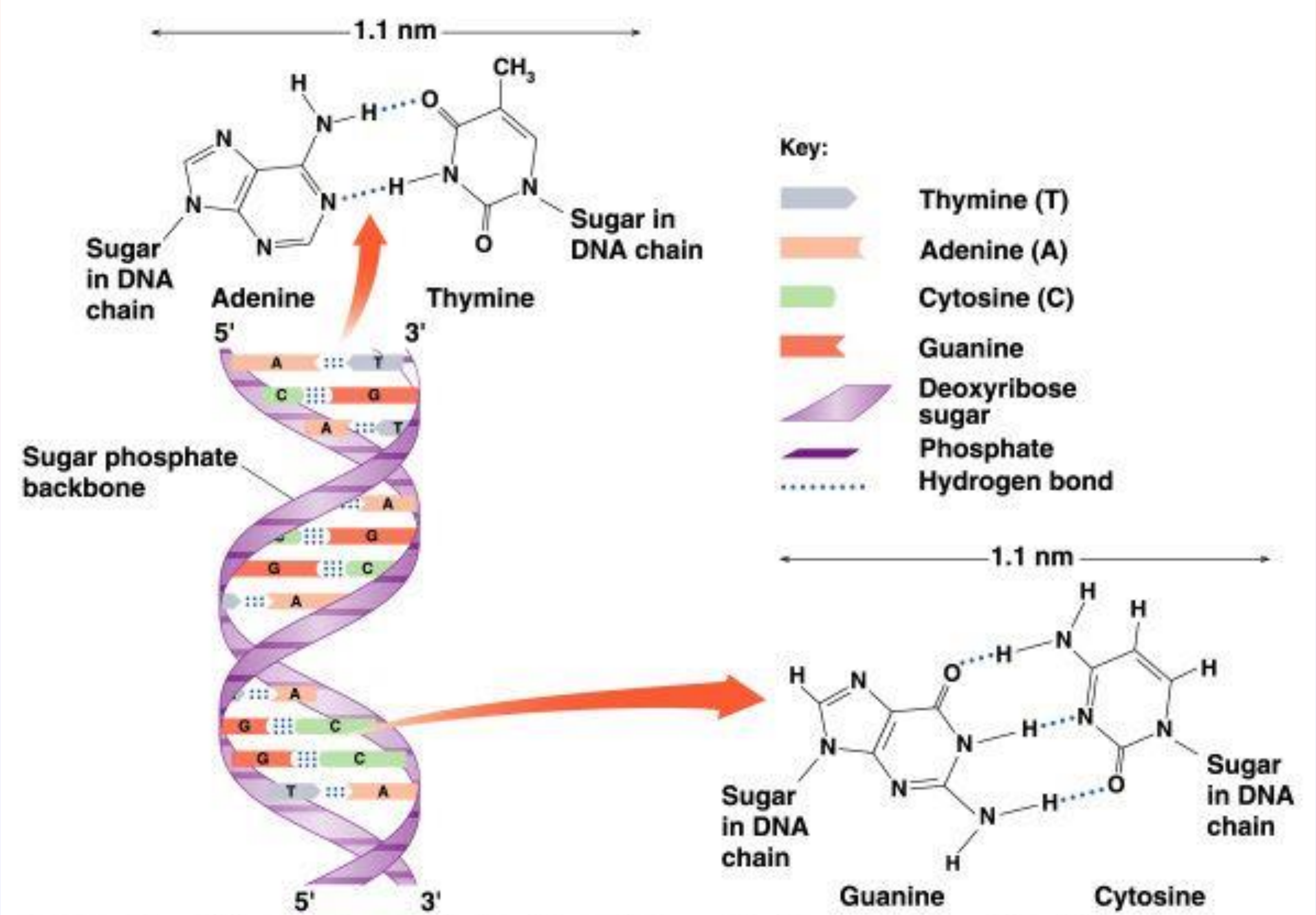
- The pairing of the bases from the two strands is very specific
- The **complimentary base pairs** are **A-T** and **G-C**

  - two hydrogen bonds form between A and T

  - three hydrogen bonds form between G and C

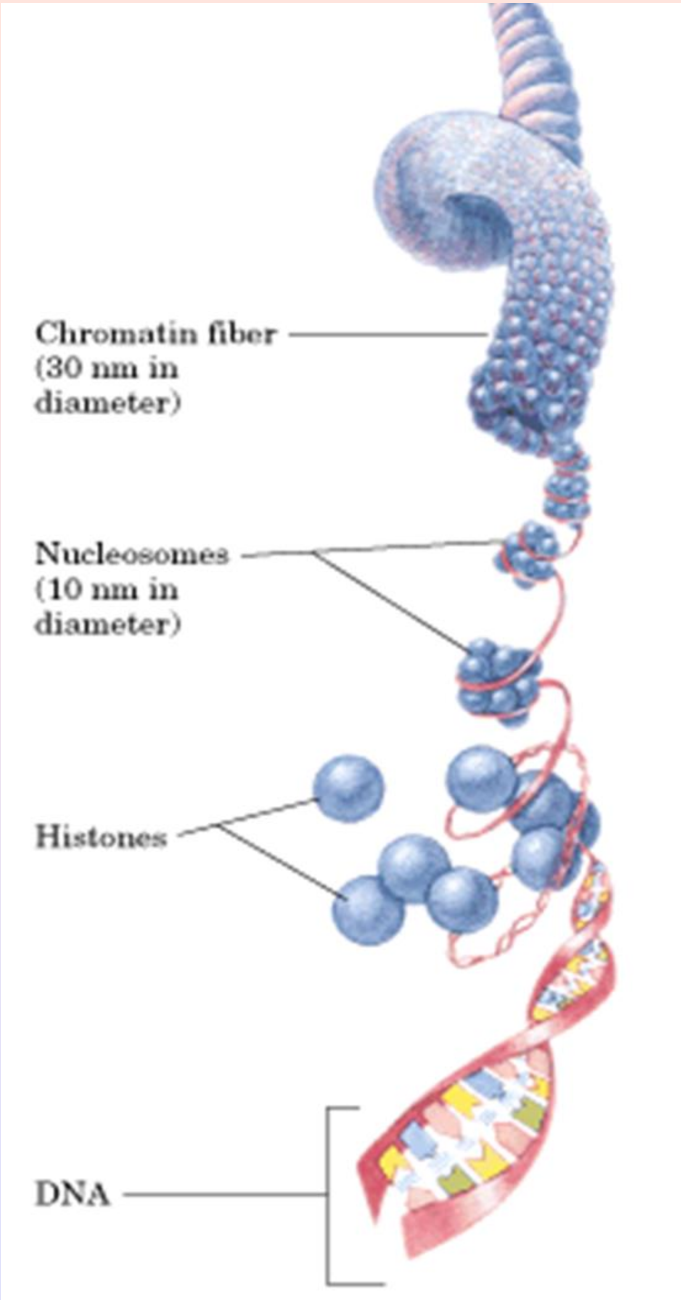
- Each pair consists of a purine and a pyrimidine, so they are the same width, keeping the two strands at equal distances from each other

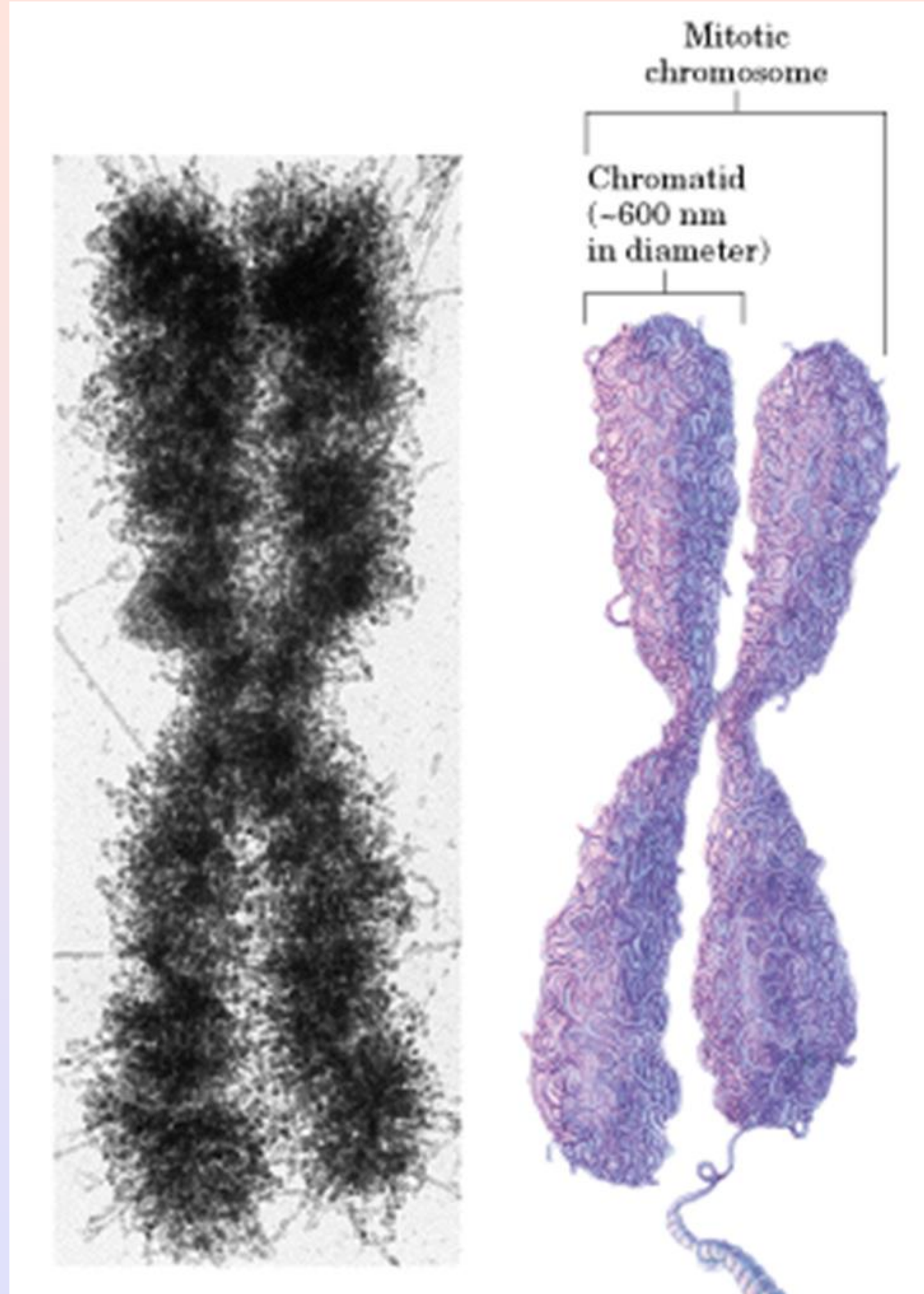
Base Pairing in the DNA Double Helix



Storage of DNA

- In **eukaryotic** cells (animals, plants, fungi) DNA is stored in the **nucleus**, which is separated from the rest of the cell by a semipermeable membrane
- The DNA is only organized into **chromosomes** during cell replication
- Between replications, the DNA is stored in a compact ball called **chromatin**, and is wrapped around proteins called **histones** to form **nucleosomes.**


Chromatin fiber
(30 nm in
diameter)
Nucleosomes
(10 nm in
diameter)
Histones
DNA


Mitotic
chromosome
Chromatid
(~600 nm
in diameter)

Nucleic Acid Function

DNA

Genetic material - sequence of nucleotides encodes different amino acids

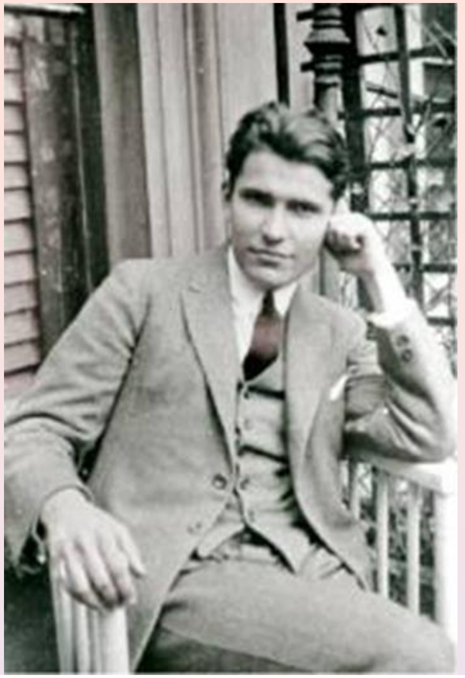
RNA

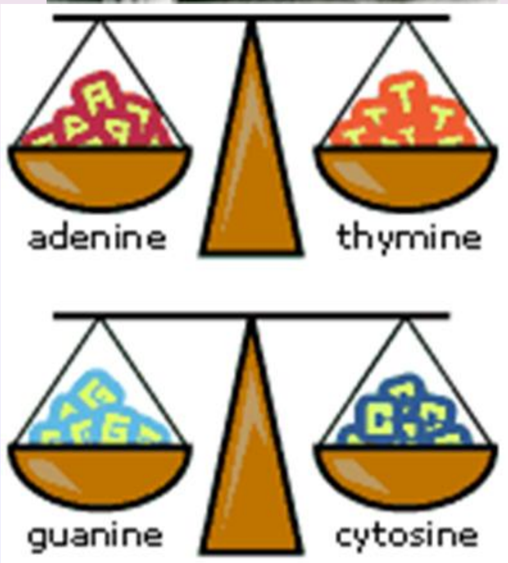
Involved in the transcription/translation of genetic material (DNA)

Genetic material of some viruses

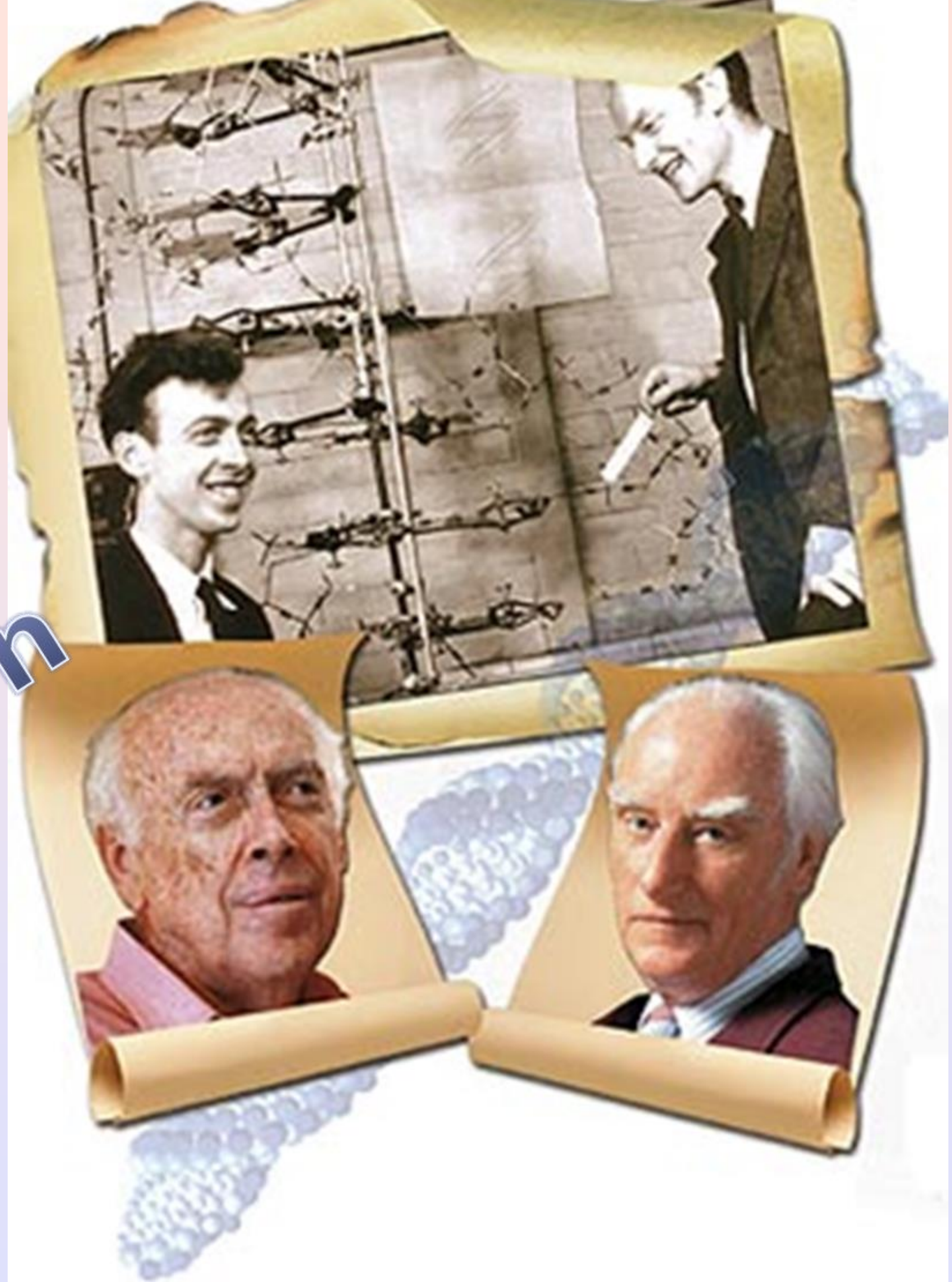
- Chargaff showed:
  - Amount of adenine relative to guanine differs among species
  - Amount of adenine always equals amount of thymine and amount of guanine always equals amount of cytosine

A=T and G=C





Watson



Crick

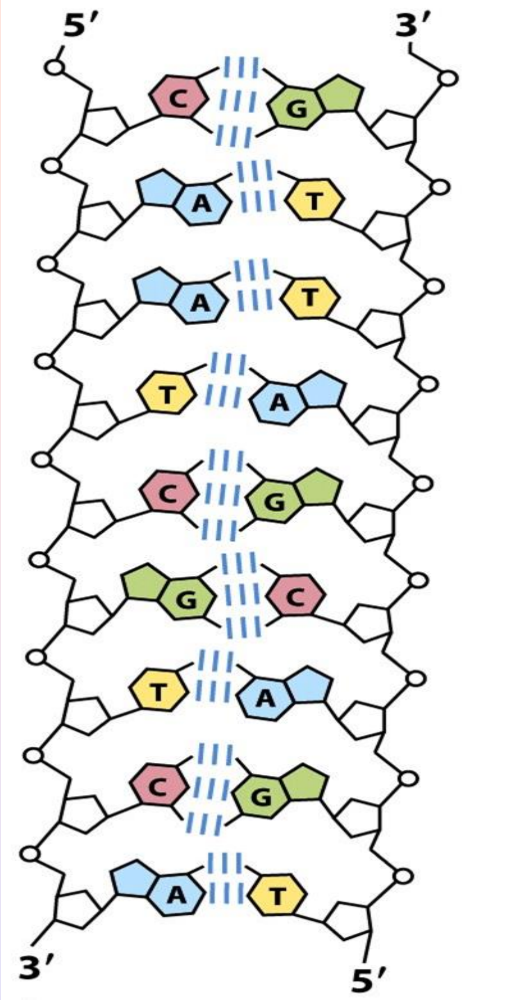
28.10 Base Pairing in DNA: The Watson–Crick Model

- In 1953 Watson and Crick noted that DNA consists of two polynucleotide strands, running in opposite directions and coiled around each other in a double helix
- Strands are held together by hydrogen bonds between specific pairs of bases
- Adenine (A) and thymine (T) form strong hydrogen bonds to each other but not to C or G
- (G) and cytosine (C) form strong hydrogen bonds to each other but not to A or T

The sugar-phosphate chains are on the outside and the strands are held together by chemical bonds between the bases

The Difference in the Strands

- The strands of DNA are complementary because of H-bonding
- Whenever a G occurs in one strand, a C occurs opposite it in the other strand
- When an A occurs in one strand, a T occurs in the other



Properties of a DNA double helix

The strands of DNA are antiparallel

The strands are complimentary

There are Hydrogen bond forces

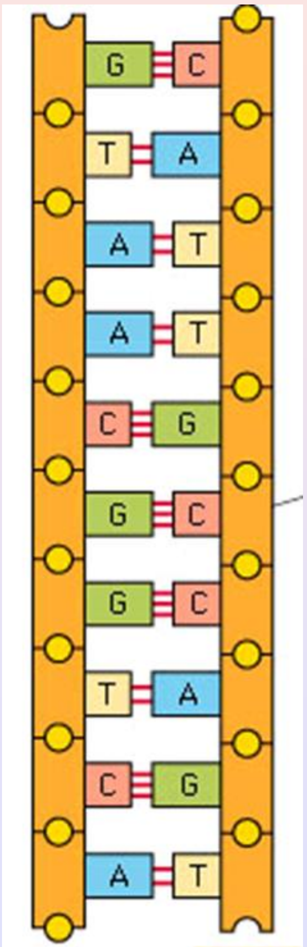
There are base stacking interactions

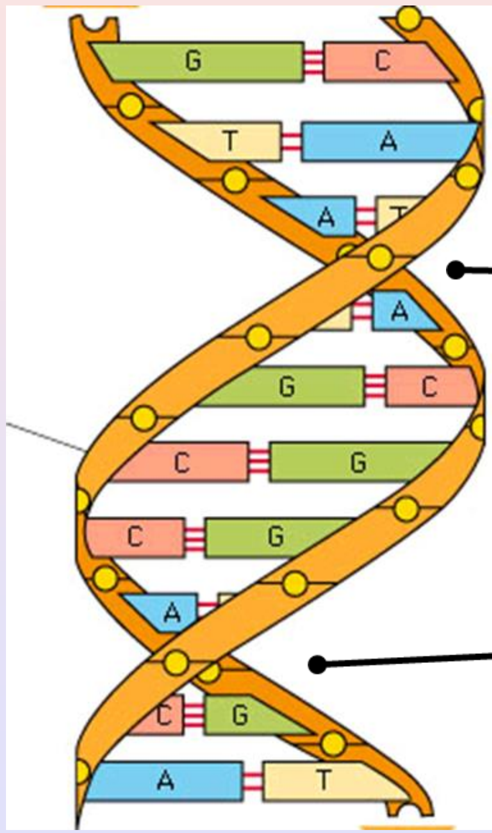
There are 10 base pairs per turn

Structure consists of "major" grooves and "minor" grooves

Major grooves are critical for binding proteins that regulate DNA function

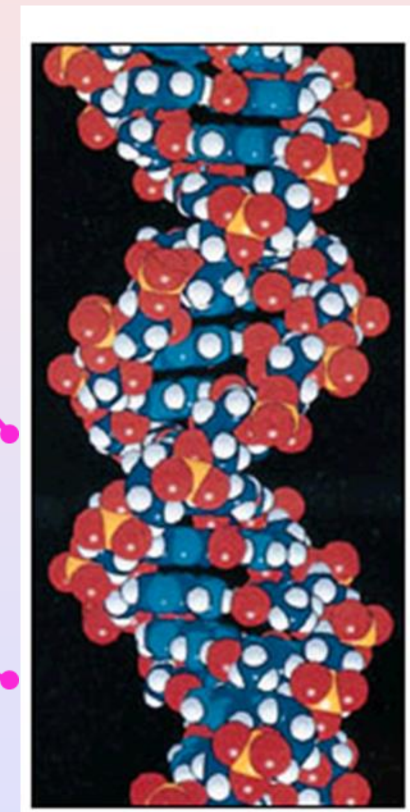
Nucleic Acid Structure
The double helix





Minor Groove

Major Groove



Factors Affecting DNA Denaturation

- The midpoint of melting (T_m) depends on base composition
  - high CG increases T_m
- T_m depends on DNA length
  - Longer DNA has higher T_m
  - Important for short DNA
- T_m depends on pH and ionic strength
  - High salt increases T_m

Mutagenesis

Four classes of mutations are (1) spontaneous mutations (molecular decay), (2) mutations due to error prone replication bypass of naturally occurring DNA damage (also called error prone translesion synthesis), (3) errors introduced during DNA repair, and (4) induced mutations caused by mutagens. Scientists may also deliberately introduce mutant sequences through DNA manipulation for the sake of scientific experimentation.

Mutations occur when copying **errors** cause a **change** in the **sequence** of DNA nucleotide bases

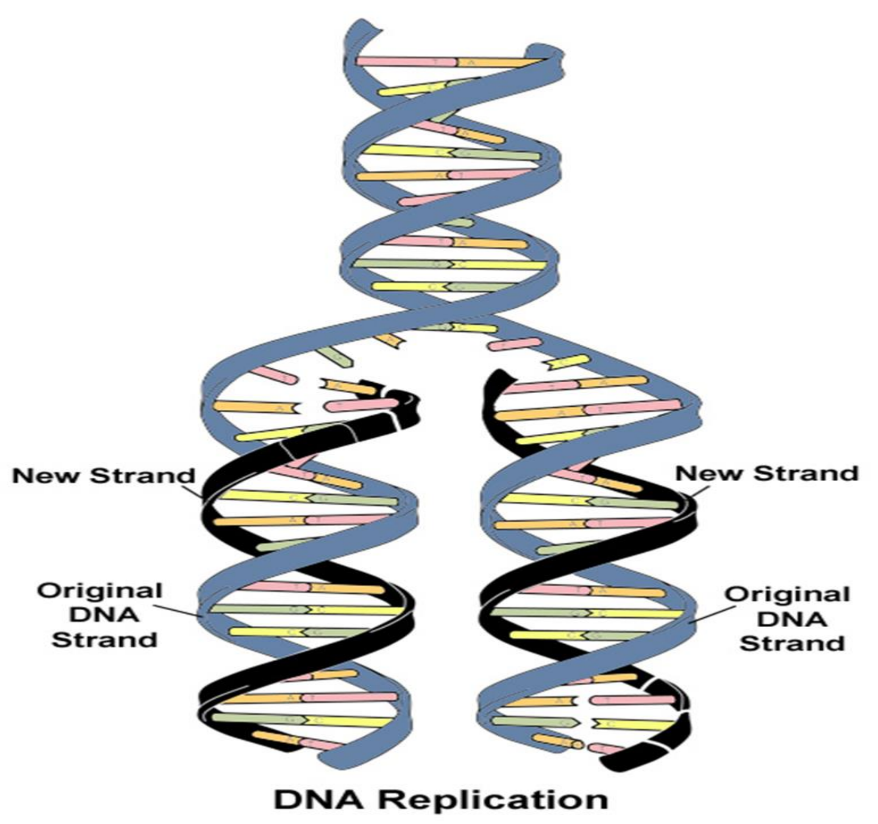
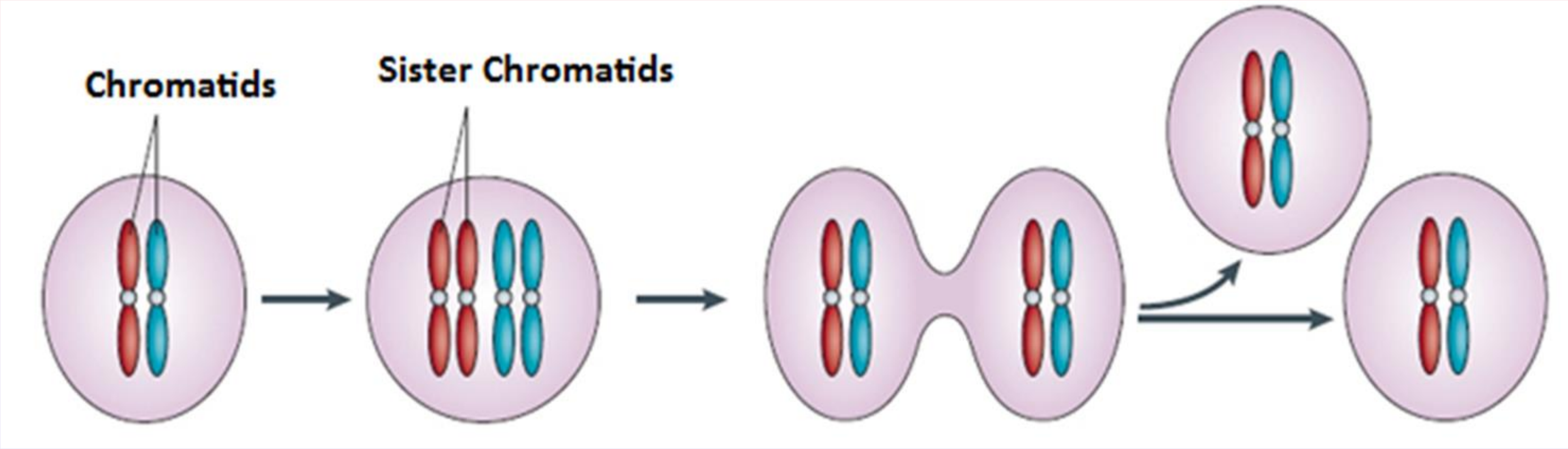


Image adapted from: National Human Genome Research Institute.

In biology, a **mutation** is a permanent change of the nucleotide sequence of the genome of an organism, virus, or extrachromosomal DNA or other genetic elements. Mutations result from damage to DNA which is not repaired or to RNA genomes (typically caused by radiation or chemical mutagens), errors in the process of replication, or from the insertion or deletion of segments of DNA by mobile genetic elements.[1][2][3] Mutations may or may not produce discernible changes in the observable characteristics (phenotype) of an organism. Mutations play a part in both normal and abnormal biological processes including: evolution, cancer, and the development of the immune system, including junctional diversity.

DNA Replication

- When a eukaryotic cell divides, the process is called **mitosis**
  - the cell splits into two identical daughter cells
  - the DNA must be replicated so that each daughter cell has a copy

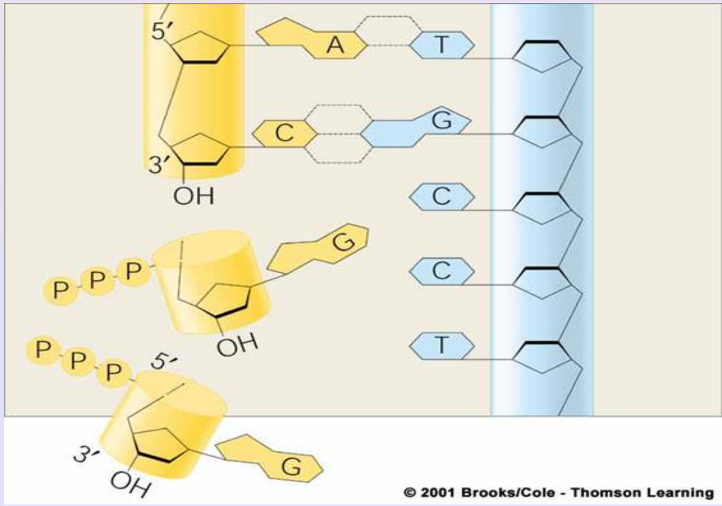


DNA replication involves several processes:

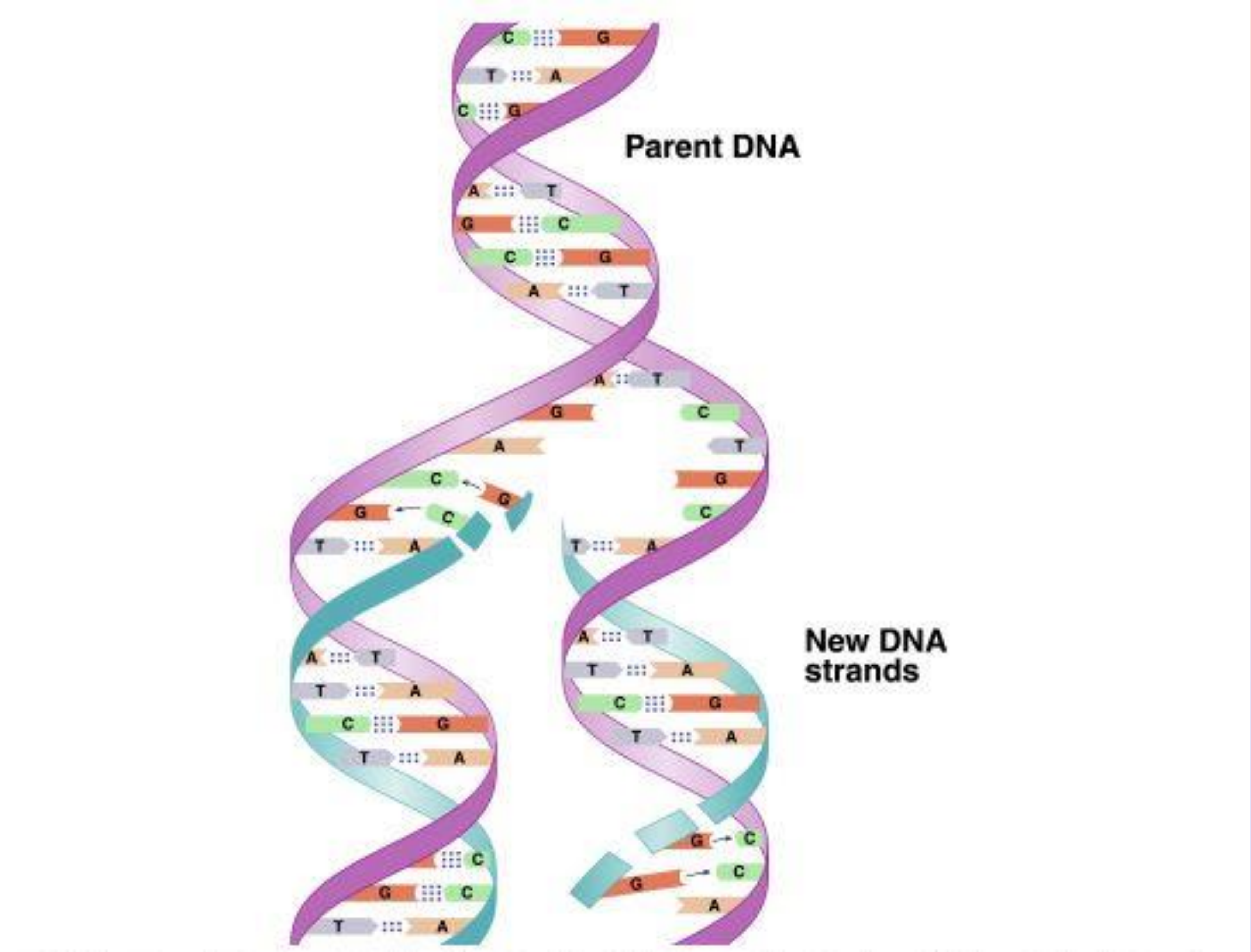
- first, the DNA must be unwound, separating the two strands
- the single strands then act as templates for synthesis of the new strands, which are complimentary in sequence
- bases are added one at a time until two new DNA strands that exactly duplicate the original DNA are produced

The process is called semi-conservative replication because one strand of each daughter DNA comes from the parent DNA and one strand is new.

Energy for strand assembly is provided by removal of two phosphate groups from free nucleotides

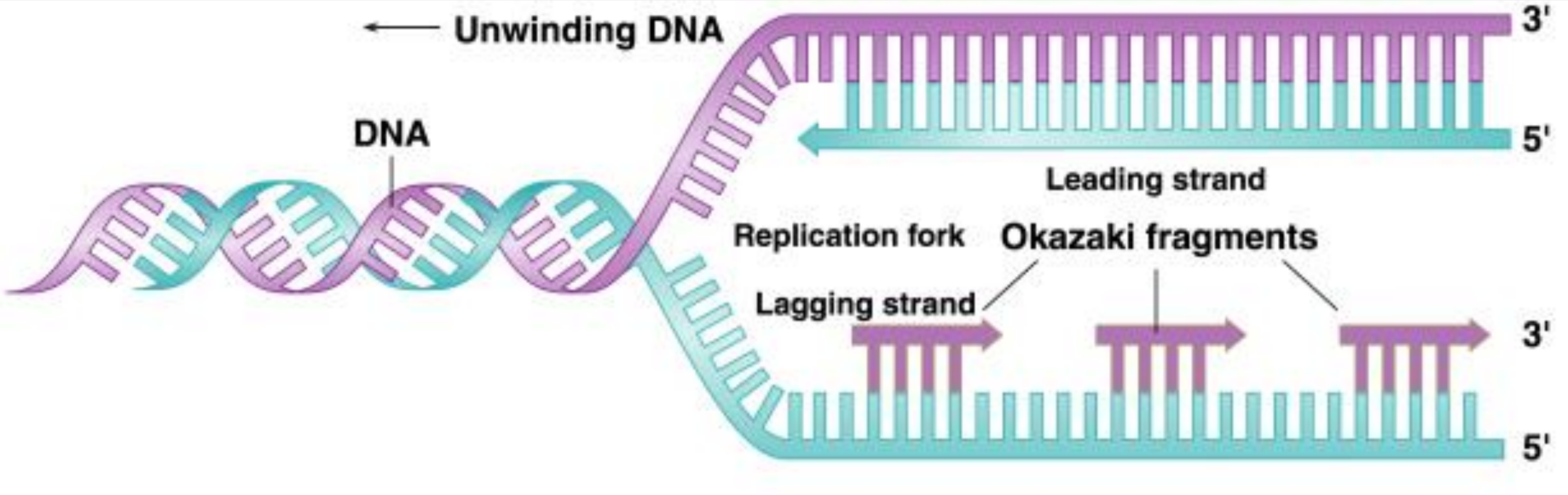


Semi-Conservative DNA Replication

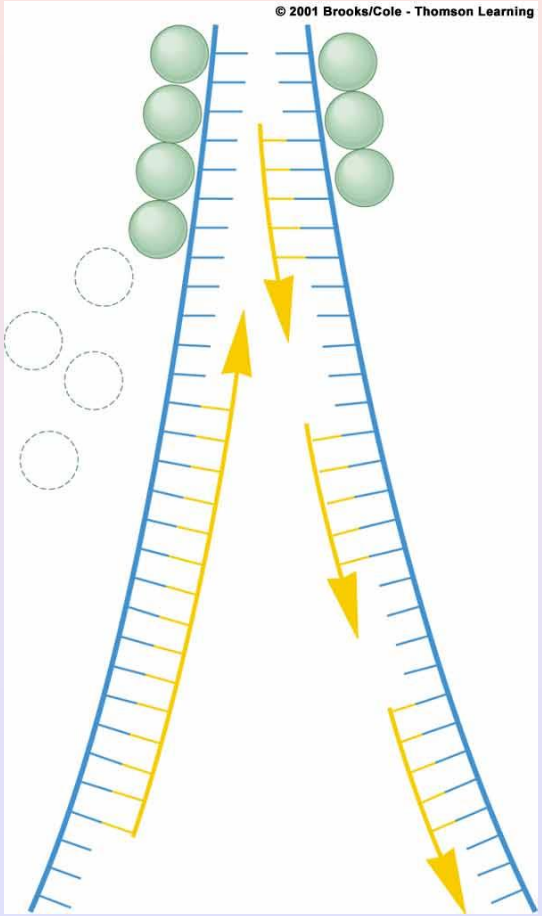


Direction of Replication

- The enzyme *helicase* unwinds several sections of parent DNA
- At each open DNA section, called a **replication fork**, DNA *polymerase* catalyzes the formation of 5'-3'ester bonds of the **leading strand**
- The **lagging strand**, which grows in the 3'-5' direction, is synthesized in short sections called **Okazaki fragments**
- The Okazaki fragments are joined by DNA *ligase* to give a single 3'-5' DNA strand



Strands can **only** be assembled in the 5’ to 3’ direction



Ribonucleic Acid (RNA)

- RNA is much more abundant than DNA
- There are several important differences between RNA and DNA:

  - the pentose sugar in RNA is ribose, in DNA it's deoxyribose
  - in RNA, uracil replaces the base thymine (U pairs with A)
  - RNA is single stranded while DNA is double stranded
  - RNA molecules are much smaller than DNA molecules
- There are three main types of RNA:

  - ribosomal (rRNA), messenger (mRNA) and transfer (tRNA)

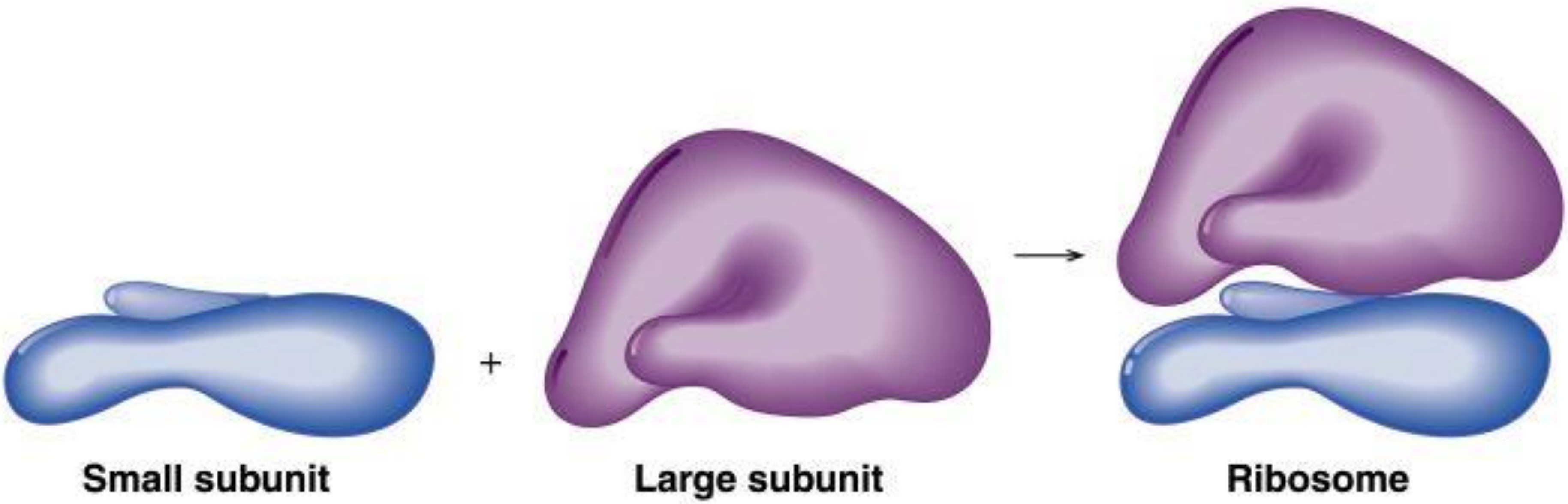
Types of RNA

Table 22.3 Types of RNA Molecules

Type	Abbreviation	Percentage of Total RNA	Function in the Cell
Ribosomal RNA	rRNA	75	Major component of the ribosomes
Messenger RNA	mRNA	5–10	Carries information for protein synthesis from the DNA in the nucleus to the ribosomes
Transfer RNA	tRNA	10–15	Brings amino acids to the ribosomes for protein synthesis

Ribosomal RNA and Messenger RNA

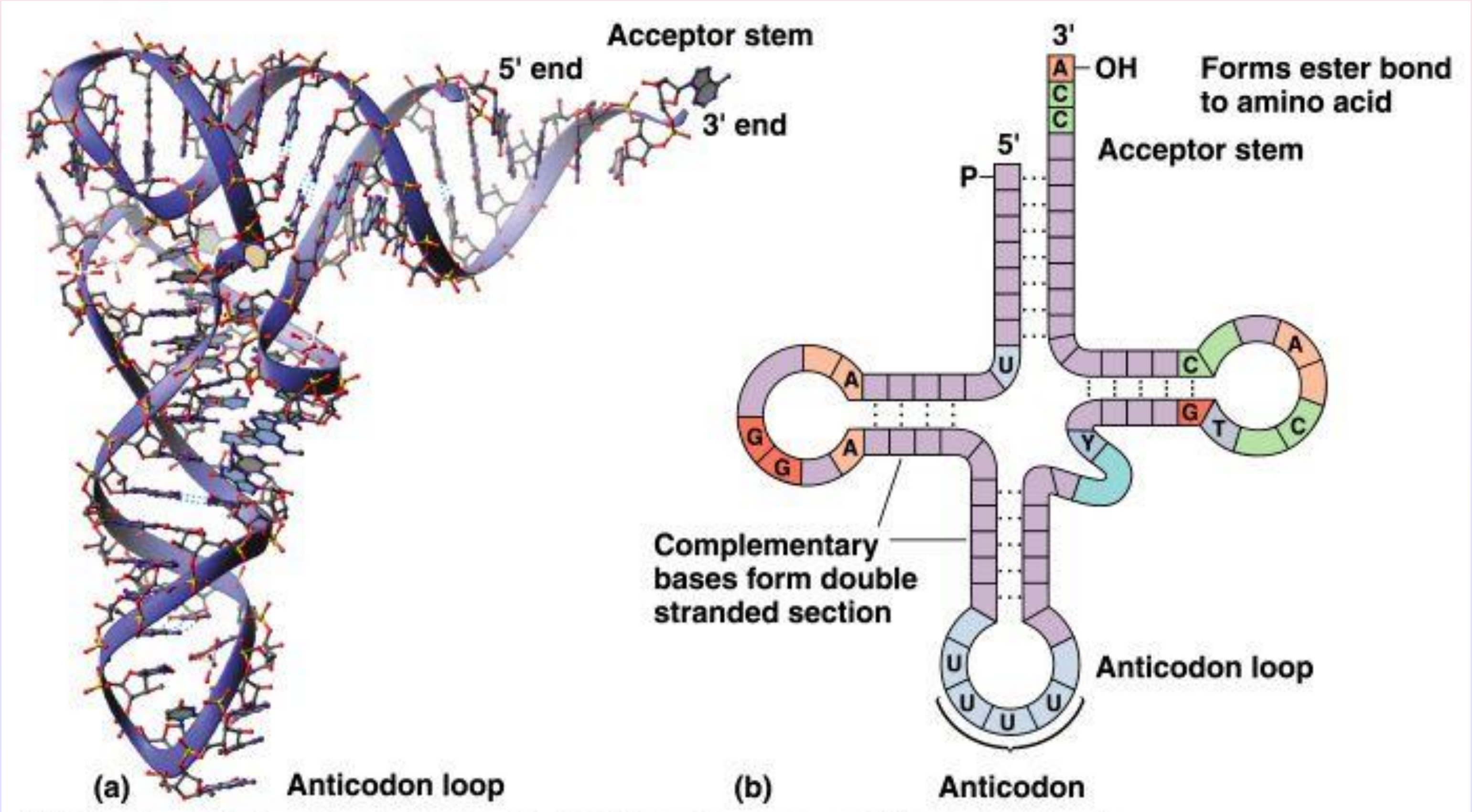
- **Ribosomes** are the sites of protein synthesis
  - they consist of **ribosomal RNA** (65%) and proteins (35%)
  - they have two subunits, a large one and a small one
- **Messenger RNA** carries the genetic code to the ribosomes
  - they are strands of RNA that are complementary to the DNA of the gene for the protein to be synthesized



Transfer RNA

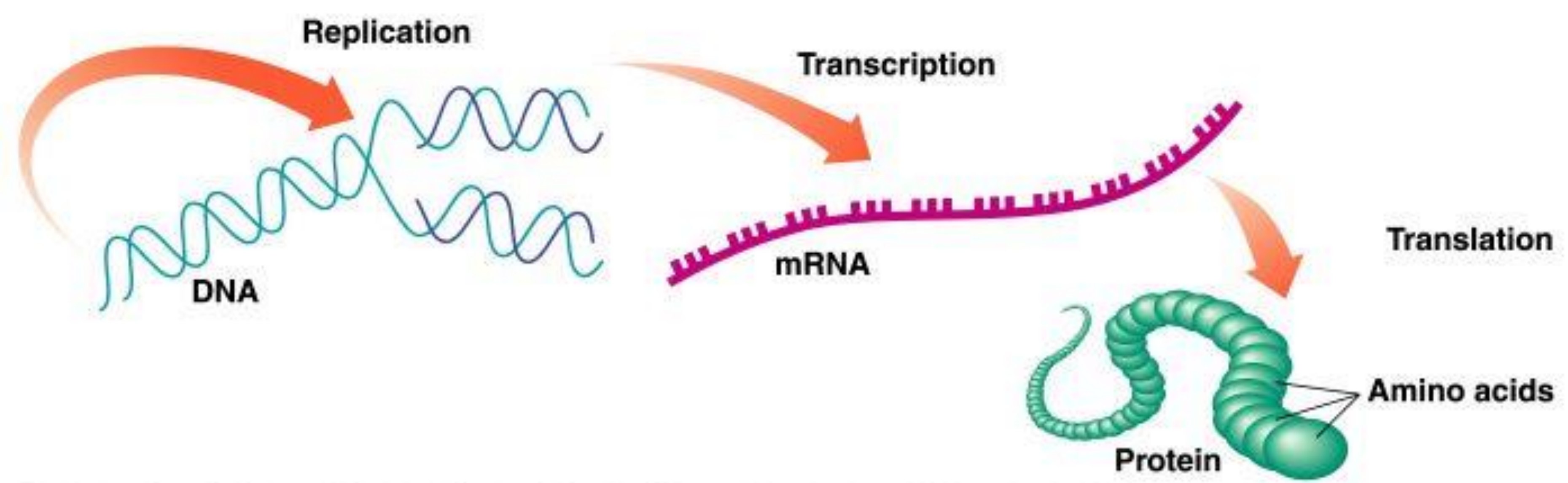
- **Transfer RNA** translates the genetic code from the messenger RNA and brings specific amino acids to the ribosome for protein synthesis
- Each amino acid is recognized by one or more specific tRNA
- tRNA has a tertiary structure that is L-shaped

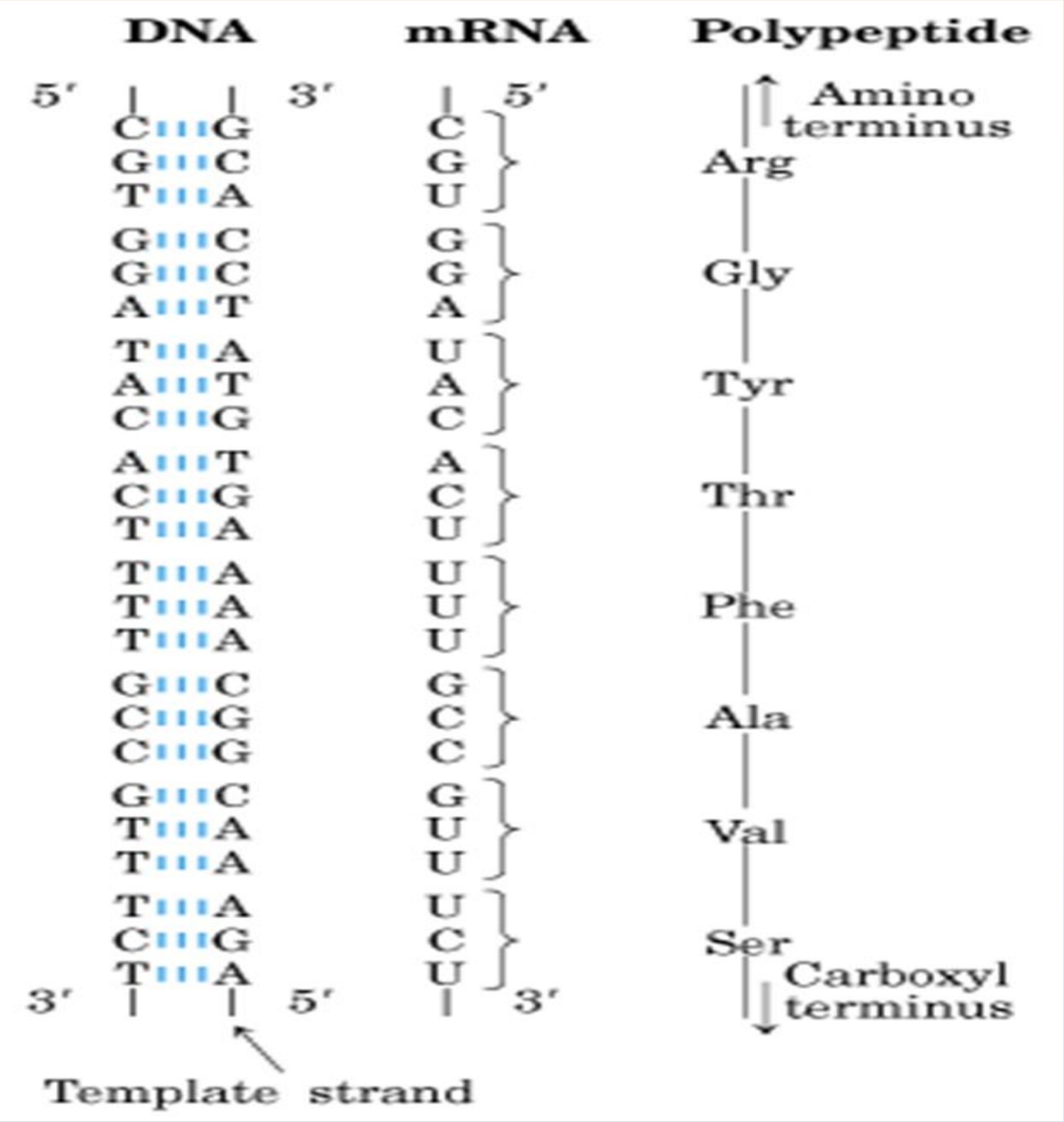
  - one end attaches to the amino acid and the other binds to the mRNA by a 3-base complimentary sequence



Protein Synthesis

- The two main processes involved in **protein synthesis** are
  - the formation of mRNA from DNA **(transcription**)
  - the conversion by tRNA to protein at the ribosome (**translation**)
- Transcription takes place in the nucleus, while translation takes place in the cytoplasm
- Genetic information is transcribed to form mRNA much the same way it is replicated during cell division




DNA
mRNA
Polypeptide
5′
3′
5′
Amino terminus
C G C
G C G
T A U
Arg
G C G
G C G
A T A
Gly
T A U
A T A
C G C
Tyr
A T A
C G C
T A U
Thr
T A U
T A U
T A U
Phe
G C G
C G C
C G C
Ala
G C G
T A U
T A U
Val
T A U
C G C
T A U
Ser
3′
5′
3′
Carboxyl terminus
Template strand

Transcription

- Several steps occur during **transcription**:

  - a section of DNA containing the gene unwinds

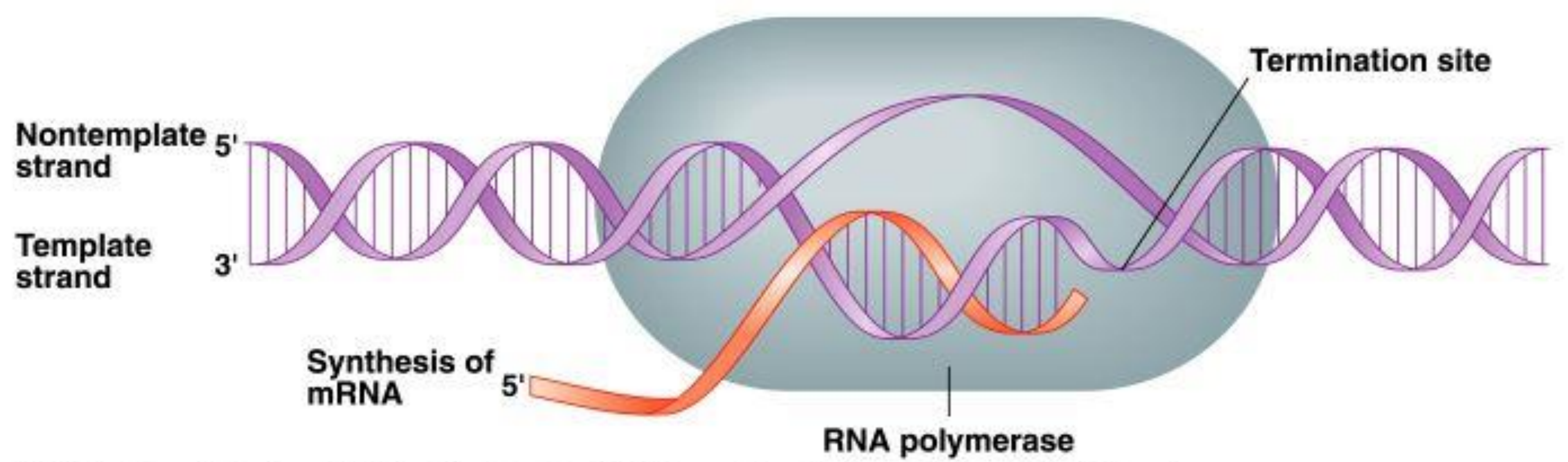
  - one strand of DNA is copied starting at the initiation point, which has the sequence TATAAA

  - an mRNA is synthesized using complementary base pairing with uracil (U) replacing thymine (T)

  - the newly formed mRNA moves out of the nucleus to ribosomes in the cytoplasm and the DNA re-winds

RNA Polymerase

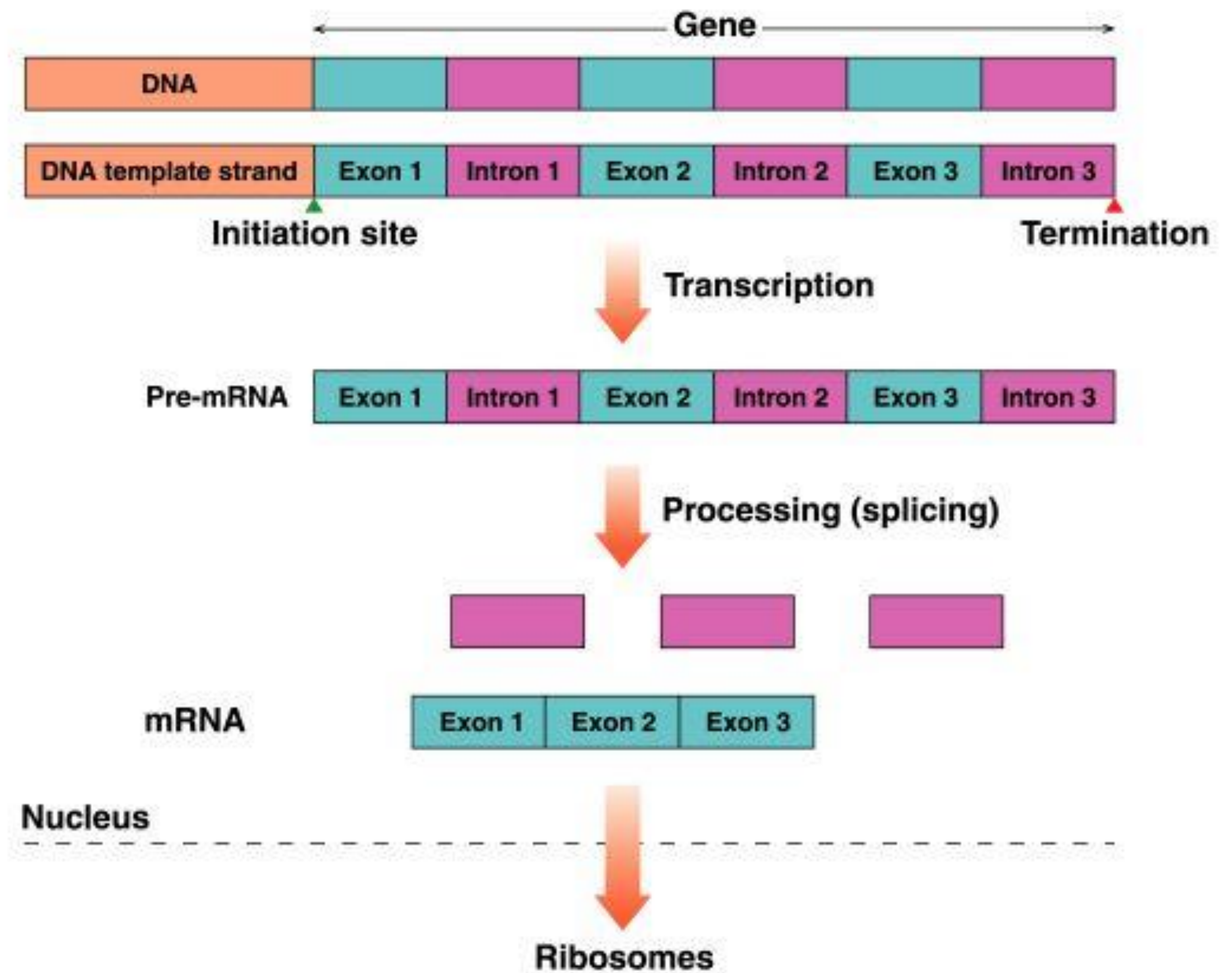
- During transcription, RNA *polymerase* moves along the DNA template in the 3’-5’direction to synthesize the corresponding mRNA
- The mRNA is released at the termination point



Processing of mRNA

- **Genes** in the DNA of eukaryotes contain **exons** that code for proteins along with **introns** that do not
- Because the initial mRNA, called a **pre-RNA**, includes the noncoding introns, it must be processed before it can be read by the tRNA
- While the mRNA is still in the nucleus, the introns are removed from the pre-RNA
- The exons that remain are joined to form the mRNA that leaves the nucleus with the information for the synthesis of protein

Removing Introns from mRNA



Regulation of Transcription

- A specific mRNA is synthesized when the cell requires a particular protein
- The synthesis is regulated at the transcription level:

  - **feedback control**, where the end products speed up or slow the synthesis of mRNA

  - **enzyme induction**, where a high level of a reactant induces the transcription process to provide the necessary enzymes for that reactant
- Regulation of transcription in eukaryotes is complicated and we will not study it here

The Genetic Code

- The **genetic code** is found in the sequence of nucleotides in mRNA that is translated from the DNA
- A **codon** is a **triplet** of bases along the mRNA that codes for a particular amino acid
- Each of the 20 amino acids needed to build a protein has at least 2 codons
- There are also codons that signal the “start” and “end” of a polypeptide chain
- The amino acid sequence of a protein can be determined by reading the triplets in the DNA sequence that are complementary to the codons of the mRNA, or directly from the mRNA sequence
- The entire DNA sequence of several organisms, including humans, have been determined, however,

  - only primary structure can be determined this way

  - doesn’t give tertiary structure or protein function

mRNA Codons and Associated Amino Acids

First Letter	Second Letter: U		C		A		G		Third Letter
U	UUU	Phe	UCU	Ser	UAU	Tyr	UGU	Cys	U
	UUC	Phe	UCC	Ser	UAC	Tyr	UGC	Cys	C
	UUA	Leu	UCA	Ser	UAA	STOP	UGA	STOP	A
	UUG	Leu	UCG	Ser	UAG	STOP	UGG	Trp	G
C	CUU	Leu	CCU	Pro	CAU	His	CGU	Arg	U
	CUC	Leu	CCC	Pro	CAC	His	CGC	Arg	C
	CUA	Leu	CCA	Pro	CAA	Gln	CGA	Arg	A
	CUG	Leu	CCG	Pro	CAG	Gln	CGG	Arg	G
A	AUU	Ile	ACU	Thr	AAU	Asn	AGU	Ser	U
	AUC	Ile	ACC	Thr	AAC	Asn	AGC	Ser	C
	AUA	Ile	ACA	Thr	AAA	Lys	AGA	Arg	A
	[a]AUG	Met/start	ACG	Thr	AAG	Lys	AGG	Arg	G
G	GUU	Val	GCU	Ala	GAU	Asp	GGU	Gly	U
	GUC	Val	GCC	Ala	GAC	Asp	GGC	Gly	C
	GUA	Val	GCA	Ala	GAA	Glu	GGA	Gly	A
	GUG	Val	GCG	Ala	GAG	Glu	GGG	Gly	G

[a] Codon that signals the start of a peptide chain.
STOP codons signal the end of a peptide chain.

Reading the Genetic Code

- Suppose we want to determine the amino acids coded for in the following section of a mRNA

5'—CCU —AGC—GGA—CUU—3'

- According to the genetic code, the amino acids for these codons are:

CCU = Proline	AGC = Serine
GGA = Glycine	CUU = Leucine

- The mRNA section codes for the amino acid sequence of Pro—Ser—Gly—Leu

Three Stages of Translation

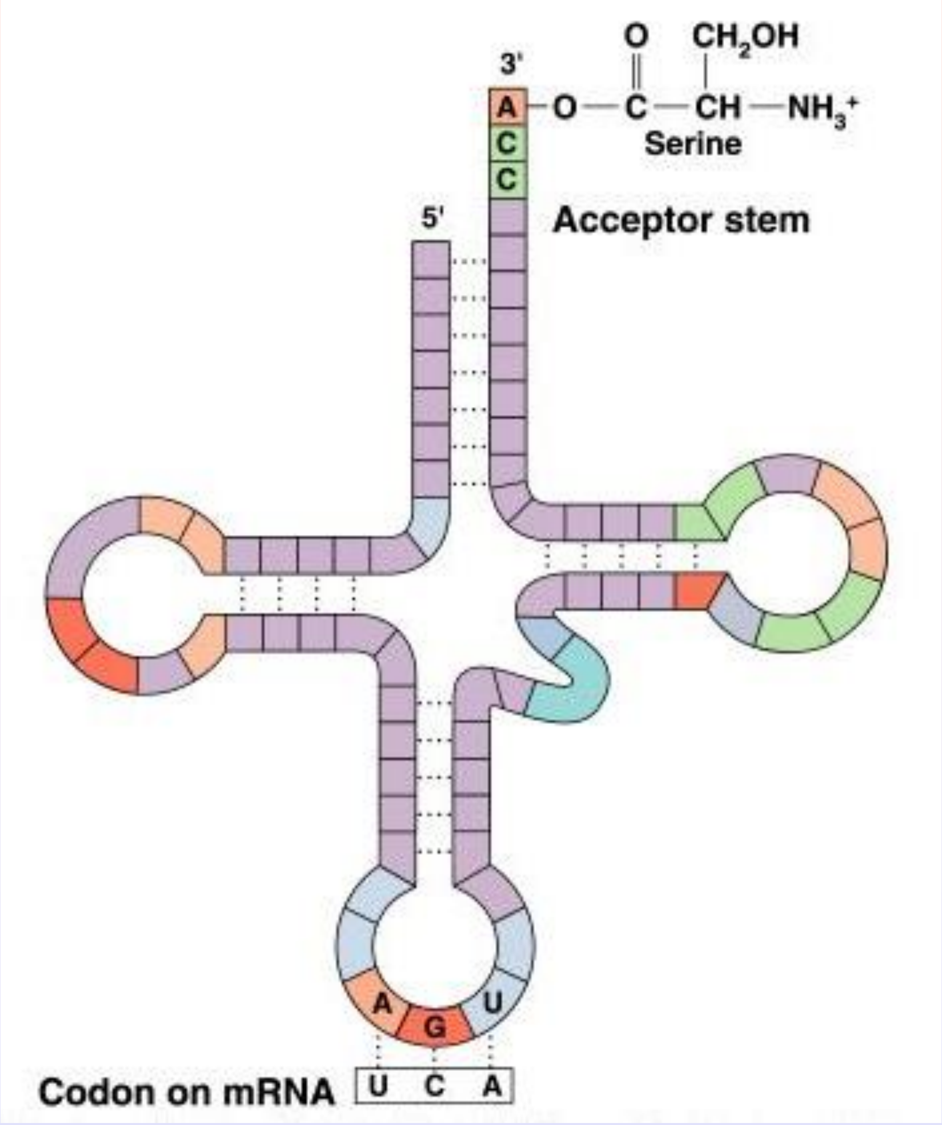
Initiation

Elongation

Termination

Translation and tRNA Activation

- Once the DNA has been transcribed to mRNA, the codons must be tranlated to the amino acid sequence of the protein
- The first step in **translation** is activation of the tRNA
- Each tRNA has a triplet called an **anticodon** that complements a codon on mRNA
- A *synthetase* uses ATP hydrolysis to attach an amino acid to a specific tRNA



Initiation and Translocation

- Initiation of protein synthesis occurs when a mRNA attaches to a ribosome
- On the mRNA, the **start codon (AUG)** binds to a tRNA with methionine
- The second codon attaches to a tRNA with the next amino acid
- A peptide bond forms between the adjacent amino acids at the first and second codons
- The first tRNA detaches from the ribosome and the ribosome shifts to the adjacent codon on the mRNA (this process is called **translocation**)
- A third codon can now attach where the second one was before translocation

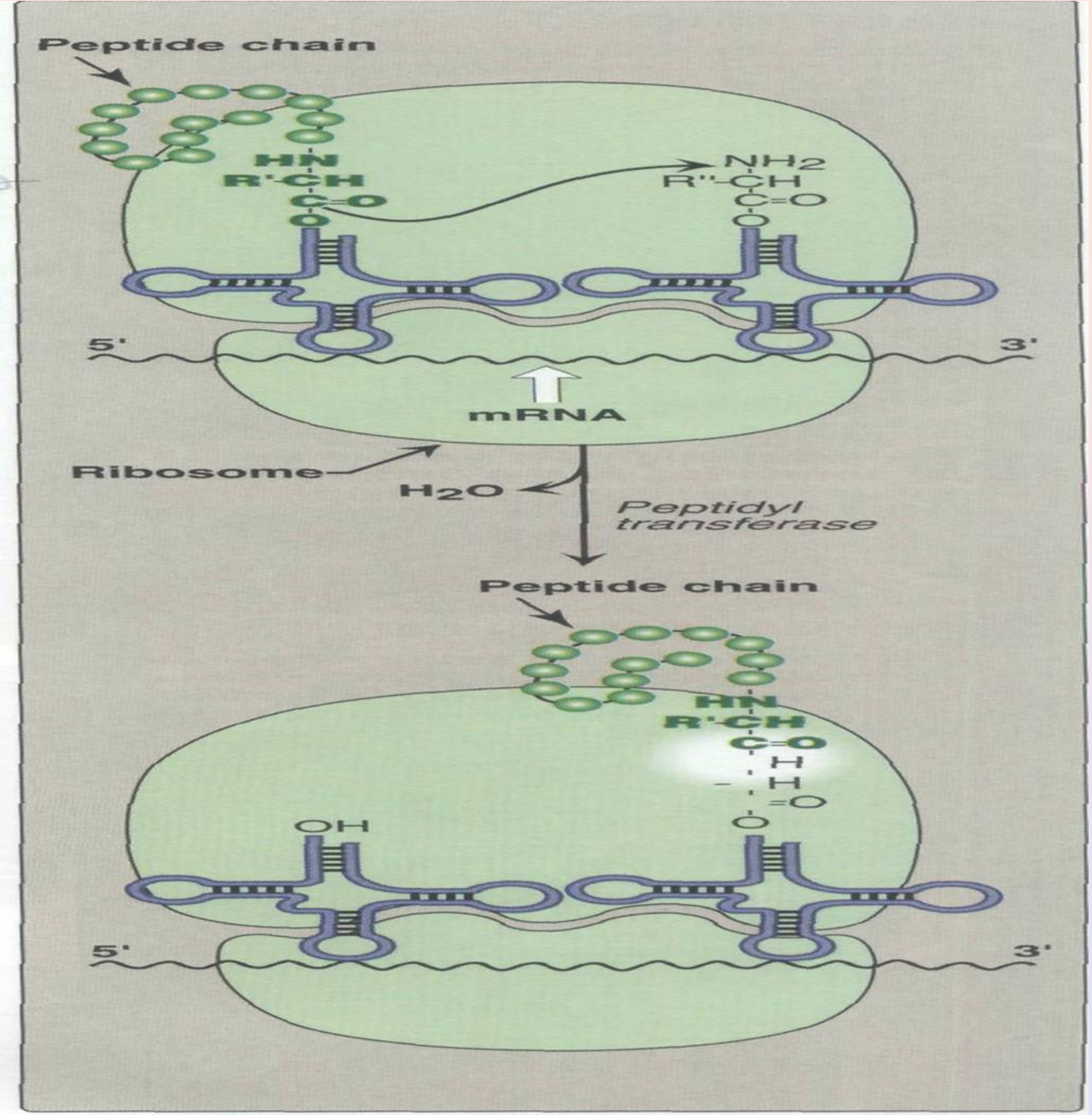


Figure 31.12
Formation of a peptide bond.

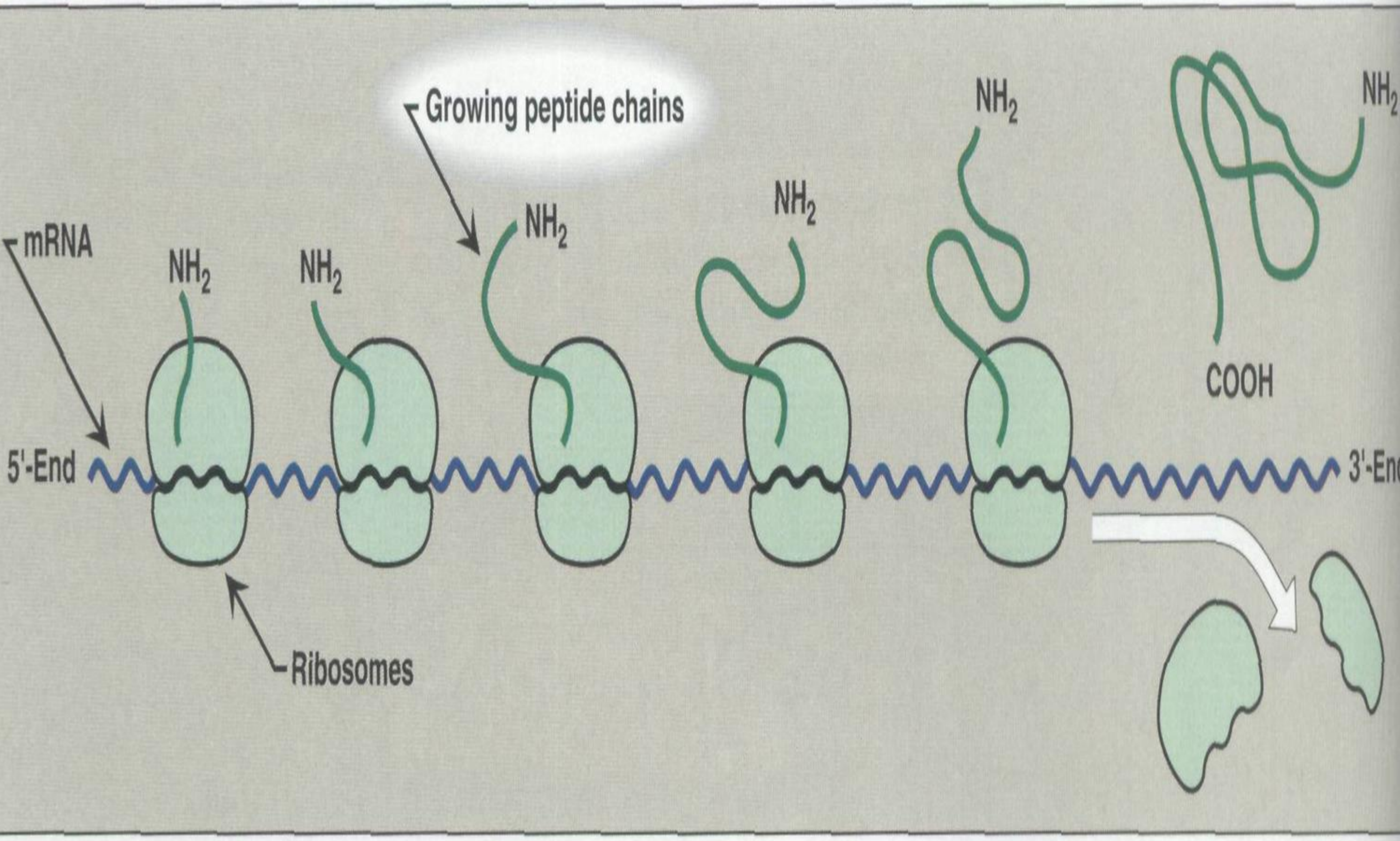


Figure 31.14
A polyribosome consists of several ribosomes simultaneously translating one mRNA.

Termination

- After a polypeptide with all the amino acids for a protein is synthesized, the ribosome reaches the the **“stop” codon**: UGA, UAA, or UAG
- There is no tRNA with an anticodon for the “stop” codons
- Therefore, protein synthesis ends (**termination**)
- The polypeptide is released from the ribosome and the protein can take on it’s 3-D structure

  (some proteins begin folding while still being synthesized, while others do not fold up until after being released from the ribosome)

Transfer RNA (tRNA)

- Transports amino acids to the ribosomes where they are joined together to make proteins
- There is a specific tRNA for each amino acid
- Recognition of the tRNA at the anti-codon communicates which amino acid is attached