قسم التقنيات الاحيائية
المادة:البايولوجبي الجزيئي
المرحلة الثالثة
الفصل الدراسي الاول
ا. د. نهى جوزيف نجيب قندلا
ا.م.د. دلال صالح قادر الربيعي
م.د. ابتهال عبد الهادي مجيد

Lecture 1: :

History of molecular biology

Molecular biology is the study of molecular underpinnings of the process of replication, transcription and translation of the genetic material. The central dogma of molecular biology where genetic material is transcribed into RNA and then translated into protein, despite being an oversimplified picture of molecular biology,

Molecular biology is the study of biology at a molecular level. The field overlaps with other areas of biology and chemistry, particularly genetics and biochemistry. Molecular biology chiefly concerns itself with understanding the interactions between the various systems of a cell, including the interactions between DNA, RNA and protein biosynthesis as well as learning how these interactions are regulated.

The field of molecular biology studies macromolecules and the macromolecular mechanisms found in living things, such as the molecular nature of the gene and its mechanisms of gene replication, mutation, and expression. Given the fundamental importance of these macromolecular mechanisms throughout the history of molecular biology.

The **history of molecular biology** begins in the 1930s with the convergence of various, previously distinct biological disciplines: biochemistry, genetics, microbiology, and virology. With the hope of understanding life at its most fundamental level, numerous physicists and chemists also took an interest in what would become molecular biology.

In its modern sense, molecular biology attempts to explain the phenomena of life starting from the macromolecular properties that generate them. Two categories of macromolecules in particular are the focus of the molecular biologist:

1) nucleic acids, among which the most famous is deoxyribonucleic acid (or DNA), the constituent of genes,

(2) proteins, which are the active agents of living organisms.

The basic concepts of heredity and genes can be traced back to 1865 and the studies of Gregor Mendel . From the results of his breeding experiments with peas, Mendel concluded that each pea plant possessed two **alleles** for each gene, but only displayed a single **phenotype**. Hereditary transmission through sperm and egg became known about the same time and Ernst Haeckel, noting that sperm consists largely of nuclear material, postulated that the nucleus was responsible for heredity. In its earliest manifestations, molecular biology—the name was coined by Warren Weaver of the Rockefeller

Foundation in 1938—was an ideal of physical and chemical explanations of life, rather than a coherent discipline.

Identified DNA as the primary agent of genetic material

In the early 1900's many people thought that **protein** must be the **genetic material** responsible for inherited characteristics. One of the reasons behind this belief was the knowledge that proteins were quite complex molecules and therefore, they must be specified by molecules of equal or greater complexity (i.e. other proteins). DNA was known to be a relatively simple molecule, in comparison to proteins, and therefore it was hard to understand how a complex molecule (a protein) could be determined by a simpler molecule (DNA). What were the key experiments which identified DNA as the primary genetic material?

1928 Frederich . Griffith

Diplococcus pneumoniae,*Streptococcus pneumoniae* or pneumococcus, is a nasty little bacteria which, when injected into mice, will cause pneumonia and **death** in the mouse. The bacteria contains a **capsular polysaccharide** on its surface which protects the bacteria from host defences. Occassionally, variants (mutants) of the bacteria arise which have a defect in the production of the capsular polysaccharide. The mutants have two characteristics: 1) They are **avirulent**, meaning that without proper capsular polysaccharide they are unable to mount an infection in the host (they are destroyed by the host defences), and 2) Due to the lack of capsular polysaccharide the surface of the mutant bacteria appears **rough** under the microscope and can be distinguished from the wild type bacteria (whose surface appears **smooth**).

Wild type
Smooth, virulent

Mutant type
Rough, avirulent

The virulent smooth wild type pneumococcus can be **heat treated** and rendered **avirulent** (still appears smooth under the microscope however).

Controls:

- **Wild type (w.t.) (smooth) + mouse = dead mouse**
- **mutant (rough) + mouse = live mouse**
- **heat treated w.t. (smooth) + mouse = live mouse**

Combinations:

- heat treated w.t. (smooth) + mutant (rough) + mouse = **dead mouse**

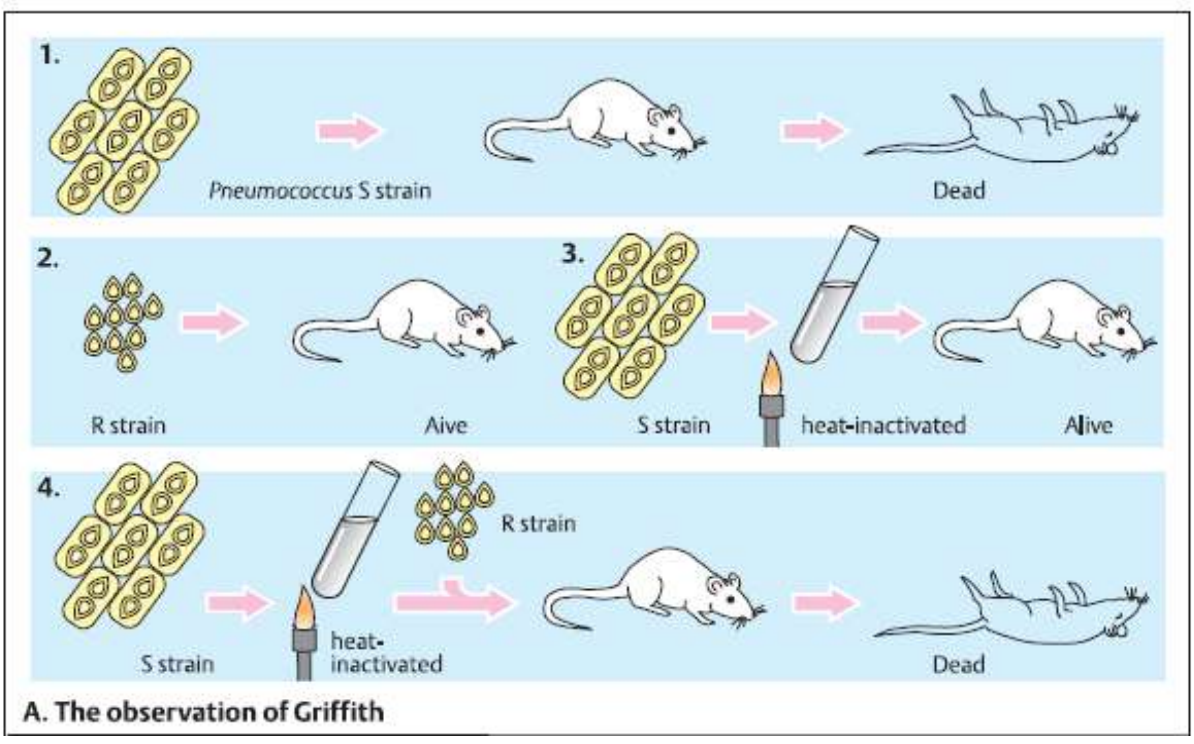
In this case when the bacteria were recovered from the cold lifeless mouse they were smooth virulent pneumococcus (i.e. indistinguishable from wild type).

A closer look at what is going on, by keeping using, and keeping track of, different *subtypes*

- heat treated w.t. (smooth) type I + mutant(rough) type II + mouse = **dead mouse**

*In this case when the bacteria were isolated from the cold lifeless mouse they were smooth virulent **type I** pneumococcus.*

The overall conclusions from these experiments was that there was a "transforming agent" in the the heat treated type I bacteria which transfomed the live mutant (rough) type II bacteria to be able to produce type I capsule polysaccharide.



A. The observation of Griffith

The first major piece of evidence supporting DNA as the genetics material

1944 (Oswaled Avery , Colin Macleod and Maclin MacCarty)

The experiment of Griffith could not be taken further until methods were developed to separate and purify DNA and protein cellular components. Avery and their colleagues utilized methods to extract relatively pure DNA from pneumococcus to determine whether it was the "transforming agent" observed in Griffith's experiments.

The experiment:

- Wild Type (smooth) type I -> extract the DNA component
- Mutant (rough) type II + type I DNA + mouse = dead mouse

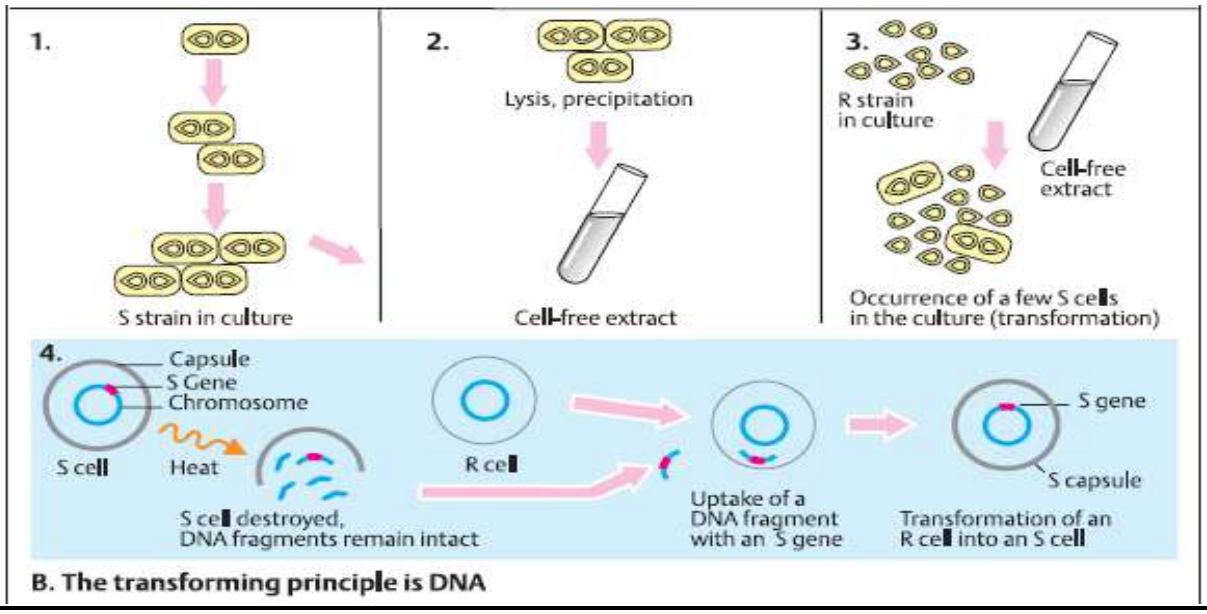
Isolation of bacteria from the dead mouse showed that they were type I . Wild Type *(smooth) bacteria*

Purified type I DNA was divided into two aliquots. One aliquot was treated with **DNAse** - an enzyme which non-specifically degrades DNA. The other aliquot was treated with **Trypsin** - a protease which (relatively) non-specfically degrades proteins.

- Type I DNA + DNAse + mutant (rough) type II + mouse = live mouse
- Type I DNA + Trypsin + mutant (rough) type II + mouse = **dead mouse**

Conclusion:

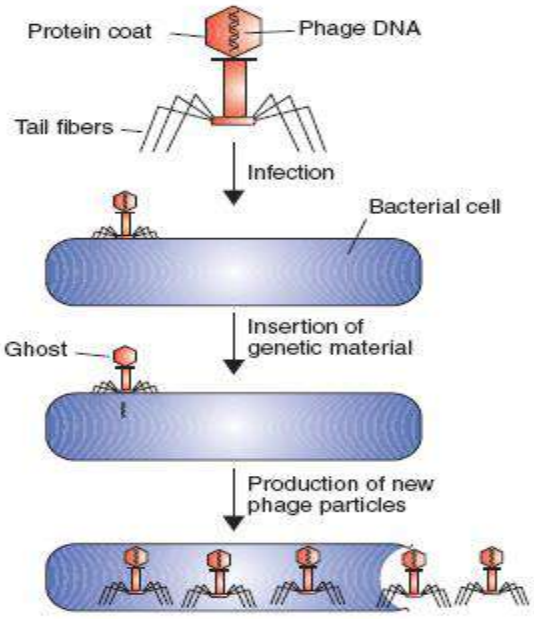
The work of Avery provided strong evidence that the "transforming agent" was in fact DNA (and not protein). However, not everyone was convinced. Some people felt that a residual amount of protein might remain in the purified DNA, even after Trypsin treatment, and could be the "transforming agent".



The second major piece of evidence supporting DNA as the genetic material .

1952 Alfred Hershey and Martha Chase

Hershey and Chase conducted their experiment on the T2 is a virus which attacks the bacteria *E. coli*. The virus, or *phage*, looks like a tiny lunar landing module:



The viral particles adsorb to the surface of the *E. coli* cells. It was known that some material then leaves the phage and enters the cell. The "empty" phage particles on the

surface cells can be physically removed by putting the cells into a blender and whipping them up. In any case, some 20 minutes after the phage adsorb to the surface of the bacteria the bacteria (lysis) and releases a multitude of progeny virus.

1- In a first experiment, they labeled the DNA of the phages with radioactive phosphorus -32 (this element its present in the DNA but not present in the 20 amino acids from which protein are made).they allowed the phage to infect *E.coli* then removed the protein shells from the infected cells with a blender and separated the cells and viral coats by using a centrifuge .they found that radioactive tracer was visible only in the pellet of bacterial cells and not in the supernatant containing the protein shells . Where the ^{32}P label went:

- Adsorbed phage shells 30%
- Infected cells (prior to lysis) 70%

In a second experiment, they labeled the phages with radioactive -35 (sulfur is present in the amino acids cysteine and methionine , but not in DNA).After separation the radioactive tracer then was found in the protein shells , but not in the infected bacteria , supporting the hypothesis that the genetic material which infects the bacteria is DNA

Where the ^{35}S label went:

- Adsorbed phage shells 85%
- Infected cells (prior to lysis) 15%

Conclusion:

The material which was being transfered from the phage to the bacteria during infection appeared to be mainly DNA.

This work , together with that of Avery ,Macleod and McCarty , provided evidence that DNA was the molecule responsible for heredity .

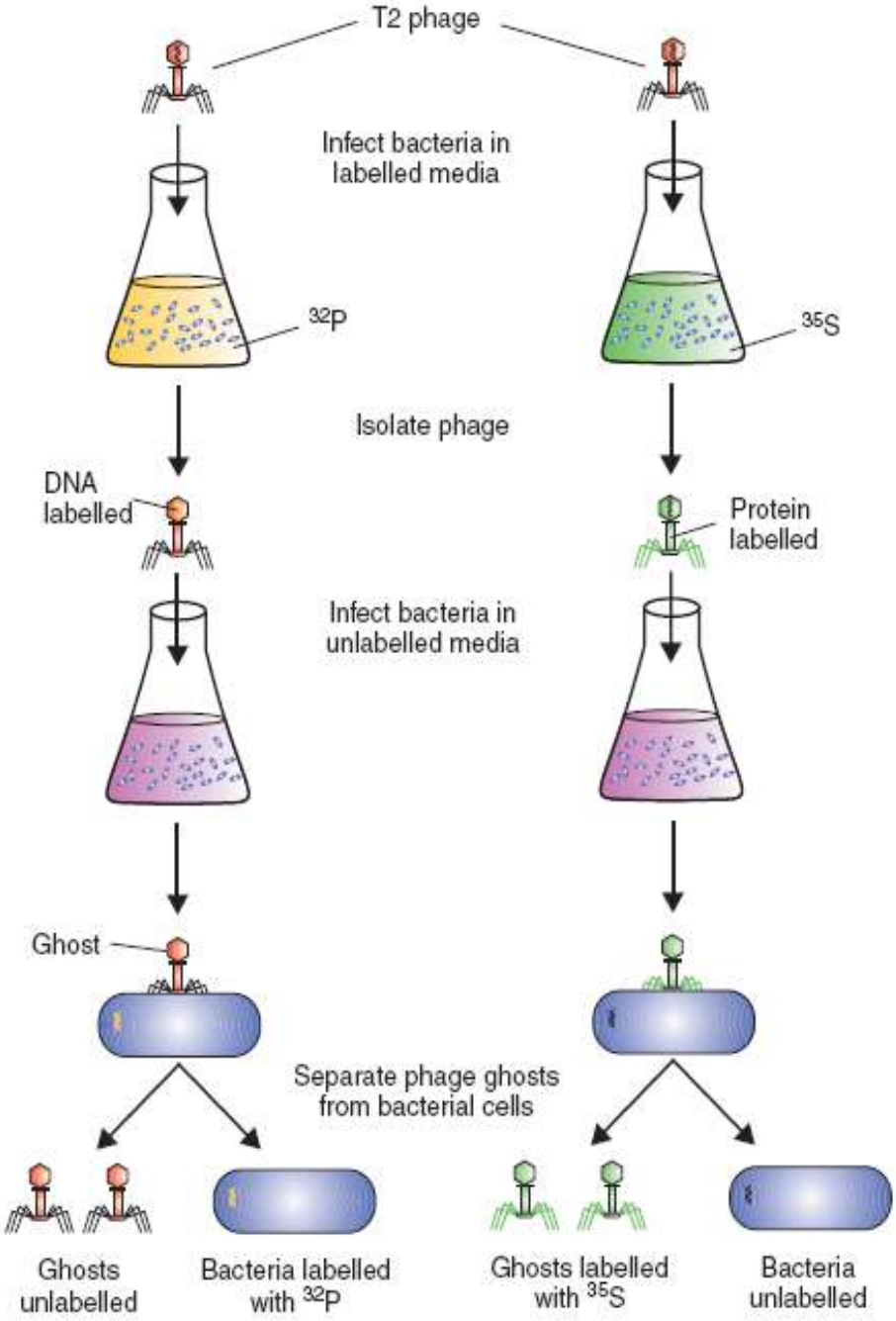
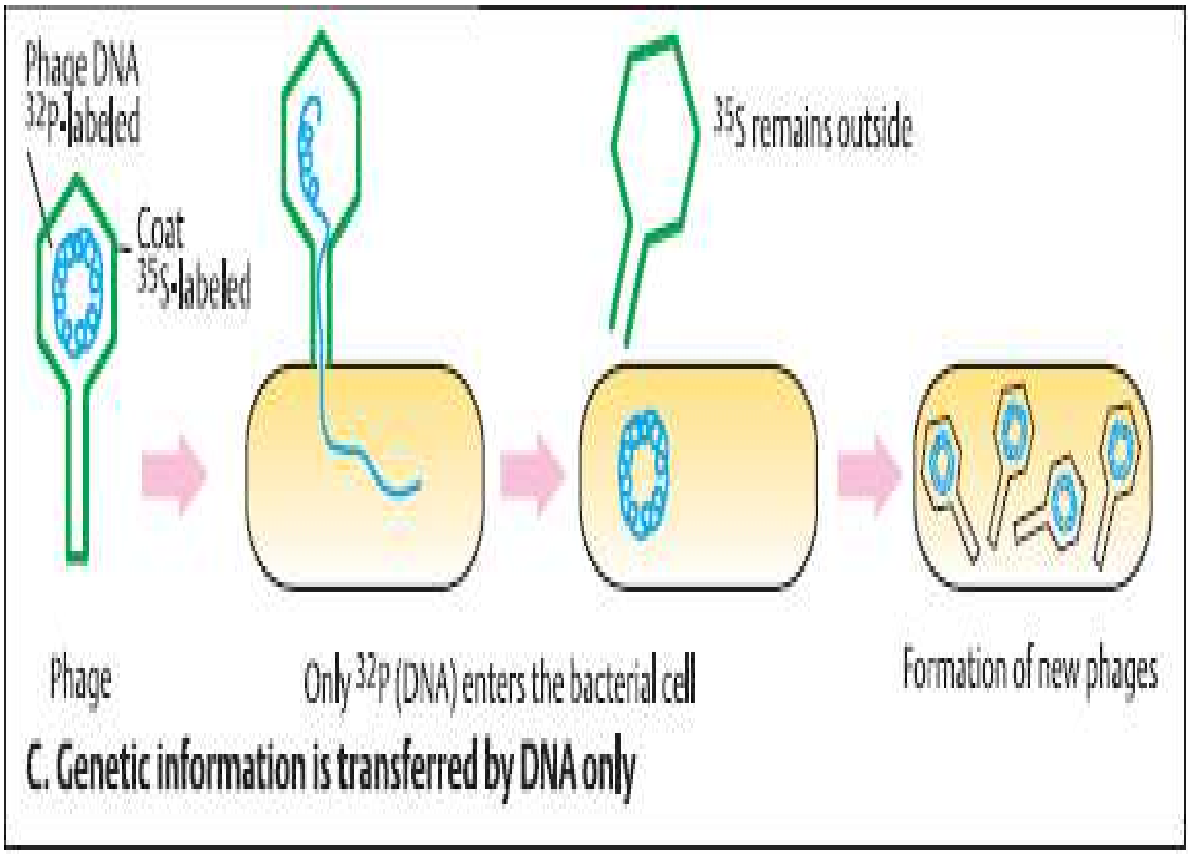


Figure 1.4. The Hershey–Chase blender experiment to show that nucleic acid was the genetic material. Hershey and Chase grew T2 bacteriophages on bacteria whose media contained either ^{32}P (to label the phosphorus of nucleic acid) or ^{35}S (to label the sulphur of proteins – the side chains of the amino acids methionine and cysteine both contain sulphur). They used their radio-labelled bacteriophages to infect a new culture of unlabelled bacteria. After a brief incubation, the bacteria were harvested by centrifugation and put into a blender to shear the bacteria away from the phage particles attached to their surface. They found that, when the DNA was labelled, the label was transferred to the bacterial cell, while the labelled protein remained with the phage ghosts. They concluded, therefore, that the material of heredity – i.e. the material passed on to make new offspring – was nucleic acid

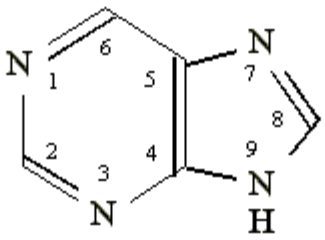


Lecture 2:

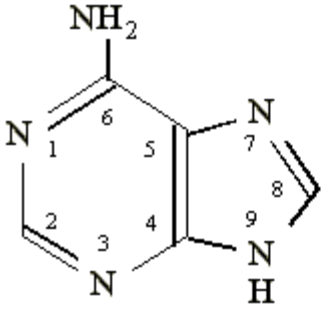
The chemical composition of nucleic acid

The structure of DNA and RNA.

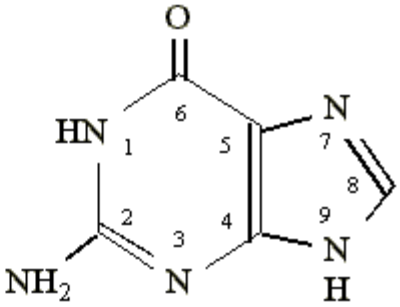
DNA (deoxyribonucleic acid) and RNA (ribonucleic acid) are composed of two different classes of nitrogen containing bases: the **purines** and **pyrimidines**. The most commonly occurring purines in DNA are **adenine** and **guanine**:



Purine

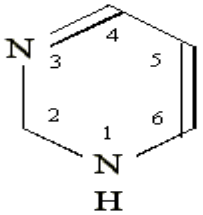


Adenine
135.1 gm/mol

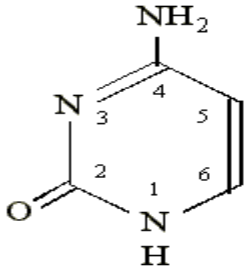


Guanine
151.1 gm/mol

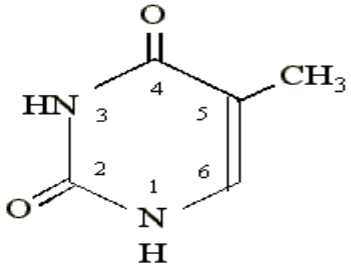
The most commonly occurring pyrimidines in DNA are **cytosine** and **thymine**:



Pyrimidine



Cytosine
111.1 gm/mol

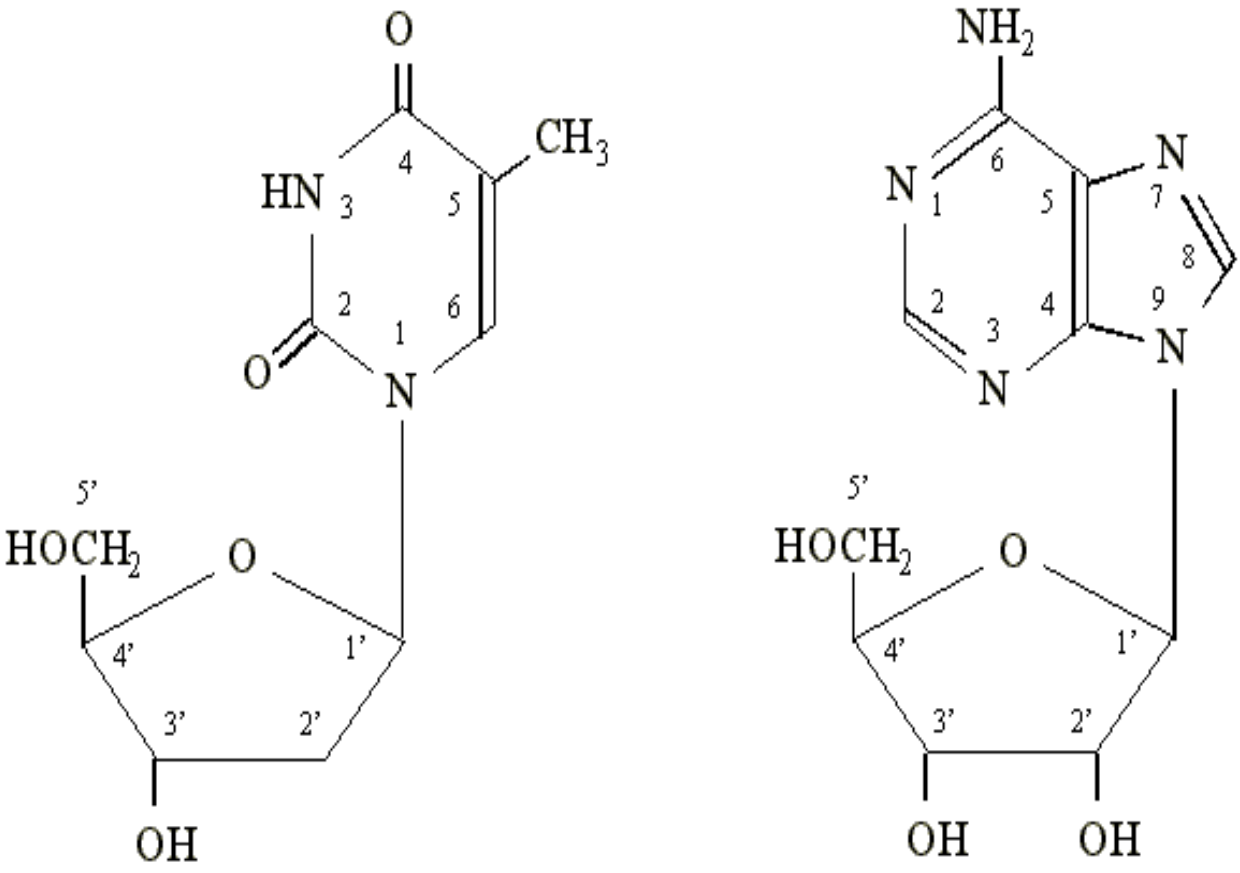


Thymine
126.1 gm/mol

RNA contains the same bases as DNA with the exception of thymine. Instead, RNA contains the pyrimidine **uracil**:

Adenine, guanine, cytosine, thymine and uracil are usually abreviated using the single letter codes A, G, C, T and U, respectively.

Purines and pyrimidines can form chemical linkages with pentose (5-carbon) sugars. The carbon atoms on the sugars are designated 1', 2', 3', 4' and 5'. It is the 1' carbon of the sugar that becomes bonded to the nitrogen atom at **position N1 of a pyrimidine** or **N9 of a purine**. DNA precursors contain the pentose deoxyribose. RNA precursors contain the pentose ribose (which contains an additional OH group at the 2' position): The resulting molecules are called **nucleoside s** and can serve as elementary precursors for DNA and RNA synthesis, *in vivo*.

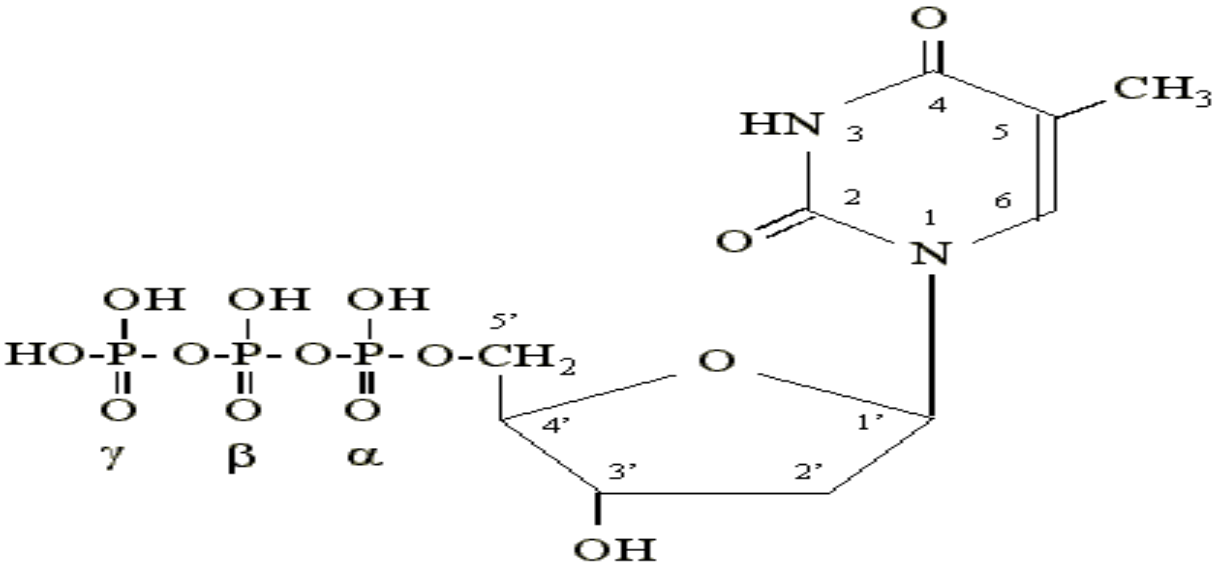


2'-deoxy Thymidine
(2'-deoxy ribose sugar
DNA precursor)

Adenosine
(ribose sugar
RNA precursor)

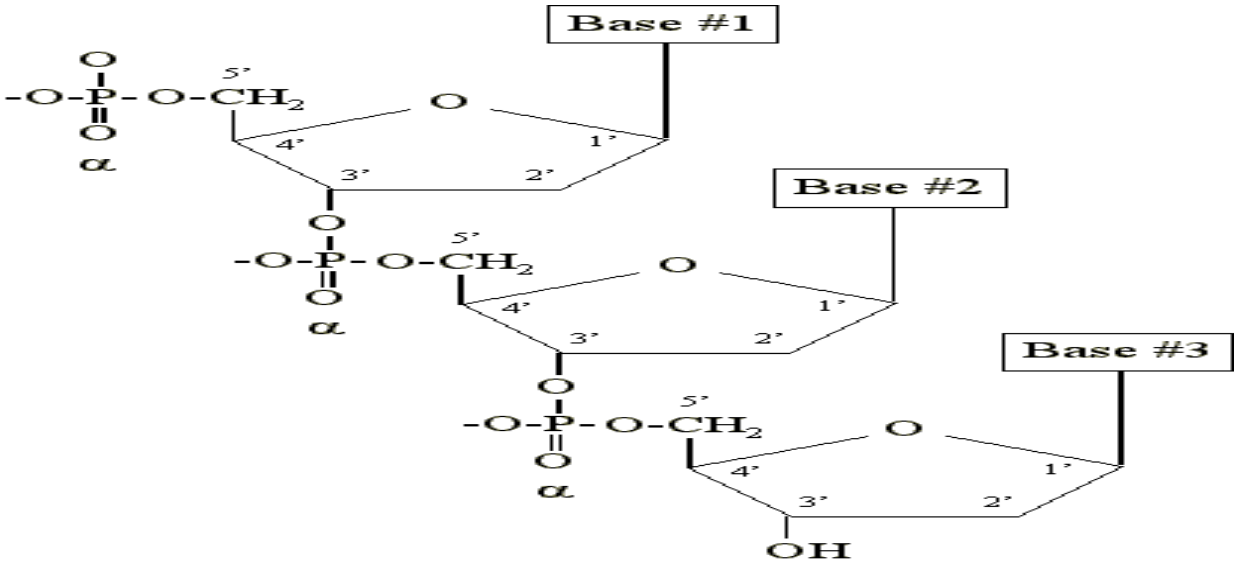
Before a nucleoside can become part of a DNA or RNA molecule it must become complexed with a phosphate group to form a **nucleotide** (either a deoxyribonucleotide or ribonucleotide). Nucleotides can posess **1, 2 or 3 phosphate groups**, e.g. the nucleotides adenosine monophosphate (AMP), adenoside diphosphate (ADP) and adenosine triphosphate (ATP). The phosphate groups are attached to the **5' carbon** of the ribose

sugar moiety. Beginning with the phosphate group attached to the 5' ribose carbon, they are labeled □, □ **and** □ **phosphate**. It is the **tri-phosphate nucleotide** which is incorporated into DNA or RNA.



2'-deoxy Thymidine triphosphate (nucleotide)

DNA and RNA are simply long polymers of nucleotides calle **polynucleotides**. Only the □ **phosphate** is included in the polymer. It becomes chemically bonded to the **3' carbon** of the sugar moiety of another nucleotide:



In other words, the polynucleotide is connected by a series of 5' to 3' phosphate linkages. Note the sequence of the bases in the above diagram. **Polynucleotide sequences are referenced in the 5' to 3' direction**. Typically, polynucleotides will contain a 5' phosphate and 3' hydroxyl terminal groups. The common representation of polynucleotides is as an **arrow** with the **5' end at the left and the 3' end at the right**.

Summary of terms:

Base	Nucleoside	Nucleotide	RNA (monophosphate)	DNA (monophosphate)	Code
Adenine	Adenosine	(Adenylic acid)	AMP	dAMP	A
Guanine	Guanosine	(Guanylic acid)	GMP	dGMP	G
Cytosine	Cytidine	(Cytidylic acid)	CMP	dCMP	C
Thymine	Thymidine	(Thymidylic acid)		dTMP	T
Uracil	Uridine	(Uridylic acid)	UMP		U

What is the structure of DNA? How is the structure related function?

1950's

The primary chemical structure of polynucleotides was known (i.e. the 3'-5' phosphate linkage).

1951 E. Chargaff

The experiment:

Take DNA from a variety of species and hydrolyze it to yield individual pyrimidines and purines. Determine the relative concentrations of the A, T, C and G bases.

Result:

Although different species had uniquely different ratios of pyrimidines or purines, the relative concentrations of adenine always equaled that of thymine, and guanine equaled cytosine. His experiments showed that :

1- the amount of adenine residues is proportional to the amount of thymine residues in DNA (column 1,2,and 5).Also , the amount of Guanine residues is proportional to the amount of cytosine residues in DNA (column 3,4,and 6).
2- Based on the proportionality the sum of Purines (A+G) equals the sum of the Pyrimidines (C+T) as shown in column 7.
3- Te percentage of (G+C) does not necessarily equal the percentage of (A+T) as shown in column 8. (Table -2)

Chargaff's Law: A=T, G=C

1950's R.E. Franklin

X-ray diffraction studies of DNA fibers demonstrated that DNA adopted a highly ordered helical structure. Franklin concluded that two or more chains must coil around each other to form a helix. Some basic dimensions of the helix were calculated from the x-ray diffraction data.

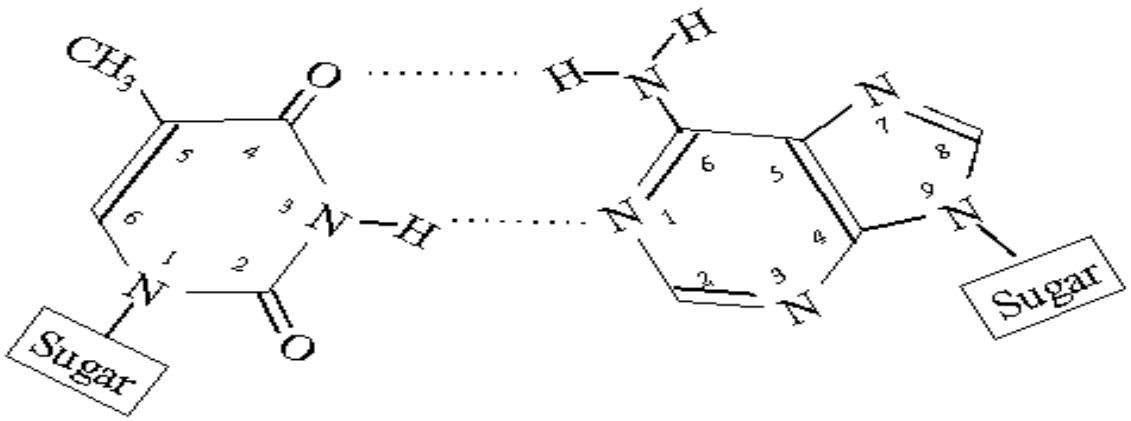
1953 L. Pauling and R.B. Corey

Propose a three chain helical structure for DNA with the phosphate backbone in the center and the bases on the outside.

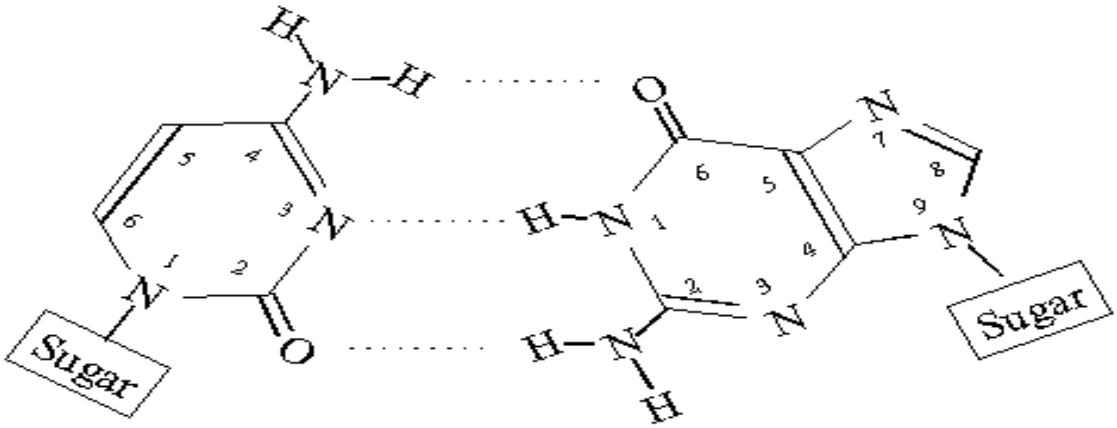
Lecture 3:

1953 J.D. Watson and F.H.C. Crick

Identified a hydrogen bonding arrangement between models of thymine and adenine bases, and between cytosine and guanine bases which fullfilled Chargaff's rule:



Thymine (T) Adenine (A)



Cytosine (C) Guanine (G)

Note that the "TA" pair can overlay the "GC" pair with the bonds to the sugar groups in similar juxtaposition. In the "double helix" model of Watson and Crick the polynucleotide chains interact to form a double helix with the chains running in **opposite directions**. The bases are directed towards the center (and stack on top of one another) and the sugar backbones face the outside of the helix.

The Watson and Crick model had the following physical dimensions:

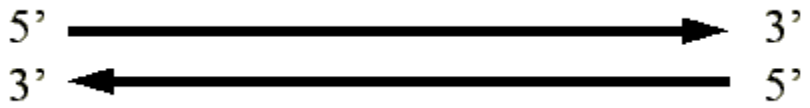
- 34 Å per helical repeat
- 10 base pairs per repeat (i.e. per turn of the helix)
- 3.4 Å inter-base stacking distance
- 20 Å diameter for the helical width

Physical characteristics of the model matched those determined by Rosalind Franklin's x-ray diffraction studies.

Consequenses of the model for genetic information:

The Watson and Crick paper was an exercise in brevity If G always paired with C, and T always paired with A, then either strand could be regenerated from the **complementary information** in the other strand.

1. The basis of the complementarity was hydrogen bonding, i.e. non-covalent interactions which could be easily broken and re-formed.
2. The information which DNA carried was within the unique **base sequence** of the DNA.
3. From the general interior location of the bases, it would appear that the double helix would **have to dissociate** in order to access the information.
4. The non-equatorial location of the sugar moieties suggested that the DNA helix would have a **major groove** and a **minor groove**.



Base Pairs and Stacking

The bases of both DNA chains are flat structures that lie approximately perpendicular to the helical axis.

1- The bases themselves are stacked upon each other.
2- The arrangement is best the base pairs are not all perpendicular to the helical axis, and that some show **propeller twist**, where the purine and pyrimidine pair do not lie flat but are twisted with respect to each other.
3- (The pairing of a purine (A or G) with a pyrimidine (T or C) within the helix is important for the integrity of the helix.
4- The constant length of the purine–pyrimidine pairing would be disrupted if purine–purine (too large) or pyrimidine–pyrimidine (too small) pairings occurred. The purine–pyrimidine pairs are said to complement each other,

5- the two strands of a single DNA molecule are thus **complementary** to one another. Thus, if the sequence 5_-ATGATCAGTACG-3_ occurs on one strand of the DNA, the other strand must have the sequence 5_-CGTACTGATCAT-3_

6- .6-Hydrophobic interactions strengthen double helix.

Alternative DNA Structures

A. Three forms of DNA

The DNA double helix does not occur as a single structure, but rather plan represents a structural family of different types.

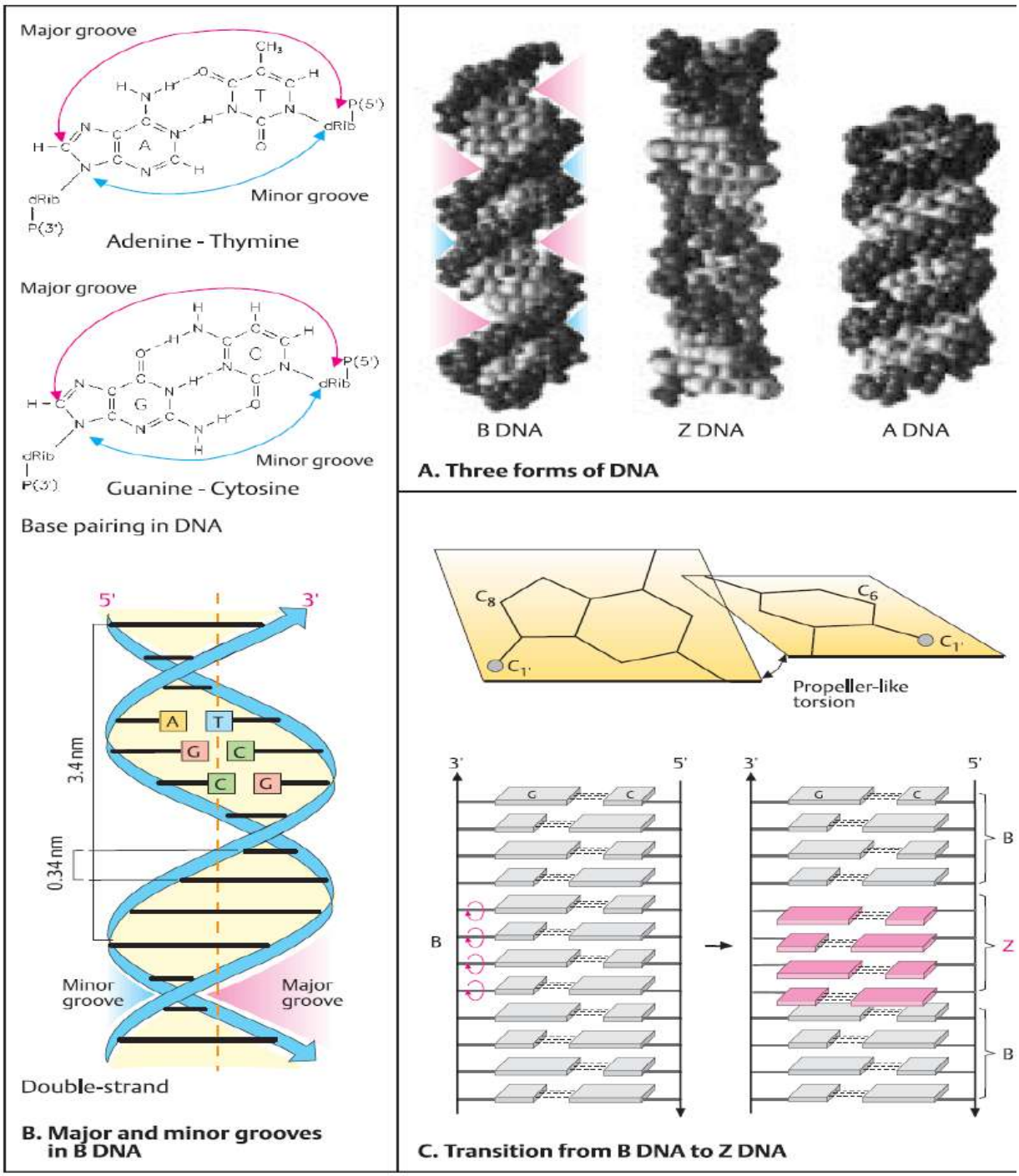
1-The original classic form, determined by Watson and Crick in 1953, is B-DNA. The essential structural characteristic of B-DNA is the formation of two grooves, one large (major groove) and one small (minor groove). the B-DNA forms a right-handed helix, (table -)

2-There are at least two further, alternative forms of the DNA double helix, Z-DNA and the rare form A-DNA. Z-DNA shows a left-handed conformation. This leads to a greater distance (0.77 nm) between the base pairs than in B-DNA and a zigzag form (thus the designation Z-DNA) The Left handed conformation Z discovered by Andrew Wang and Alexander Rich in 1979 saw the sugar phosphate back bone follow a zig - zag pattern around the axis .

3- A-DNA is rare (left-handed conformation). It exists only in the dehydrated state and differs from the B form by a 20-degree rotation of the perpendicular axis of the helix. A-DNA has a deep major groove and a flat minor groove

B- Major and minor grooves in B- DNA

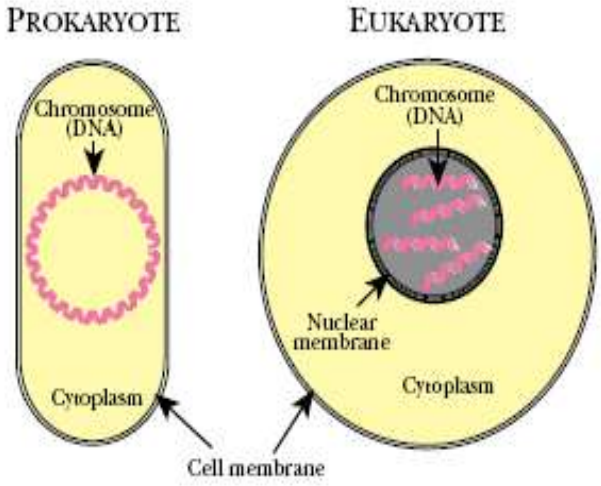
The base pairing in DNA (adenine–thymine and guanine–cytosine) leads to the formation of a large and a small groove because the glycosidic bonds to deoxyribose (dRib) are not **diametrically opposed**. In B-DNA, the purine and pyrimidine rings lie 0.34 nm apart. DNA has ten base pairs per turn of the double helix. The distance from one complete turn to the next is 3.4 nm. In this way, localized curves arise in the double helix. The result is a somewhat larger and a somewhat smaller groove..



Compare between prokaryotic and eukaryotic cells

Based on differences in compartmentalization, living cells may be divided into two types, the simpler **prokaryotic** cell and the more complex **eukaryotic** cell. By definition,

prokaryotes are those organisms whose cells are not subdivided by membranes into a separate **nucleus** and cytoplasm. All prokaryote cell components are located together in the same compartment. In contrast, the larger and more complicated cells of higher organisms (animals, fungi, plants and protists) are subdivided into separate compartments and are called eukaryotic cell.(Figure-1)compares the design of prokaryotic and eukaryotic cells



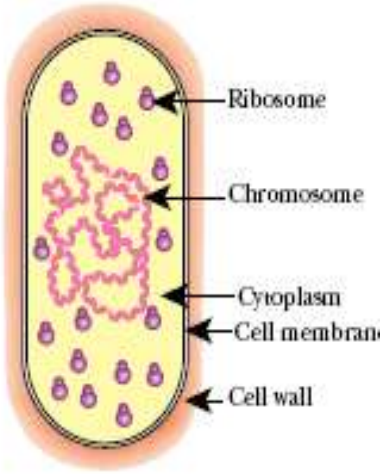
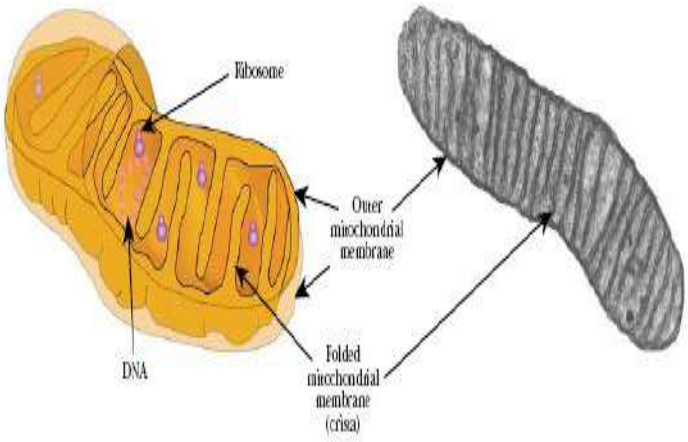
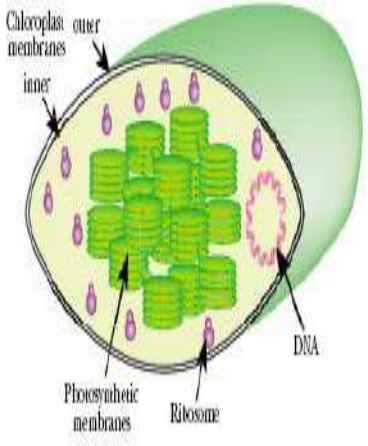


FIGURE 2.07 ***Typical Bacterium***

.





Major differences between prokaryotic and eukaryotic cells	
Prokaryotic cells	**Eukaryotic cells**
No nuclear membrane: chromosome(s) in direct contact with cytoplasm (organized at the nucleoid)	Chromosomes are enclosed in a double layered nuclear membrane (organized in the nucleus)
Simple chromosome structure single **chromosome** carrying a full set of genes closed circular	Complex chromosome structure; DNA associated with histone proteins eukaryotic cells often have multiple copies of certain genes as the result of gene duplication. linear
Cell division does not involve meiosis(replicate by binary fission)	Cell division involves mitosis and meiosis
If present, cell walls contain *peptidoglycan*, no cellulose or chitin	If present, cell walls contain *cellulose* or *chitin*, never peptidoglycan
No mitochondria or chloroplasts , but have DNA in plasmid	Mitochondria usually present, chloroplasts in photosynthetic cells each contain a circular molecule of DNA. The mitochondrial genome is similar to a bacterial chromosome, though much smaller.The mitochondrial DNA has some genes needed for mitochondrial function.
Cells contain ribosomes of only one size	Cells contains two types of ribosomes, one in cytoplasm, and smaller type in mitochondria

No introns present	introns present
Typically, bacteria have 3,000–4,000 genes, although some have as few as 500.The minimum number of genes to allow the survival of a living cell	The genome of eukaryotes consists of 10,000–50,000 genes carried on several chromosomes.(100.000 genes carried in total of human chromosomes)
About 1 mm in length	Tens or hundreds of millimeters in length
Replicate by semiconservative method Rolling circle method	Replicate by semiconservative method No evidence of Rolling circle method
Single molecule of DNA per genetic trait (haploid)	two molecule of DNA per genetic trait (diploid)

DNA and RNA Molecules in different types of cells

	chromosome of all eukaryotic organisms	long ,linear double helixes (double stranded DNA)
	chromosomes of prokaryotic ex. bacteria	covalently closed circular, double stranded DNA

organelles in eukaryotic organisms ex. mitochondria ex .chloroplast	covalently closed circular, double stranded DNA
papovaviruses (cause cancers in animals and humans)	covalently closed circular, double stranded DNA
bacteriophages øx 174 and M 13	single – stranded DNA (covalently closed circular)
mammalian parvovirus's (associated with fetal death and spontaneous abortion in human)	(linear single stranded DNA
Reovirus	double stranded RNA
Picorna virus	single – stranded RNA

Terms :

eukaryote Higher organism with advanced cells, which have more than one chromosome within a compartment called the nucleus

prokaryote Lower organism, such as a bacterium, with a primitive type of cell containing a single chromosome and having no nucleus

nucleus An internal compartment surrounded by the nuclear membrane and containing the chromosomes. Only the cells of higher organisms

have nuclei.

cytoplasm The portion of a cell that is inside the cell membrane but outside the nucleus

membrane A thin flexible structural layer made of protein and phospholipid that is found surrounding all living cells

macromolecule Large polymeric molecule; in living cells especially DNA, RNA, protein or polysaccharide

metabolism The processes by which nutrient molecules are transported and transformed within the cell to release energy and to provide new cell

materi**gene** A unit of genetic information

genome The entire genetic information from an individual

chromosome Structure containing the genes of a cell and made of a single molecule of DNA

diploid Possessing two copies of each gene

haploid Possessing only a single copy of each gene

deoxyribonucleic acid (DNA) The nucleic acid polymer of which the genes are made

nucleic acid Polymer made of nucleotides that carries genetic information

ribonucleic acid (RNA) Nucleic acid that differs from DNA in having ribose in place of deoxyribose and having uracil in place of thymine

replication Duplication of DNA prior to cell division

ribosome The cell's machinery for making proteins

messenger RNA (mRNA) The class of RNA molecule that carries genetic information from the genes to the rest of the cell

transcription Process by which information from DNA is converted into its RNA equivalent

translation Making a protein using the information provided by messenger RNA

Escherichia coli A bacterium commonly used in molecular biology

LECTURE -4

Structure of DNA in the Cell

Different types of nucleic acid are used to form the genome of organisms depending on the organism itself. For example,:

1- viruses have a genome composed of double-stranded DNA, single-stranded DNA or RNA, depending on the type of virus.
2- The chromosomes of most **eukaryotes** are composed of single linear double-stranded DNA molecules.

3- The genomes of **prokaryotic** organisms are generally composed of a circular DNA molecule. That is, rather than having free 5_- and 3_- ends, the ends are joined to each other to form a continuous ring of double-stranded DNA.
4- A number of extra-chromosomal DNA molecules, called **plasmids**, are found in prokaryotic cells. understanding how cells deal with these plasmids has played a pivotal role in advances in molecular biology and genetic engineering.

Plasmid DNA molecules are usually closed circles of either single-stranded or double-stranded DNA. Electron microscopy images of DNA molecules indicate that DNA is more string-like than rod-like, and will wrap around itself to form a variety of irregular structures. Inside a eukaryotic cell, DNA is associated with a vast array of proteins – for example,

A- proteins required for its replication.
B- its transcription into RNA .
C- its packaging within the cell.

Many of these proteins wrap DNA around themselves, or in other ways constrain or bend the DNA molecule. For short linear DNA molecules, this type of constraint is not a major problem. Stresses placed in one part of a DNA molecule by, for example, twisting the double helix can be relieved by untwisting another part of the DNA. The free ends of the linear DNA allow relatively free rotation of the DNA strands. For circular DNA molecules, however, these stresses can prove extremely problematic. Since the DNA has no free ends, it cannot simply untwist to counteract a twist elsewhere in the molecule. Twisting DNA molecules results in the formation of DNA **supercoils** (Figure 1.13).

In 1965 Jerome Vinograd and his colleagues suggested that closed-circular DNA molecules could adopt a '**twisted circular form'**. Such a form would result if, before joining the ends of a linear duplex DNA into a closed circle, one end was twisted relative to the other to introduce some strain into the molecule. Such coiling of the DNA helix upon itself is called **supercoiling**.

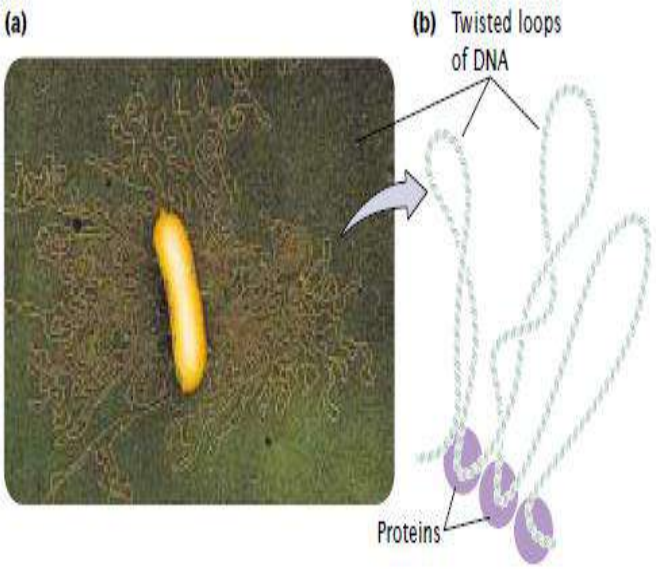
Supercoiling can only be introduced into or released from a closed-circular DNA molecule by breaking at least one of the phosphor diester backbones.

The level of supercoiling within a particular DNA molecule can be described by its linking number (*Lk).* This number corresponds to the number of double helical turns in the original linear molecule.

Supercoiling is Necessary for Packaging of Bacterial DNA

An average bacterial cell is about one millionth of a meter long. The length of the single DNA molecule needed to carry the 4,000 or so genes of a bacterial cell is about one millimeter! Thus, a stretched out bacterial chromosome is a thousand times longer than a bacterial cell.The double helical DNA inside a cell must be **supercoiled** to make it more compact.The DNA, which is already a double helix, is twisted again, Figure 5 The original double helix has a right-handed twist but the supercoils twist in the opposite sense; that is, they are left-handed or **"negative" supercoils**.There is roughly one supercoil every 200 nucleotides in typical bacterial DNA. Negative (rather than positive) supercoiling helps promote the unwinding and strand separation necessary during replication and transcription. [Eukaryotic DNA is also negatively supercoiled, however the mechanism is rather different and involves coiling it around histone proteins as discussed below.] Negative supercoils are introduced into the bacterial chromosome by DNA gyrase. In the absence of topoisomerase I and topoisomerase IV, the DNA becomes hypernegatively supercoiled. The steady-state level of supercoiling in *Escherichia coli* is maintained by a balance between topoisomerase IV, acting in concert with topoisomerase I, to remove excess negative supercoils and thus acting in opposition to DNA gyrase.A typical bacterial chromosome contains approximately 50 giant loops of supercoiled DNA arranged around a protein scaffold. In Figure 6, the single line represents a double helix of DNA and the helixes are the supercoils.

Bacterial chromosomes andplasmids are double stranded circular DNA molecules and are often referred to as **covalently closed circular DNA**, or **cccDNA**. If one strand of a double stranded circle is nicked, the supercoiling can unravel. Such a molecule is known as an **open circle**.



11.3 **Bacterial DNA is highly folded into a series of twisted loops.**
(Part a, Dr. Gopal Murti/Photo Researchers.)

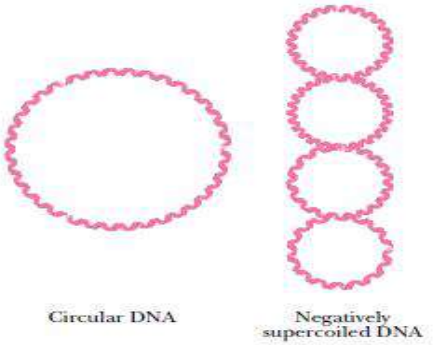


FIGURE 4.15 ***Supercoiling of DNA***

Bacterial DNA is negatively supercoiled in addition to the twisting imposed by the double helix.

FIGURE 4.16 ***Supercoiling of the Bacterial Chromosome***

Supercoiling of bacterial DNA results in giant loops of supercoiled DNA extending from a central scaffold.

The Eukaryotic Nucleosome

Each cell within our body contains a huge amount of DNA. The different chromosomes of the human genome contain approximately 3.2 × 109 base pairs of DNA. Since we are diploid organisms, having two sets of each chromosome, the total amount of DNA in most of our cells totals 6.4 × 109 base pairs. At 0.33 nm per base pair (Figure 1.7), this corresponds to an overall length of approximately 2.1 m.

How can this fit into a nucleus measuring just 5–10 µm across?

The answer is that the DNA is highly compacted. It is associated with a number of proteins that results in the wrapping of DNA into **nucleosomes**. During interphase, the genetic material (together with its associated proteins) is relatively uncoiled and dispersed throughout the nucleus as **chromatin**. When mitosis begins, the chromatin condenses greatly, and during prophase it is compressed into recognizable chromosomes. This condensation represents a contraction in length of some 10 000-fold.

The genetic material when isolated from bacteria and viruses consists of strands of DNA or RNA almost devoid of proteins. In eukaryotes, however, a substantial amount of protein is associated with the DNA to form chromatin.

Electron microscopic observations have revealed that chromatin fibers are composed of linear arrays of spherical particles. The particles occur regularly along the axis of a chromatin strand and resemble beads on a string. These particles, initially referred to as *ν*-bodies , are now called nucleosomes. The digestion of chromatin with certain nucleases, such as micrococcal nuclease, yields DNA fragments that are approximately 200 bp in length, or multiples thereofre . If the digestion of chromatin DNA were random, then a wide range of fragment sizes would be produced.

This therefore demonstrates that the DNA of chromatin consists of repeating units that are protected from enzymatic cleavage. The DNA between the units is attacked and cleaved by the nuclease, and multiples occur where two or more units are joined together.

The proteins associated with DNA in chromatin are divided into basic, positively charged **histones** and less positively charged **non-histones**.

Of the proteins associated with DNA, the histones play the most essential structural role. Histones contain large amounts of the positively charged amino acids lysine and arginine , making it possible for them to bind through electrostatic interactions to the negatively charged phosphate groups of the DNA nucleotides. There are five different types of histone protein – H1, H2A, H2B, H3 and H4. A nucleosome core particle consists of two copies each of histones H2A, H2B, H3 and H4 to form a histone octamer around which ~150 base pairs of DNA are wrapped in a left-handed superhelix, which completes about 1.7 turns per nucleosome (Figure - 9). This model is illustrated in(Figure -10). The 2 nm DNA double helix is initially coiled into a nucleosome core particle that is about 10 nm in diameter. Approximately 200 base pairs of DNA link each core particle to form the 'beads on a string' seen in electron microscopy images. Histone H1, which is not part of the core octamer, may be located at the site where DNA enters and leaves the nucleosome and possibly functions to seal the DNA around the nucleosome. The formation of nucleosomes represents the first level of packing, whereby the DNA is reduced to about one-third of its original length. In the nucleus, however, chromatin does not exist in this extended form. Instead, the 10 nm chromatin fibre is further packed into a thicker 30 nm fibre, The 30 nm fibre does, however, consist of numerous nucleosomes packed closely together, It has recently been suggested that the 30 nm fibre might adopt a compact helical zig-zag pattern with about four nucleosomes per 10 nm . The formation of the 30 nm fibre creates a second level of packaging, in which the overall length of the DNA is reduced some two fold.

The 30 nm fibre forms a series of looped domains that further condense the structure of the chromatin fibre. The fibres are then coiled into the chromosome arms that constitute a **chromatid**, which is part of the metaphase chromosome. In the overall transition from fully extended DNA helix to the extremely condensed status of the mitotic chromosome, a packaging ratio of about 500:1 must be achieve

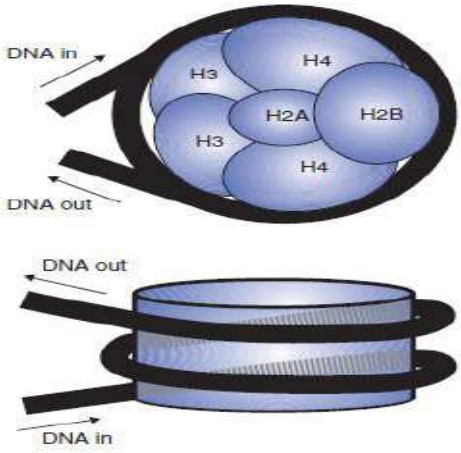
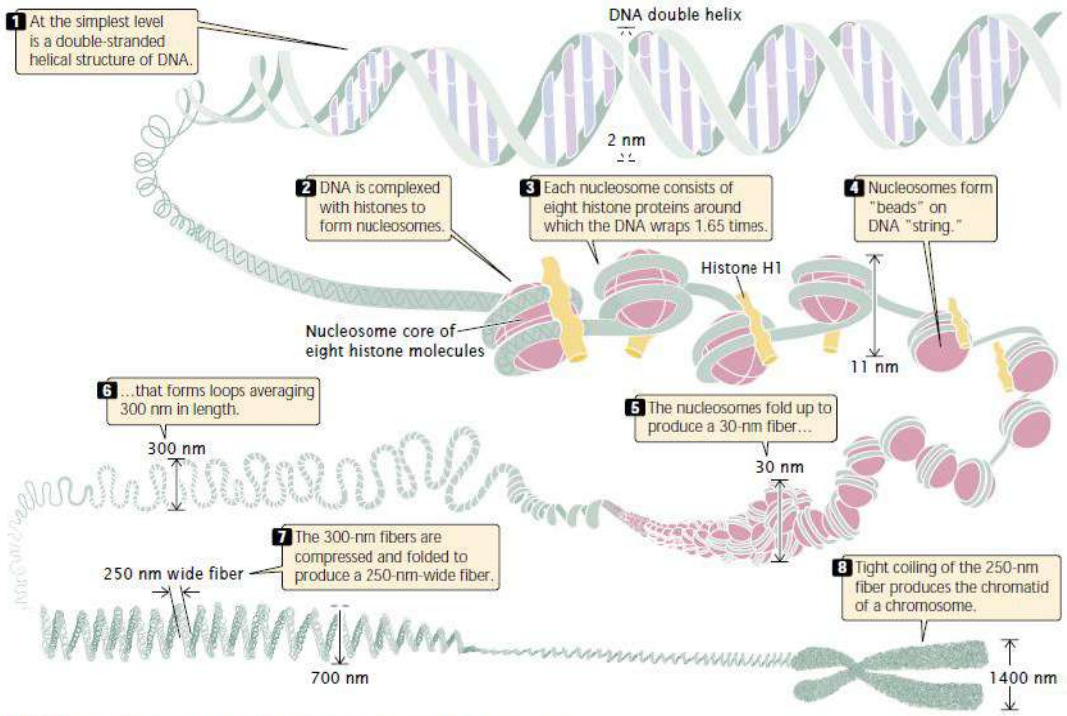


Figure 1.14. Wrapping of DNA around the nucleosome core. The nucleosome is composed of two molecules each of histones H4, H3, H2A and H2B. In the representation shown here only a monomer of H2A and H2B can be observed, the other monomers being located at the back of the octamer. Almost 150 bp of DNA wrap around the octamer core, forming approximately two turns



11.5 **Chromatin has a highly complex structure with several levels of organization.**

Table 11.3 Genome sizes of various organisms

Organism	Approximate Genome Size (bp)
λ (bacteriophage)	50,000
E. coli (bacterium)	4,600,000
Saccharomyces cerevisiae (yeast)	13,500,000
Arabidopsis thaliana (plant)	100,000,000
Drosophila melanogaster (insect)	140,000,000
Homo sapiens (human)	3,000,000,000
Zea mays (corn)	4,500,000,000
Amphiuma (salamander)	765,000,000,000

TABLE 4.01 Genome Sizes

Organism	Number of Genes	Amount of DNA (bp)	Number of Chromosomes
Viruses			
Bacteriophage MS2	4	3,600	1 (ssRNA)*
Tobacco Mosaic Virus	4	6,400	1 (ssRNA)*
ΦX174 bacteriophage	11	5,387	1 (ssDNA)
Influenza	12	13,500	8 (ssRNA)
T4 bacteriophage	200	165,000	1
Poxvirus	300	187,000	1
Bacteriophage G	680	498,000	1
Prokaryotes			
Mitochondrion (human)	37	16,569	1
Mitochondrion (*Arabidopsis*)	57	366,923	1
Chloroplast (*Arabidopsis*)	128	154,478	1
Nanoarchaeum equitans	550	490,000	1
Mycoplasma genitalium	480	580,000	1
Methanococcus	1,500	1.7 Mbp	1
Escherichia coli	4,000	4.6 Mbp	1
Myxococcus	9,000	9.5 Mbp	1
Eukaryotes (haploid genome)			
Encephalitozoon	2,000	2.5 Mbp	11
Saccharomyces	5,700	12.5 Mbp	16
Caenorhabditis	19,000	100 Mbp	6
Drosophila	12,000	140 Mbp	5
Homo sapiens	25,000	3,300 Mbp	23
Arabidopsis	25,000	115 Mbp	5
Oryza sativa (Rice)	45,000	430 Mbp	12

*ssRNA = single stranded RNA; ssDNA = single stranded DNA; all other genomes consist of double stranded DNA.

Lecture 5

DNA Replication –

DNA Replication

* DNA Replication Is a Two-Stage Process Occurring at the Replication Fork:

Replication is the process by which the DNA of the ancestral cell is duplicated, prior

to cell division. Upon cell division, each of the descendants will get one complete cop of the DNA that is identical to its predecessor.

The first stage in replication is to separate the two DNA strands of the parental DNA molecule.

The second stage is to build two new strands, using each of the two original strands as **templates**. The most fundamental aspect of replication is the base pairing of A with T and of G with C. Each of the separated parental strands of DNA serves as a template strand for the synthesis of a new complementary strand. The incoming nucleotides for the new strand recognize their partners by base pairing and so are lined up on the **template strand** . Since A pairs only with T, and since G pairs only with C, the sequence of each original strand dictates the sequence of the new complementary strand.Synthesis of both new strands of DNA occurs at the **replication fork** that moves along the parental molecule. Amazingly, in *E. coli*, DNA is made at nearly 1,000 nucleotides per second. The replication fork consists of the zone of DNA where the strands are separated, plus an assemblage of proteins that are responsible for synthesis, sometimes referred to as the **replisome**. The result of replication is two double stranded DNA molecules, both with sequences identical to the original one. One of these daughter molecules has the original left strand and the other daughter has the original right strand.

Three different strategies for replication of the double helix seemed possible :

1- The pattern of replication is **semi-conservative**, since each of the progeny conserves half of the original DNA molecule (the daughter molecules each contain one polynucleotides derived from the original molecules and one newly synthesized strand).

2- conservative replication, one of the two "daughter" double helixes would consist entirely of original DNA strands, while the other helix would consist of two newly synthesized strands.

3- dispersive replication, both "daughter" double helixes would carry blocks of original DNA interspersed with blocks of newly synthesized material.

These alternatives are less satisfactory than semiconservative replication because they do not immediately suggest a mechanism for copying the information in the sequence of bases,

Introduction to Prokaryotic replication

1953

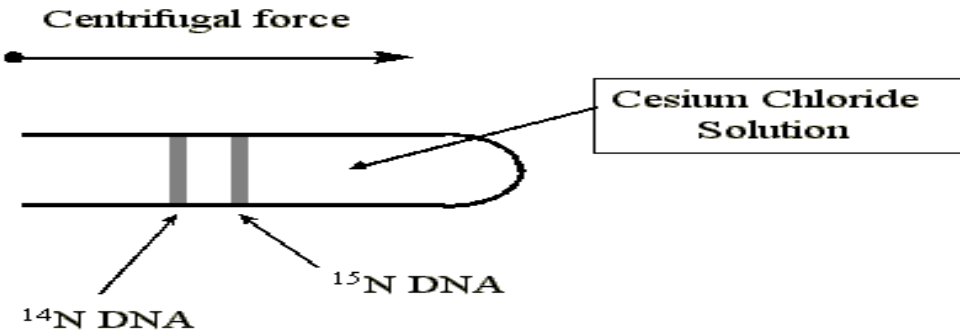
- The DNA model of Watson and Crick suggested how genetic information might be replicated: either strand of the duplex can be used as a template to replicate the sequence information.
- But, was the replication conservative (i.e. the original parental strands remain together after replication) or semi-conservative (one parental strand pairs with one newly synthesized strand)?

The answer for prokaryotic organisms (i.e. lack a true membrane bound nucleus and cellular organelles; e.g. bacteria) came from the 1958 experiment of Meselson and Stahl.

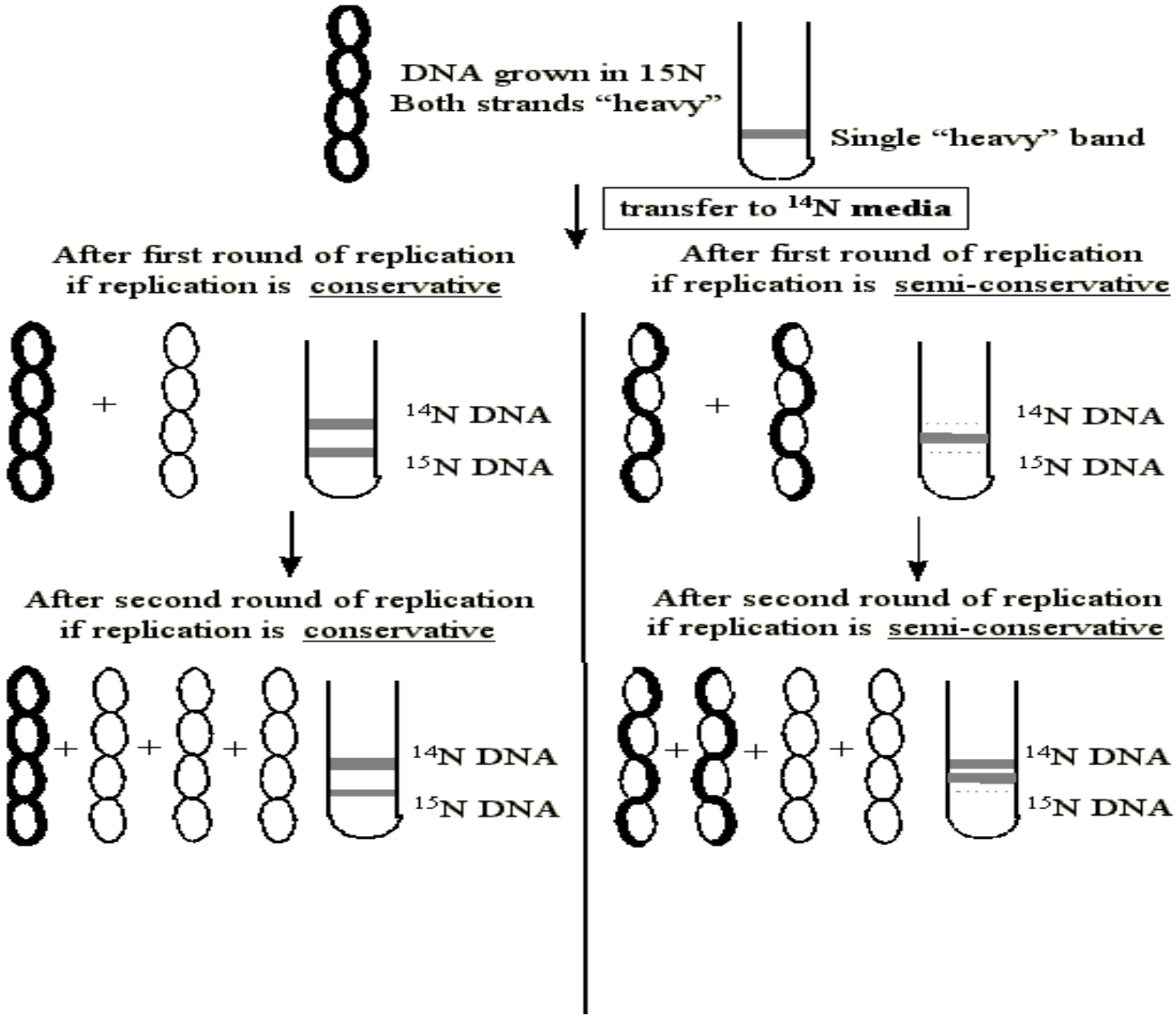
1958

The nitrogen in ammonium salts in culture broth is incorporated into DNA bases. The most common isotope of nitrogen is ^{14}N. However, ^{15}N ammonium salts (a heavier isotope) can also be obtained.

- DNA from *E. coli* cells grown with ^{15}N ammonium salts will have a *higher density* than DNA grown in "normal" (^{14}N) ammonium salts.
- Such DNA will migrate differently on cesium chloride ($CsCl_2$) equilibrium density gradient centrifugation.
- The more dense DNA will migrate as a lower band (on this type of centrifugation the characteristic migration position is a function of density, and *is independent of DNA length*).

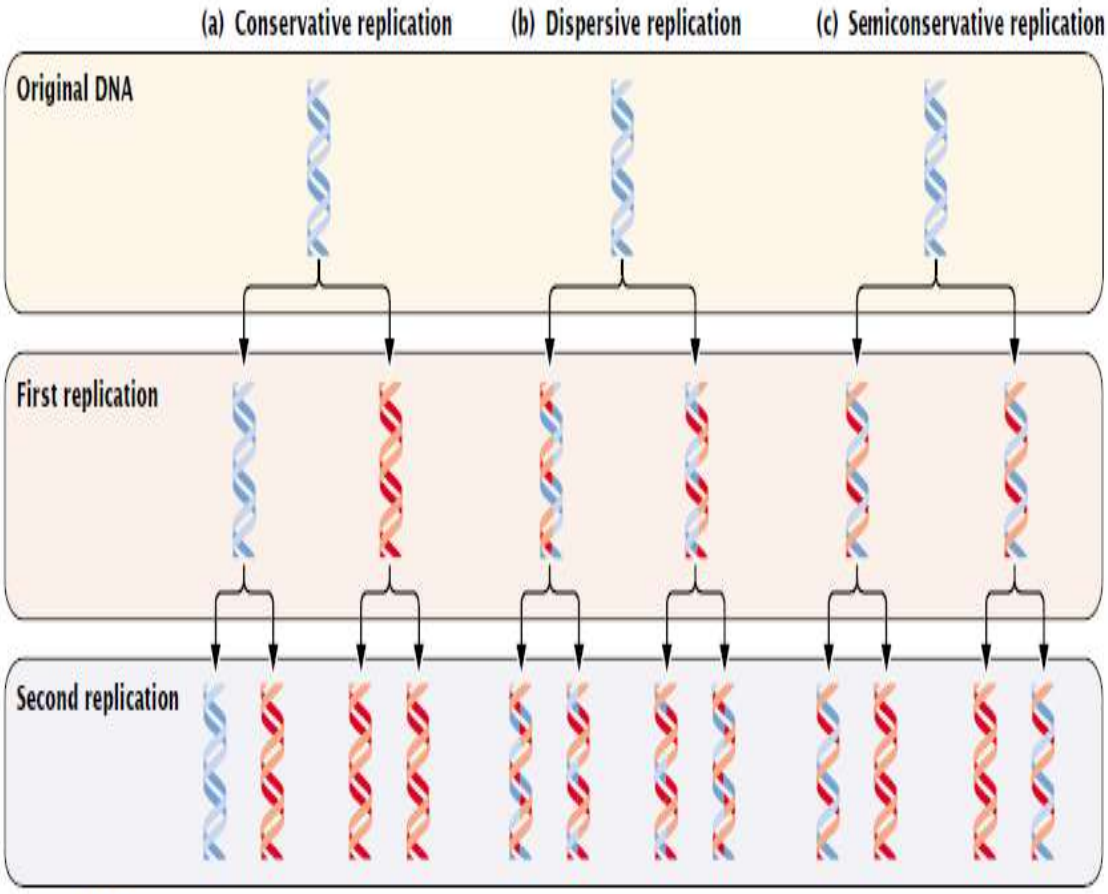


Meselson and Stahl reasoned that if they grew *E. coli* in ^{15}N salts then ***switched media to ^{14}N salts*** for additional rounds of replication, the <u>mode of replication</u> could be deduced from the <u>density</u> of the DN

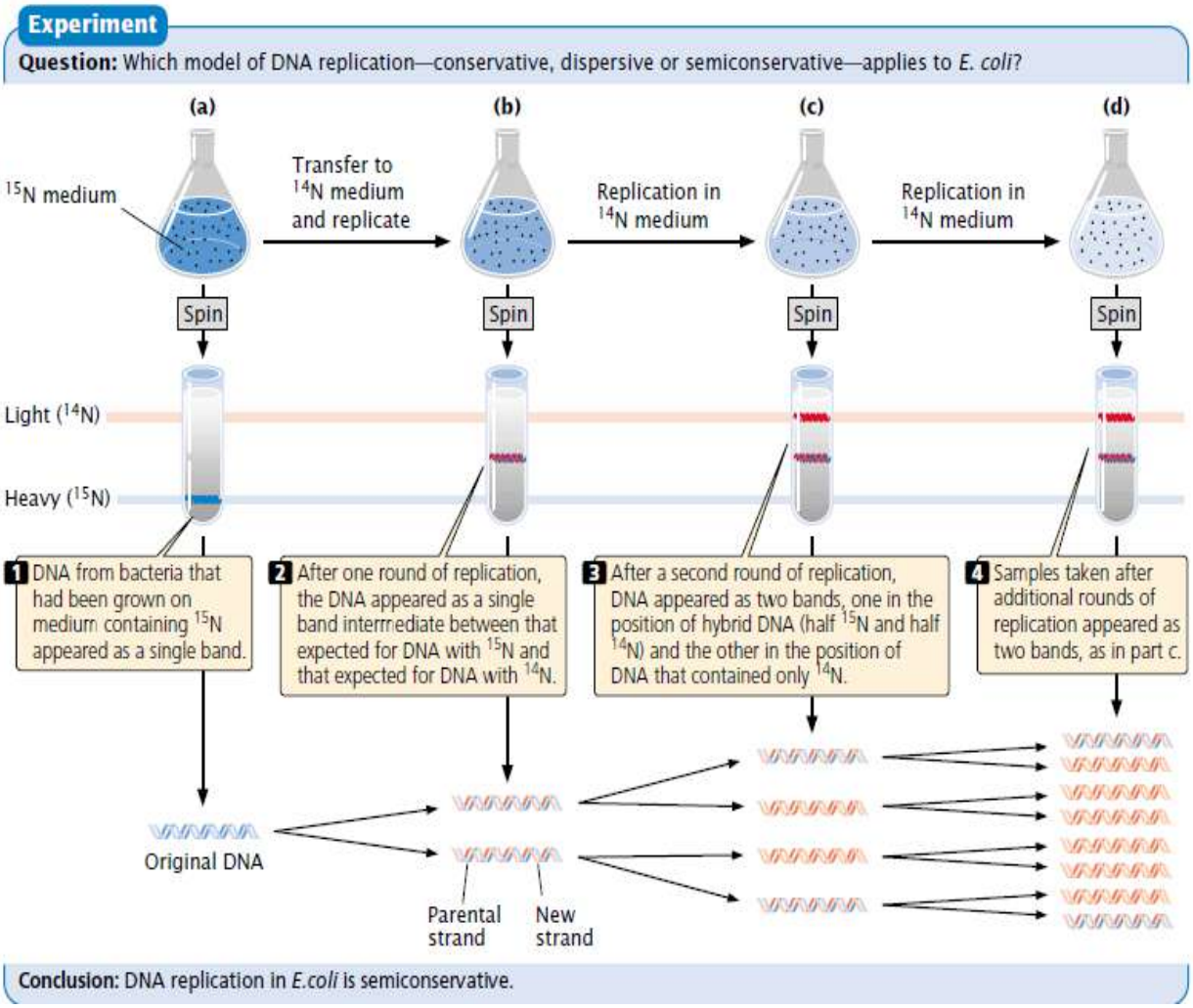


-
- After switching to the ^{14}N media and allowing the cells to go through a round of replication a single band of intermediate density was observed (i.e. between ^{14}N and ^{15}N control DNA samples).
- After a second round of replication in ^{14}N media two bands were present in approximately equimolar amounts; one was intermediate in density and the other migrated as purely ^{14}N labeled DNA.

The results were consistent with a semi-conservative mode of replication for DNA. Evidence of semi-conservative replication of DNA has since been obtained with both plant and animal DNA.



12.1 Three proposed models of replication are conservative replication, dispersive replication, and semiconservative replication.



Mechanism of DNA replication :

1-Deoxy ribonucleotide precursors synthesis

In DNA : dATP, dGTP dCTP, dTTP.

In RNA : ATP, GTP CTP, UTP.

2- Deoxynucleotide polymerization .

3- enzymes and proteins:

DNA polymerase , primase , toboiomerae 1, toboiomerae11, ligase , SS protein ,Helicase.

4- RNA primers

5- template strand .

Summary of steps in *E. coli* DNA synthesis:

1. dnaA protein melts duplex in oriC region.
2. dnaB (helicase), along with dnaC and ATP binds to replication fork (dnaC protein exits).[1] (**Pre-priming complex**)
3. Single strand binding protein (ssb protein) binds to separated strands of DNA and prevents reannealing.
4. Primase complexes with helicase, creates RNA primers ($pppAC(N)_{7-10}$) on the strands of the open duplex[2] (Primase+helicase constitute the **Primosome**).
5. After making the RNA primers, **DNA pol III holoenzyme comes in** and extends the RNA primer (laying down dNTP's) on the leading strand.
6. As the replication fork opens up (via helicase + ATP action) leading strand synthesis is an uninterrupted process, the lagging strand experiences a **gap**.
7. The gap region of the lagging strand can wind around one active site unit of the Pol III complex, and bound Primase **initiates an RNA primer in the gap region[3].**
8. On the lagging strand, Pol III extends the RNA primer with dNTP's as the lagging template strand is looped through the Pol III complex
9. After synthesis of a nascent fragment the lagging strand loop is released and the single strand region further up near the replication fork is subsequently looped through the Pol III complex.
10. Steps 7-9 are repeated.
11. Meanwhile, Pol I removes the RNA primer regions of the Okazaki fragments via 5' to 3' exonuclease activity (nick translation
12. Pol I exits and ligase joints the DNA fragments (on lagging strand).

Notes from above:

1. Polymerases are unable to open up duplex DNA, thus the requirement for helicase

2. Polymerases cannot replicate a DNA template in the absence of a primer (either DNA or RNA).

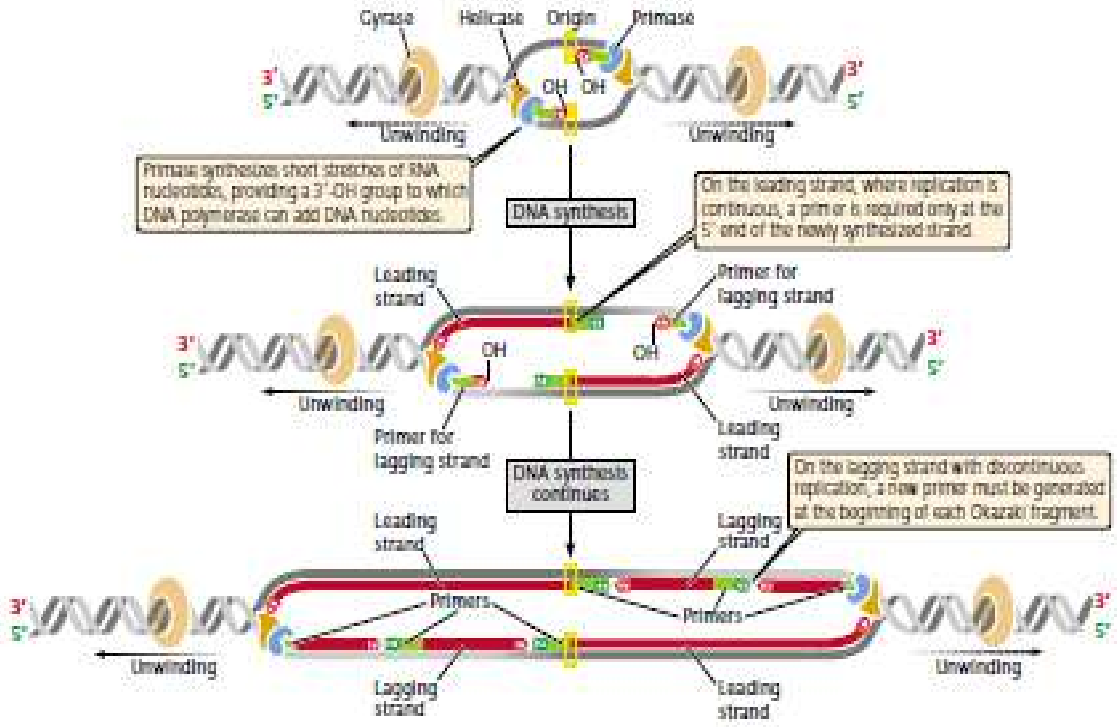
3. Polymerases extend a polynucleotide in the 5' to 3' direction only. Gaps at the 5' end must be filled by "upstream" discontinuous synthesis.

- Pol I: gap filling during DNA synthesis and repair, removal of RNA primers
- Pol II: involved in DNA synthesis of damaged templates
- Pol III: functional polymerase at the replication fork

Termination of DNA replication

- Specific termination sites of DNA replication exist in *E. coli*.

- Termination involves the binding of the *tus* gene product (*tus* protein).
- This protein may act to prevent helicase from unwinding DNA (will therefore halt *pol III* and *pol I* action).
- DNA replication produces two interlocking rings which must be separated.
- This is accomplished via the enzyme *topoisomerase*.



DNA Polymerases in prokaryotic cells

- In *E. coli* , *Bacillus subtilis*, three different DNA polymerases discovered :
- DNA polymerases I and II are a single whereas DNA polymerase III is complex at least three polypeptides for functional activity and as many as four additional polypeptide coenzymes.

- Initially when poll was discovered by Arthur Kornberg in 1958, it was thought to be the enzyme responsible for DNA replication in the cell.
- It POI III appears to be the major enzyme involved in DNA replication.

- Short segments of DNA (about 2000 nucleotides in prokaryotic cells and about 200 nucleotides in euokaryotic cells) known as Okazaki fragments formed by the DNA polymerase running opposite the direction of unwinding of the parental DNA lecule. The okazaki fragments are later joined by action of DNA ligase , which establishes a phosodiester bond between the 3-OH and 5-P ends of nucleotides.

-

Lecture 6:

Replication in Eukaryotes :

Eukaryotic DNA polymerases

A significant difference in the processes of bacterial and eukaryotic replication is in the number and functions of DNA polymerases. Eukaryotic cells contain a number of different DNA polymerases that function in replication, recombination, and DNA repair (Table 12.5).

***DNA polymerase** (alpha)**,** which contains primase activity, initiates nuclear DNA synthesis by synthesizing anRNA primer, followed by a short string of DNA nucleotides.

*After DNA polymerase has laid down from 30 to 40 nucleotides, **DNA polymerase** (delta)completes replication on the leading and lagging strands. ***DNA polymerase** (beta) does not participate in replication but is associated with the repair and recombination of nuclear DNA.

***DNA polymerase** (gamma) replicates mitochondrial DNA;

* a (gamma) like polymerase also replicates chloroplast DNA. Similar in structure and function to DNA polymerase (delta),

***DNA polymerase**((epsilon)_ appears to take part in nuclear replication of both the leading and the lagging strands, but its precise role is not yet clear. *OtherDNA polymerases ($\zeta, \eta, \theta, \kappa, \lambda, \mu$) allow replication to bypass damaged DNA (called translesion replication) or play a role in DNA repair. Many of the DNA polymerases have multiple roles in replication and DNA repair .

Table 12.5 DNA polymerases in eukaryotic cells

DNA Polymerase	5′→3′ Polymerase Activity	3′→5′ Exonuclease Activity	Cellular Function
α (alpha)	Yes	No	Initiation of nuclear DNA synthesis and DNA repair
β (beta)	Yes	No	DNA repair and recombination of nuclear DNA
γ (gamma)	Yes	Yes	Replication of mitochondrial DNA
δ (delta)	Yes	Yes	Leading- and lagging-strand synthesis of nuclear DNA, DNA repair, and translesion DNA synthesis
ε (epsilon)	Yes	Yes	Unknown; probably repair and replication of nuclear DNA
ζ (zeta)	Yes	No	Translesion DNA synthesis
η (eta)	Yes	No	Translesion DNA synthesis
θ (theta)	Yes	No	DNA repair
ι (iota)	Yes	No	Translesion DNA synthesis
κ (kappa)	Yes	No	Translesion DNA synthesis
λ (lambda)	Yes	No	DNA repair
μ (mu)	Yes	No	DNA repair
σ (sigma)	Yes	No	Nuclear DNA replication (possibly), DNA repair, and sister-chromatid cohesion

Eukaryotic Chromosomes Have Multiple Origins

Eukaryotic chromosomes are often very long and have numerous replication origins scattered along each chromosome. Replication is bi-directional, as in bacteria. A pair of replication forks starts at each origin of replication and the two forks then move in opposite directions (Fig.).The bulges where the DNA is in the process of division are often called **replication bubbles**. A vast number of replication origins function simultaneously during eukaryotic DNA replication. For example, there are estimated to be between 10,000 and 100,000 replication origins in a dividing human somatic cell. This creates major problems in synchronization. Synthesis at each origin must be coordinated to make sure that each chromosome is completely replicated. Conversely, each origin must initiate once and once only during each replication cycle in order to avoid

duplication of DNA segments that have already been replicated. This is achieved by a protein complex, known as replication licensing factor (RLF), which binds to the DNA next to each origin before each replication cycle and is displaced during replication. Only when RLF is present is DNA replication permitted

Synthesis of Eukaryotic DNA

The synthesis of DNA in eukaryotes is less well investigated than in bacteria. Nonetheless, the same general principles apply, although there are differences in detail from the bacterial scheme. In eukaryotes, **semi-conservative replication** occurs. One new strand is made continuously and the other in fragments. Both strands are made simultaneously by a replisome consisting of a helicase plus two DNA polymerase assemblies. A sliding clamp holds the polymerase on the DNA.An RNA primer is required. In animal cells, two DNA polymerases (a and d) are involved in chromosome replication . DNA **polymerase** a is responsible for initiation of new strands. It is accompanied by two smaller proteins that make the RNA primer. After the RNA primer has been made, polymerase a elongates it by a short piece of DNA only three or four bases long (the **initiator DNA**, or "iDNA"). Another protein, **Replication factor C** (**RFC**), then binds to the iDNA and loads DNA **polymerase** d plus its sliding clamp (**PCNA protein**) onto the DNA.Two assemblies of DNA polymerase d elongate the two new strands.The sliding clamp of animal cells is a trimer (not a dimer as in bacteria) that forms a ring surrounding the DNA. It was named PCNA, for proliferating cell nuclear antigen, before its role was fully known. Linking of the Okazaki fragments differs significantly between animal and bacterial cells. In animals, there is no equivalent of the dual function polymerase I of bacteria. The RNA primers are removed by an exonuclease (MF1) and the gaps are filled by the DNA polymerase d that is working on the lagging strand. As in bacteria, the nicks are sealed by DNA ligase.

Cell Division in Higher Organisms

The eukaryotic cells of higher organisms face further problems during cell division. Not only do they have multiple chromosomes, but these are inside the nucleus, separated from the rest of the cell by the nuclear membrane. Consequently, an elaborate process is needed to disassemble the nucleus, replicate the chromosomes and partition them among the daughter cells.This process is mitosis and involves several operations:

1. Disassembly of the nuclear membrane of the mother cell

2. Division of the chromosomes

3. Partition of the chromosomes

4. Reassembly of nuclear membranes around each of the two sets of chromosomes

5. Final division of the mother cell, or **cytokinesis**.

Mitosis itself is only one of several phases of the eukaryotic **cell cycle**

The process of DNA replication described above takes place in the synthetic, or **Sphase**, of the cell cycle. The S-phase is separated from the actual physical process of cell division (**mitosis**) by two gap phases, or **G-phases**, in which nothing much appears to happen (except the normal processes of cellular activity and metabolism).Together, G1, S and G2 constitute **interphase**.

cytokinesis Cell division
DNA polymerase alpha Enzyme that makes short segment of initiator DNA during replication of animal chromosomes
DNA polymerase dalta Enzyme that makes most of the DNA when animal chromosomes are replicated
initiator DNA (iDNA) Short segment of DNA made just after the RNA primer during replication of animal chromosomes
PCNA protein The sliding clamp for the DNA polymerase of eukaryotic cells (PCNA = proliferating cell nuclear antigen)
replication factor C (RFC) Eukaryotic protein that binds to initiator DNA and loads DNA polymerase d plus its sliding clamp onto the DNA
semi-conservative replication Mode of DNA replication in which each daughter molecule gets one of the two original strands and one new complementary strand

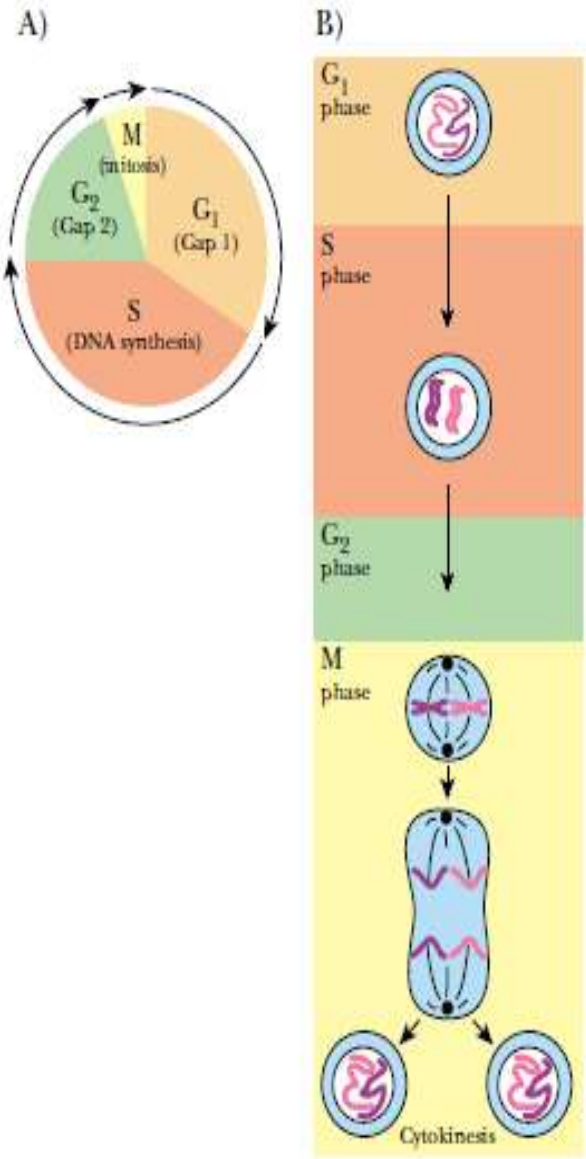
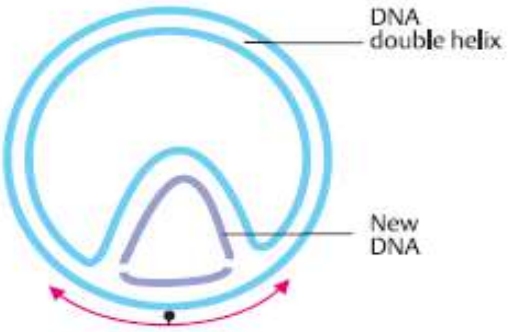
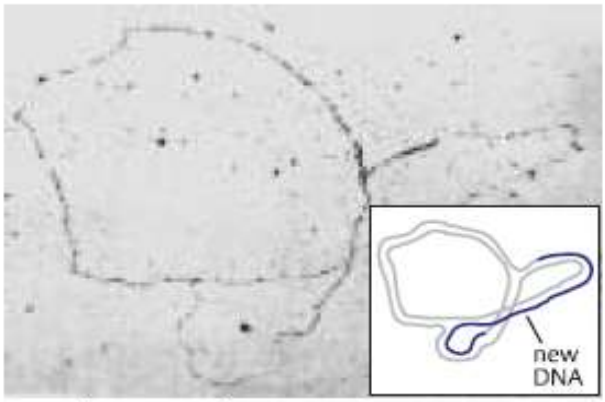


FIGURE 5.32 ***The Eukaryotic Cell Cycle***

DNA replication occurs during the S phase of the cell cycle but the chromosomes are actually separated later, during mitosis or M phase. The S and M phases are separated by G1 and G2.

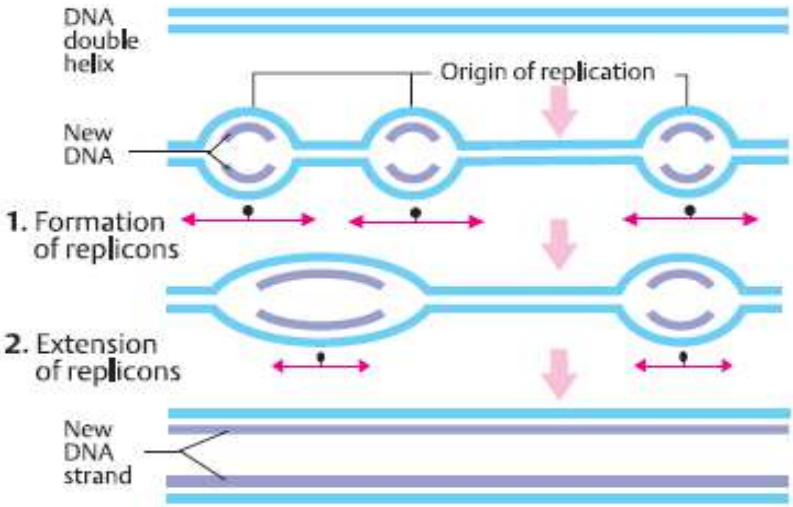


1. DNA replication in the bacterial chromosome

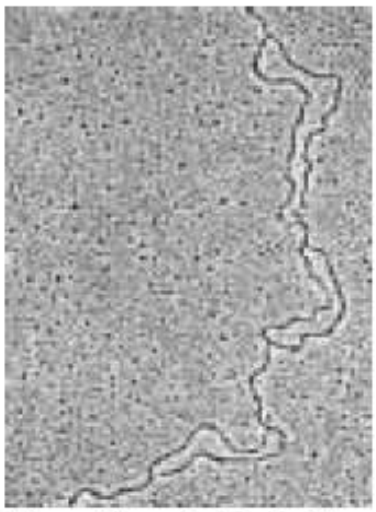


2. Prokaryotic replication in an autoradiogram in *E. coli* (J. Cairns)

A. Prokaryotic replication begins at one site

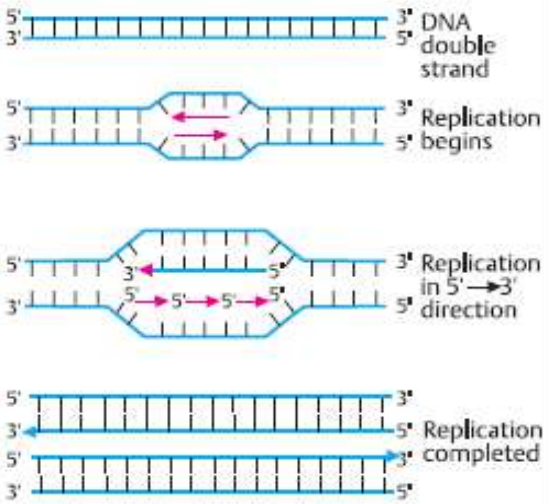


3. Replication completed

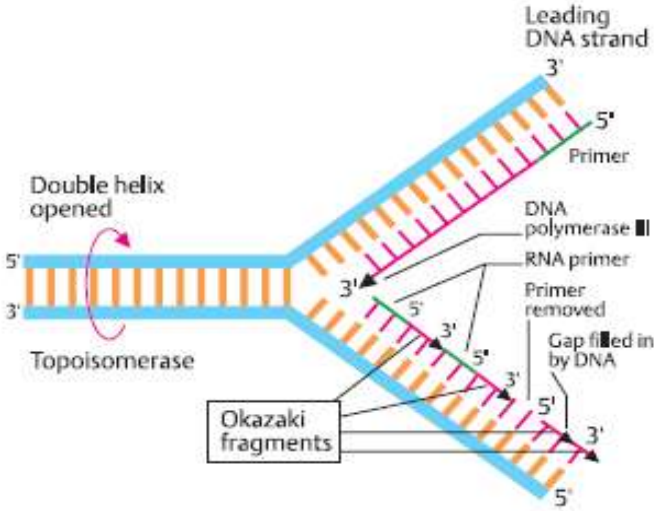


4. Eukaryotic replication in the EM (D. S. Hogness)

B. Eukaryotic replication begins at several sites



C. Scheme of replication

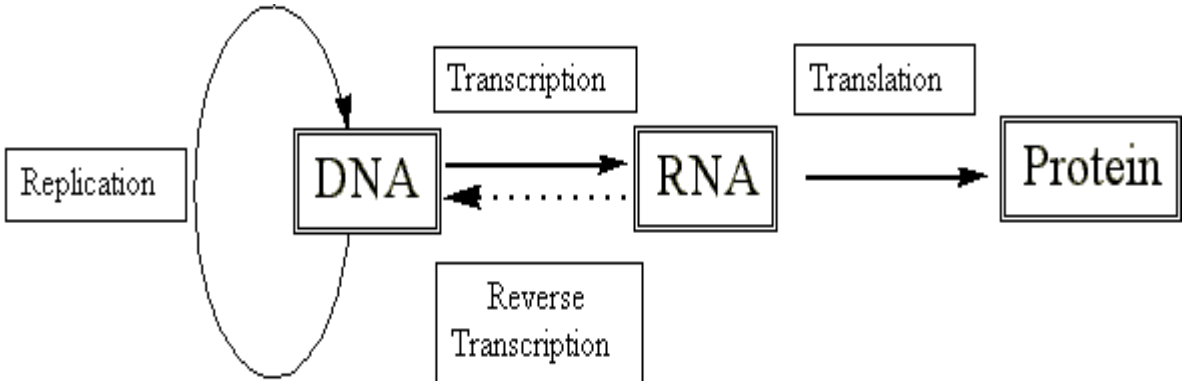


D. Replication fork

Lecture 7: Dr.Nuha Joseph Kandala

Introduction to Gene Expression :

The "central dogma"



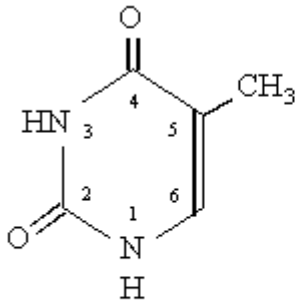
- We have seen how DNA, with the aid of specific polymerases and accessory proteins, is able to replicate.
- We have also seen how we can use this information to create autonomously replicating extra-chromosomal elements (i.e. **plasmids**).

*However, the real utility of such systems arises when we use them to create **proteins** of interest. To get to proteins we have to go through **RNA** first.*

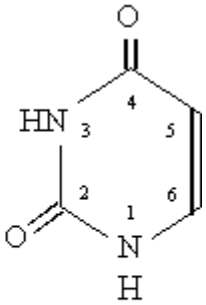
Structural features of RNA:

1. Similar to DNA except it contains a 2' hydoxyl group (**makes phosphodiester bond more labile than DNA**).

2. **Thymine** in DNA is replaced by **Uracil** in RNA



Thymine
(DNA)



Uracil
(RNA)

3. RNA's can adopt regular ***three-dimensional structures*** which allow them to function in the process of genetic expression (i.e. the production of proteins).

- This ability to adopt defined three dimensional structures which impart functionality places RNA in a unique class - somewhat akin to proteins, and different from DNA.
- For example certain RNA molecules, when folded, exhibit ***catalytic capacities*** (e.g. the cleavage of RNA molecules).
- The majority of RNA in cells is found in <u>complexes with proteins</u>. The most common example is <u>ribosomes</u> (involved in protein synthesis).

Transcription: the copying of DNA by an RNA polymerase to make RNA.

<u>RNA polymerase</u>:

- <u>*Can initiate*</u> a new nucleic acid strand given a template.
- DNA polymerases **cannot**; they require a **primer** (or more typically, an RNA polymerase to provide the primer).

- Three kinds of RNA molecules perform different functions in the protein synthesizing aparatus:

1. *Messenger RNA* (mRNA) encodes the genetic information copied from DNA in the form of a sequence of bases that specifices a sequence of amino acids

2-Ribosomal RNA (rRNA) forms a complex with specific proteins to form the **ribosome** which is the key translational component

- the ribosome complexes with mRNA and directs appropriate tRNA's and the synthesis of the polypeptide bond.

3-Transfer RNA (tRNA) is part of the structural machinery which deciphers the mRNA code.

- They carry specific amino acids which are transfered to a nacent polypeptide according to the instructions contained within the mRNA.

Structure and function of Ribosomal RNA:

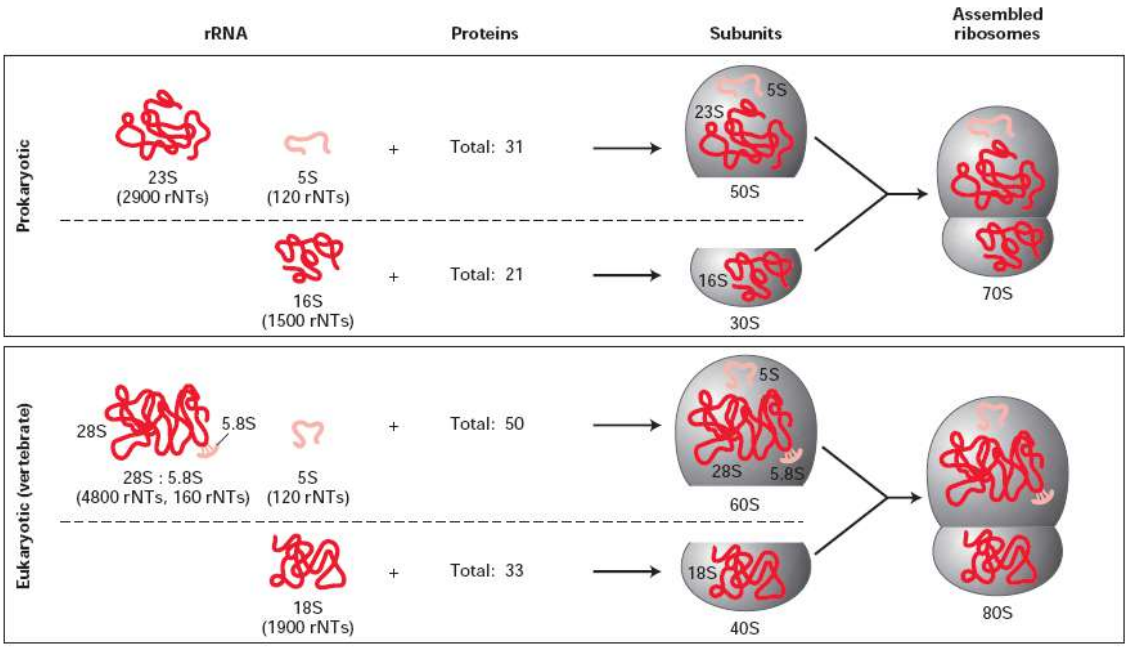


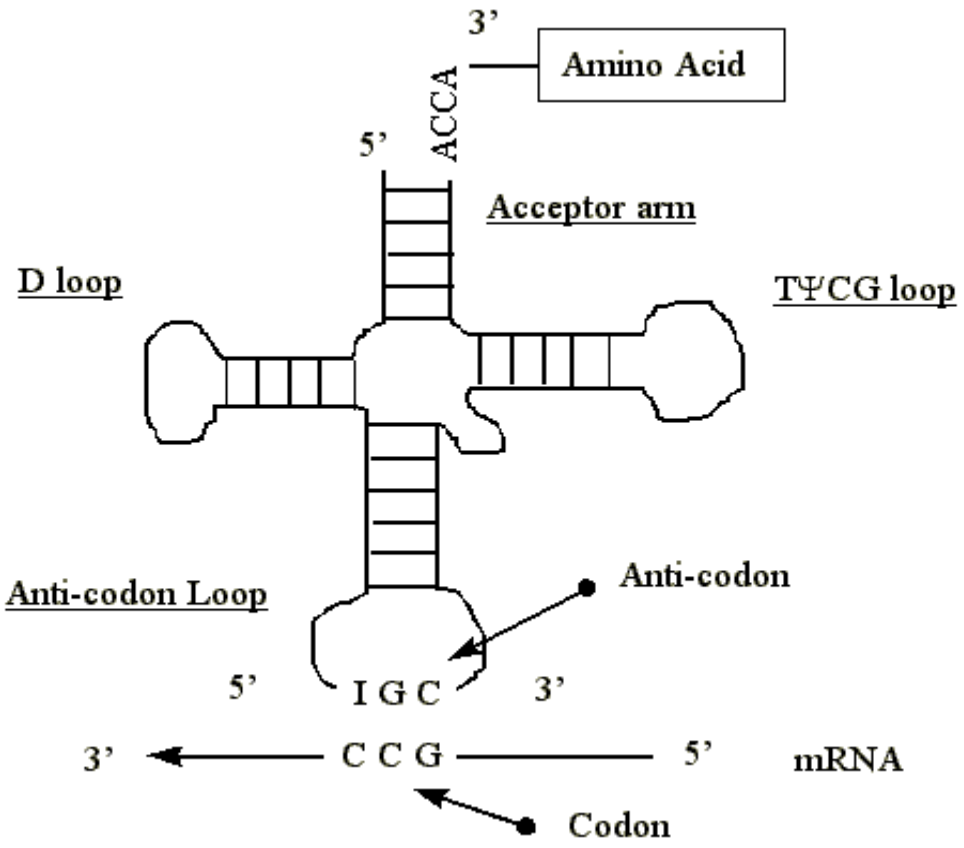
Table 14.3 Composition of ribosomes in bacterial and eukaryotic cells

Cell Type	Ribosome Size	Subunit	rRNA Component	Proteins
Bacterial	70S	Large (50S)	23S (2900 nucleotides) 5S (120 nucleotides)	31
		Small (30S)	16S (1500 nucleotides)	21
Eukaryotic	80S	Large (60S)	28S (4700 nucleotides) 5.8S (160 nucleotides) 5S (120 nucleotides)	49
		Small (40S)	18S (1900 nucleotides)	33

Structure and function of transfer RNA's

Structure of tRNA's

- 70-80 nucleotides long
- Form a series of stem/loop secondary structures
- tRNA's are synthesized with the standard bases AGCU. However, after synthesis several bases may be modified:
    1. Uridylate may be **methylated** to produce **Thymidylate**
    2. Uridylate may be **rearranged** to produce **pseudouridylate** (i.e. ribose attached to Carbon 5 instead of Nitrogen 1).
    3. Guanidylate may be **methylated** at different positions.
- The amino acid is attached at the 3' end of the tRNA to either the 2' hydroxyl or the 3' hydroxyl.
    1. ***Class I*** *amino-acyl tRNA synthetases attach their associated amino acids to the tRNA* ***2' hydroxyl*** (NOTE: typically the hydrophobic amino acids)
    2. ***Class II*** *amino-acyl tRNA synthetases attach their associated amino acids to the tRNA* ***3' hydroxyl*** (NOTE: typically hydrophilic amino acids)



- **tRNA's** have ***two*** functions:
    1. To chemically link to a particular amino acid (***covalent***)
    2. To recognize a specific **codon in mRNA** (***non-covalent***) so that its attached amino acid can be added to a growing peptide chain

Amino-acyl tRNA synthetases

- Function is to ***"charge"*** tRNA molecules; i.e. *to chemically link a specific amino acid to its associated tRNA molecule*.

Amino Acids	Amino-acyl tRNA synthetases	tRNA's	Codons
20	20	30-40 (prokaryotes) 50 (eukaryotes)	61 (3 stop codons)

Conclusions:

1. *There is one amino-acyl tRNA synthetase per amino acid* (they are quite specific).
2. *There is potentially more than one tRNA per amino acid*.

Therefore, amino-acyl tRNA synthetases must be able to recognize more than one tRNA.

1. *There is potentially more than one codon per tRNA*.

Therefore each tRNA must be able to recognize more than one codon (there is not a unique tRNA for each codon).

- If perfect Watson-Crick base pairing were required at the codon/anti-codon triplet then ***61 different tRNA's would be required***.
- We know this is not the case, therefore a single tRNA anti-codon must be able to recognized several different mRNA codon triplets.

- This greater recognition of tRNA is possible due to "**wobble**" basepair interactions at the third base in the codon/first base in the anti-codon:

wobble hypothesis

Clark (1966) proposed Wobble Hypothesis. According to this hypothesis, the base in first position of anti-codon on tRNA is usually an abnormal base, like inosine, pseudouridine,. These abnormal bases are able to pair with more than one type of nitrogenous base in the third position of the mRNA.
For example inosine (I) can pair with A , C or U. this base is called Wobble base .

- Wobble occurs at position 1 of the anti-codon and position 3 of codon

Possible "wobble" codon base pairing (in addition to Watson-Crick):

1. **U - G**
2. **I - C**
3. **I - A**
4. **I - U**

- Where U, G, A and C can be in ***either*** the codon (mRNA) or anti-codon (tRNA)
- **I** (inosine) can be found in the ***anti-codon***.

For example, the codons UUU and UUC are both recognized by the tRNA which has GAA in the anti-codon position (making either G - C, or G - U base pairings).

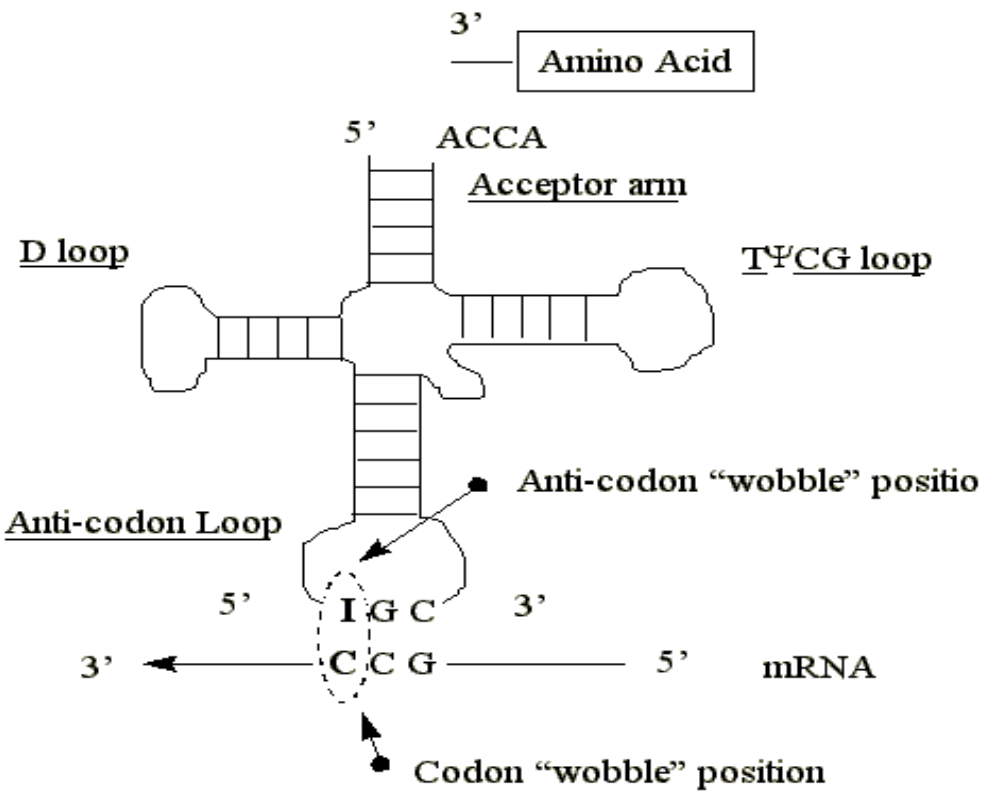
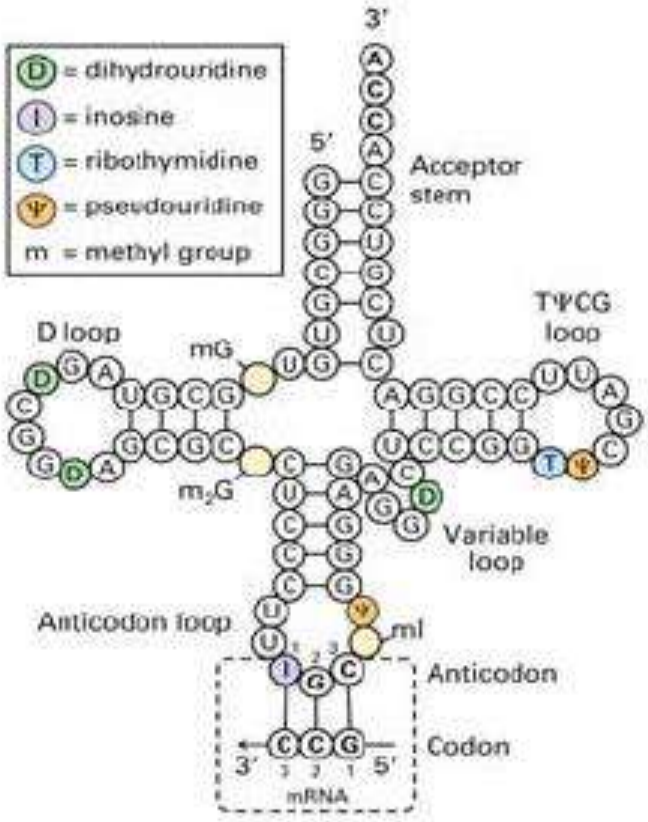
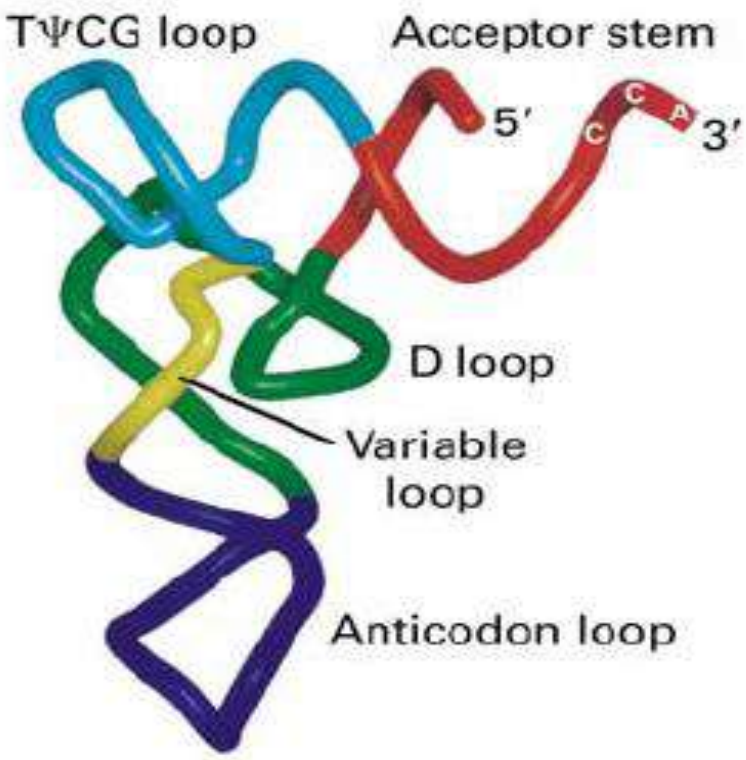


TABLE 8.01 Wobble Rules for Codon/Anticodon Pairing

	Pairs with Third Codon Base	
First Anticodon Base	**normal**	**by wobble**
G	C	U
U	A	G
I	—	C or U or A
C	G	no wobble
A	U	no wobble





Characteristics of the Genetic Code

number of characteristics of the genetic code. these characteristics.

1. The genetic code consists of a sequence of nucleotides in DNA or RNA. There are four letters in the code, corresponding to the four bases—A, G, C, and U (T in DNA).

2. The genetic code is a triplet code. Each amino acid is encoded by a sequence of three consecutive nucleotides, called a codon.

3. The genetic code is degenerate—there are 64 codons but only 20 amino acids in proteins. Some codons are synonymous, specifying the same amino acid.

4. Isoaccepting tRNAs are tRNAs with different anticodons that accept the same amino acid; wobble allows the anticodon on one type of tRNA to pair with more than one type of codon on mRNA.

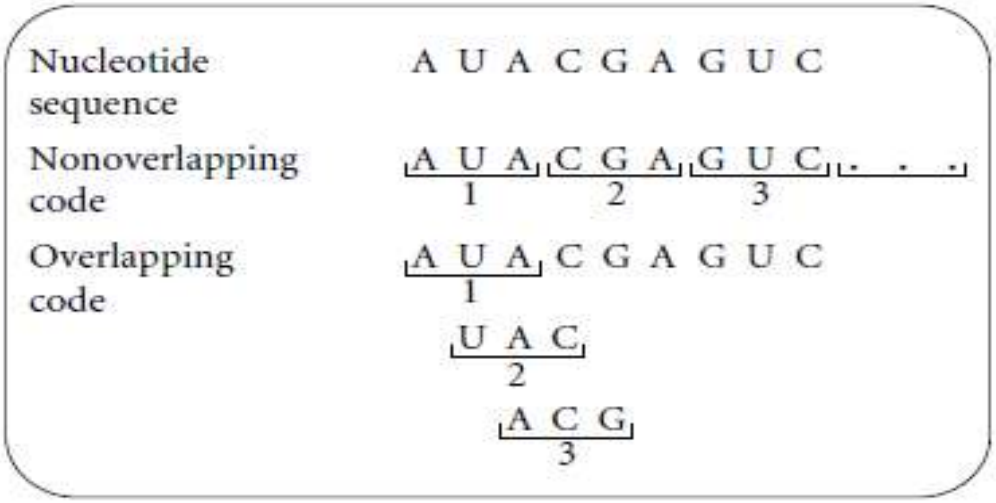
5. The code is generally nonoverlapping; each nucleotide in an mRNA sequence belongs to a single reading frame.

6. The reading frame is set by an initiation codon, which is usually AUG.

7. When a reading frame has been set, codons are read as successive groups of three nucleotides.

8. Any one of three termination codons (UAA, UAG, and UGA) can signal the end of a protein; no amino acids are encoded by the termination codons.

9. The code is almost universal.



15.14 The genetic code is generally nonoverlapping. In a nonoverlapping code, each nucleotide belongs to only one codon. In an overlapping code, some nucleotides belong to more than one codon. The genetic code used in almost all living organisms is nonoverlapping.

Lecture8:

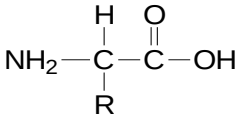
The proteins:

The sequence of nucleotides in DNA or RNA that determines the specific amino acid sequence in the synthesis of proteins. It is the biochemical basis of heredity and nearly universal in all organisms.

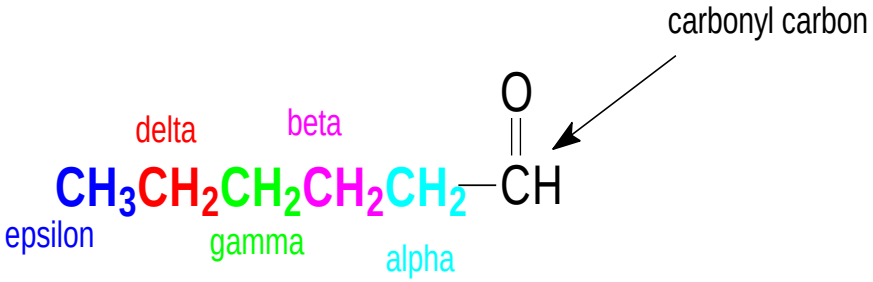
To understand how information in a gene specifies the production of a particular protein, it is necessary to review the chemical composition of protein. Proteins are about 50% of the dry weight of most cells, and are the most structurally complex molecules known. Each type of protein has its own unique structure and function.

Proteins are polymers of amino acids in which different amino acids are linked together through peptide bonds.Polymersare any kind of large molecules made of repeating identical or similar subunits called monomers.

Proteins are polymers of about 20 amino acids monomers. **Amino Acids are the common compounds which contain both carboxylic and amino group. Carboxylic group is acidic whereas, the amino group is basic.**



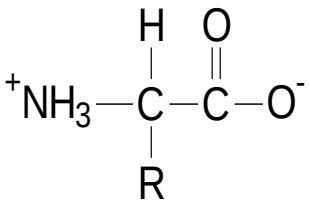
The amino acids referred to as **alpha amino acids**. The reason is that the central carbon is in an **alpha** position in relation to the carbonyl carbon. The carbon adjacent to the carbonyl carbon is designated the **alpha carbon**. Each carbon in the chain will be designated with a different letter of the Greek alphabet. See the example below.



Zwitter Ion

When the hydrogen ion of carboxylic group is accepted by amino group in amino acid then Zwitter ion is formed.

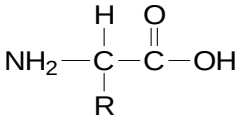
Zwitter ion is a German word, its meaning is two charged ion, i.e. both positive and negative charge. This ion is electrically neutral. In solution amino acid, exist in following ionic form.



zwitter ion

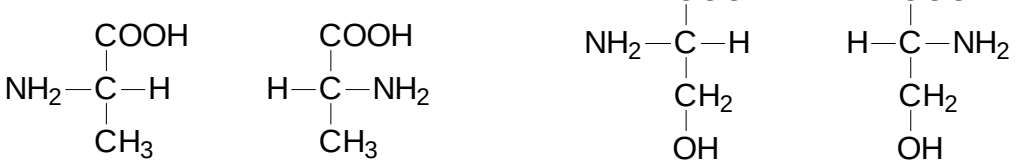
Amino acid as chiral compounds

A chiral compound must contain a carbon that is bonded to four different atoms/groups. If you see the main structure of amino acids(except glycine) as below you will see that, each structure is chiral around the carbon with the R group.



Each amino acid will come in two structural formats, called an L and a D. To determine L and D form of amino acid, it should look to the location of the

hydrogen on the chiral carbon. If the hydrogen is on the left, then the amine group is on the right, this is the D form. If the hydrogen is on the right, then the amine group is on the left, this is the L form. as the below :



The importance of chiral compounds is that their chemical reactivity, in our bodies enzyme only metabolize the D form that mean the enzymes used in the metabolism of amino acids are built to fit this D form but not the L form.

General Formula

The general formula of amino acid is
R- NH2CH-COOH
Where,
R = alkyl Group
NH2 = Amino Group
COOH = Carboxylic Group

Classification of Amino Acids

On the basis of number of carboxylic group (-COOH) and amino group (-NH2). Amino acids are classified into the following three classes.

1. **Neutral Amino Acids**: Those amino acids, which have equal number of -COOH and -NH2 groups, are called Neutral Amino Acids. Example : Glycine (CH2NH2COOH) , Alanine (CH3-CH-NH2-COOH) .

2. Acidic Amino Acids : Those amino acids, which have greater number of -COOH than -NH2 groups are called Acidic Amino Acids.
Example :Asparite [COOHCH2NH2CHCOOH] , Glutamic [HOOC-CH2-CH2-CH-NH2-COOH]

3. Basic Amino Acids: Those amino acids, which contain greater number of -NH3 than -COOH groups, are called Basic Amino Acids.
Example: Lysine H2N-(CH2)4-CH-NH2-COOH .**There is anotherclassification for amino acid (Table 1): 1-Charged**

side chains:
Glutamic acid and aspartic acid have additional carboxyl groups and usually impart a negative charge to proteins. Lysine has an 6-amino group arginine a guanidine group and histidine an imidazole group. These three basic amino acids generally impart a positive charge to proteins.

2-Polar uncharged side chains

Serine and threonine have hydroxyl group asparagine and glutamine have amide groups and cysteine has a thiol group.

3-Nonpolar aliphatic side chains:

Glycine is the simplest amino acid with no side chain. Proline is a secondary amino acid
(iminoacid). Alanine ,valine, leucine and isoleucine have hydrophobic alkyl groups.
Methionine has a thioether sulfur atom.

4-Aromatic side chains:

Phenylalanine..tyrosine and tryptophan have bulky aromatic side chains which absorb ultraviolet light.

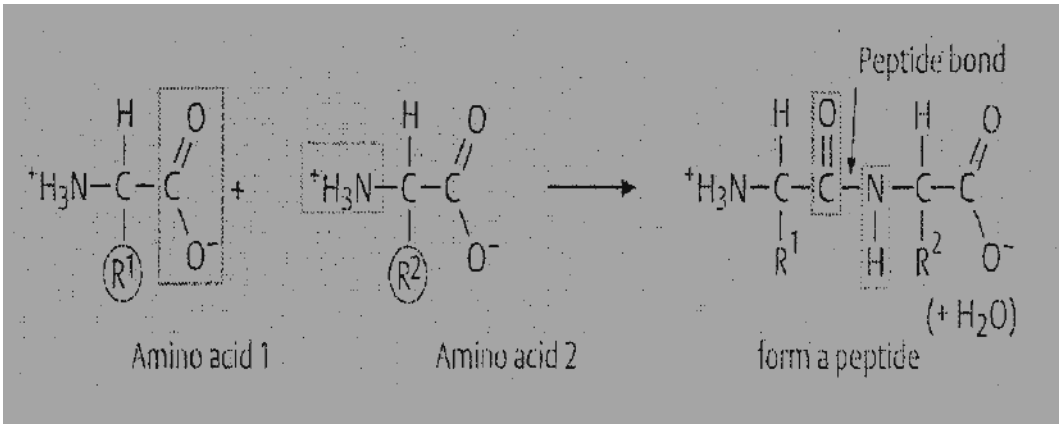
Properties of Amino Acids

The properties of amino acids are as follows.
1. Amino acids are soluble in water, but insoluble in organic solvents.
2. Amino acids are solids.
3. Amino acids have high melting point.
4. They donate or accept proton in the medium in which they are dissolved.

Function of Amino Acids

The role of amino acids in the human body is as follows.
1. Amino Acids can be synthesized back into Proteins.
2. Amino Acids oxidize to provide energy.
3. Amino Acids may be transformed into carbon hydrates or fats **Peptide Bond Or Peptide Linkage**

In the polymers of amino acids, different amino acids are linked together through a bond or linkage, which is known as peptide bond or Peptide Linkage.
When the hydrogen of amino acid combine with the OH group of other amino acid to form water, the bond between Nitrogen of one amino acid and carbon of other amino
acid is formed which is known as Peptide Bond or Linkage. This peptide bond is formed in proteins which are the polymers of amino acids.



Protein Structure Types:

Some proteins contain only one polypeptide chain while others, such as hemoglobin, contain several polypeptide chains all twisted together. The sequence of amino acids in each polypeptide or protein is unique to that protein, this is call the **primary structure**The**primary structure** is created through the linking of amino acids. This linking is accomplished by the formation of a peptide bond. This is a dehydration reaction. In other words the peptides combine and lose a water molecule.

As show below each peptide chain will have an amine end and a carboxylic acid end and each amino acid is referred to as a residue. So, the ends are named, n-terminal residue and c-terminal residue or the n-terminus and c-terminus

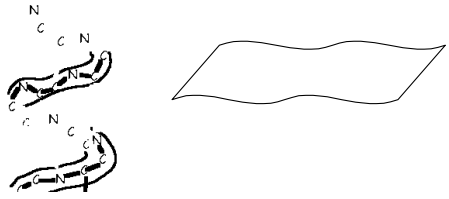
The nature of sequence of amino acids will cause the protein to have 3 shapes, 1- **Secondary structures:** This structure include (alpha-helical, beta-pleated sheet) These structures are created by molecular interactions between amino acids. Normally the interactions are hydrogen bonds.(Hydrogen bonds, in the most simple explanation, from between hydrogens attached to an oxygen or nitrogen and the lone pairs found on an oxygen or nitrogen).

2-Tertiary structure. The **tertiary structure** gives the protein its function. If the tertiary structure is deformed the protein **will not** function. The primary structure is sequenced in a way as to form the tertiary structure. The side chains of the amino acids cause them to interact with the other parts of the chain. These interactions include hydrogen bonding, hydrophobic interactions, electrostatic interactions and van der Waals forces. An egg white is all protein, when it comes

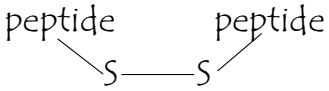
out of the shell it is clear, when you cook the egg you destroy its Tertiary Structure and the protein unfolds and becomes white, this destroys the proteins secondary, tertiary and quaternary structures.

3- Quaternary structure. The quaternary structure occurs in proteins composed of more than one peptide chain. Meaning two or more proteins come together to form one large protein. This large protein has a quaternary structure as it is composed of four myoglobin subunits. Each subunit is a separate polypeptide chain. **<u>In summery</u>:**

- Primary structure
  - amino acid sequence SNHEEVADLLAQIQ
- Secondary structure
  - alpha helices or beta sheets



- Tertiary structure
  - disulfide bonds, H-bonds, salt bridges



- Quaternary structure
  - dimers, polymers, complexes

Protein Functions: In general we have two type of protein ,function protein and structural protein ,so we can summarize the function of protein as follow:-

Type of Protein	Function	Examples
Structure	structural support	collagen in tendons and cartilage keratin in hair and nails
Contractile	muscle movement	actin, myosin, tubulin
Transportation	movement of compounds	hemoglobin carries O_2 and lipoproteins carry lipids
Hormone	chemical communication	insulin regulates blood sugar
Enzyme	Catalyze biological reactions	lactase breaks down lactose trypsin breaks down proteins
Protection	Recognized and destroy foreign substances	immunoglobulins stimulate immune system

Lecture -9

Gene expression

The expression of the genetic material occurs for the most part through theproduction of proteins, involving two consecutive steps in which the informationis converted from one form to another: transcription and translation.

Transcription

Transcription is the synthesis of RNA molecules, with DNA as a template, and it is the first step in the transfer of genetic information from genotype to phenotype. The process is complex, and requires a number of protein components.

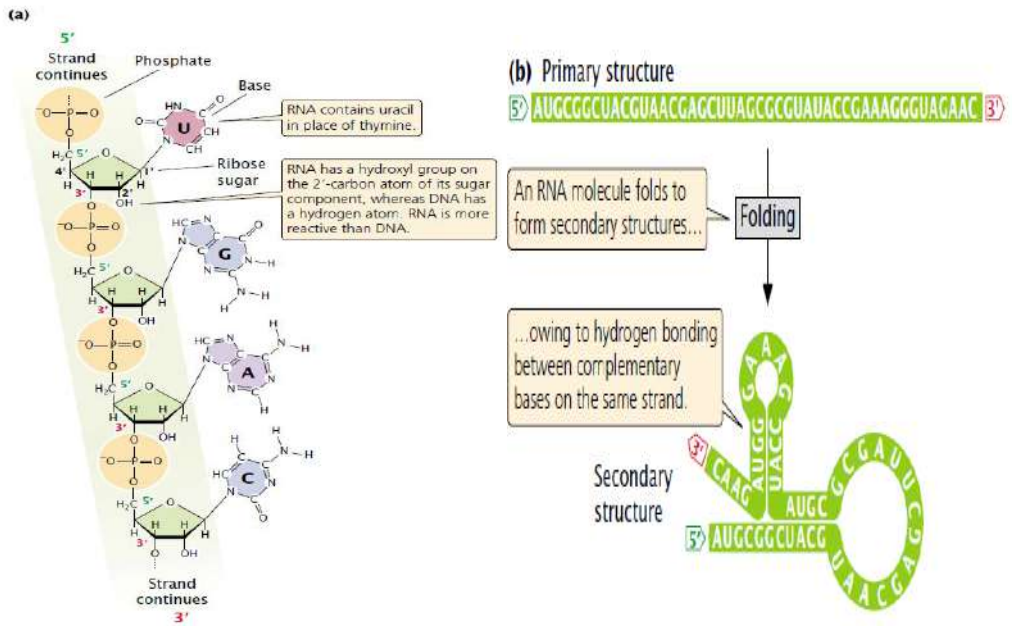
RNA Molecules

The Structure of RNA

RNA, like DNA, is a polymer consisting of nucleotidesjoined together by phosphodiesterbonds,there are severalimportant differences in the structures of DNA andRNA. Whereas DNA nucleotides contain deoxyribose sugars,RNA nucleotides have ribose sugars .With a free hydroxyl group on the 2 carbon atom of the ribose sugar, RNA is degraded rapidly under alkaline conditions. The deoxyribose sugar of DNA lacks this freehydroxyl group; so DNA is a more stable molecule. Anotherimportant difference is that thymine, one of the two pyrimidinesfound in DNA, is replaced by uracil in RNA.A final difference in the structures of DNA and RNA isthat RNA is usually single stranded, consisting of a singlepolynucleotide strand, whereas DNA normallyconsists of two polynucleotide strands joined byhydrogen bonding between complementary bases. Someviruses contain double-stranded RNA genomes Although RNA is usually singlestranded, short complementary regions within a nucleotidestrand can pair and form secondary structures .These RNA secondary structures are often calledhairpin-loops or stem-loop structures. When two regionswithin a single RNA molecule pair up, the strands in thoseregions must be antiparallel, with pairing between cytosineand guanine and between adenine and uracil (althoughoccasionally guanine pairs with uracil) (Table-1)(Figure-1).

Table -1 : Similarities and differences in DNA and RNA structures

Characteristic	DNA	RNA
Composed of nucleotides	Yes	Yes
Type of sugar	Deoxyribose	Ribose
Presence of 2′-OH group	No	Yes
Bases	A, G, C, T	A, G, C, U
Nucleotides joined by phosphodiester bonds	Yes	Yes
Double or single stranded	Usually double	Usually single
Secondary structure	Double helix	Many types
Stability	Quite stable	Easily degraded



Figure(1) :**RNA has a primary(a) and a secondary structure(b)**

Classes of RNA

RNA molecules perform a variety of functions in the cell (Table-2) .

1-Ribosomal RNA (rRNA), along with ribosomal proteinsubunits, makes up the ribosome, the site of protein assembly.

2-Messenger RNA (mRNA): carries the codinginstructions for polypeptide chains from DNA to the ribosome.After attaching to a ribosome, an mRNA moleculespecifies the sequence of the amino acids in apolypeptide chain and provides a template for joiningamino acids. Bacterial cells do not possess premRNA;in these cells, transcription takes place concurrentlywithtranslation.

3-Transfer RNA (tRNA): serves as the link between thecoding sequence of nucleotides in the mRNA and theamino acid sequence of a polypeptide chain. Each tRNAattaches to one particular type of amino acid and helps toincorporate that amino acid into a polypeptide chain.

Table.2. The different classes of RNA molecules

Class of RNA	Cell Type	Location of Function* in Eukaryotic Cells	Function
Ribosomal RNA (rRNA)	Bacterial and eukaryotic	Cytoplasm	Structural and functional components of the ribosome
Messenger RNA (mRNA)	Bacterial and eukaryotic	Nucleus and cytoplasm	Carries genetic code for proteins
Transfer RNA (tRNA)	Bacterial and eukaryotic	Cytoplasm	Helps incorporate amino acids into polypeptide chain
Small nuclear RNA (snRNA)	Eukaryotic	Nucleus	Processing of pre-mRNA
Small nucleolar RNA (snoRNA)	Eukaryotic	Nucleus	Processing and assembly of rRNA
Small cytoplasmic RNA (scRNA)	Eukaryotic	Cytoplasm	Variable

Transcription:Synthesizing RNAfrom a DNA Template

All cellular RNAs are synthesized from a DNA templatethrough the process of transcription .Transcription is in many ways similar to the process ofreplication, but one fundamental difference relates to the length of the template used. During replication, all thenucleotides in the DNA template are copied, but, duringtranscription, only small parts of the DNA moleculeusually a single gene or, at most, a few genes are transcribedintoRNA. Because not all gene products are neededat the same time or in the same cell, it would be highlyinefficient for a cell to constantly transcribe all of its genesFurthermore, much of the DNA does not code for a functionalproduct, and transcription of such sequences wouldbe pointless. Transcription is, in fact, a highly selectiveprocessindividual genes are transcribed only as theirproducts are needed. But this selectivity imposes a fundamentalproblem on the cell—the problem of how to recognizeindividual genes and transcribe them at the propetime and place.

1-The first step is the conversion of the information into messenger RNA (mRNA).
2- This process (transcription) is carried out by RNA polymerase.
3-The RNA strand is made in the 5‘ to 3‘ direction.

There are major differences between transcription and replication:
1-Firstly, only a comparatively short molecule is produced,
2-secondly, only one of the DNA strands is transcribed.
3-Since only a single strand is made, it can be produced continuously using a single enzyme; there is no need for lagging strand synthesis.
4-In addition the production of relatively short single-stranded RNA causes fewer topological problems: the enzyme and the RNA product can essentially rotate around the helix, so there is no need for the helicases and topoisomerases that are essential for replication.
5- RNA polymerase can start synthesis without needed any primer.
Transcription is therefore considerably simpler than replication. Since transcription results in the synthesis of comparatively short mRNA molecules there must be a large number of signals around the chromosome that direct the RNA polymerase to start transcription at the required place and to stop when the block of genes has been transcribed.

The start signals (promoters) also convey the information as to the direction in which transcription should proceed.

In E. coli, depending on growth conditions, (2000–5000) copies of RNA polymerase may be engaged on mRNA synthesis at any time.

- Transcription requires three majorcomponents:

1. DNA template.
2. The raw materials (substrates) needed to build a newRNA molecule;
3. The transcription apparatus, consisting of the proteinsnecessary to catalyze the synthesis of RNA.

- **Transcribe: to make an RNA that is complementary to a region within a strand of DNA.**
- **The transcription unit**

A **transcription unit** is a stretch ofDNA that codes for an RNA molecule and the sequences necessaryfor its transcription (Figure – 2).

Included within a transcription unit are three criticalregions:

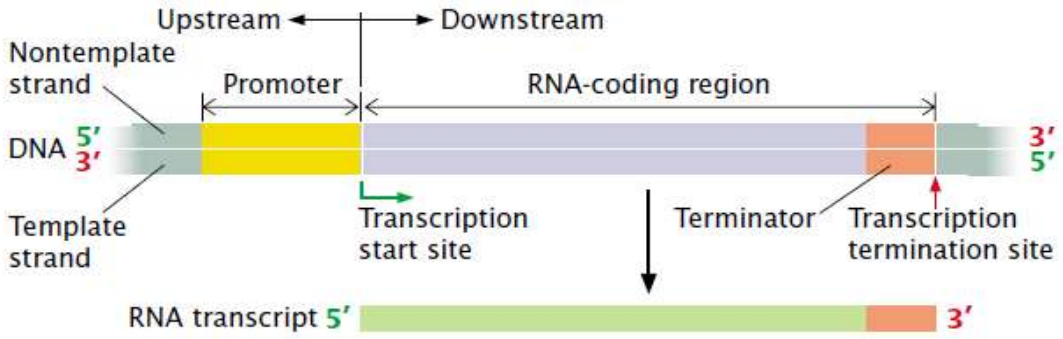
1-A promoter

2-RNA coding sequence,

3- Terminator

- **The promoter**is a DNA sequencethat the transcription apparatus recognizes and binds. Itindicates which of the two DNA strands is to be read as thetemplate and the direction of transcription. The promoteralso determines the transcription start site, the firstnucleotide that will be transcribed into RNA. In most transcriptionunits, the promoter is located next to the transcriptionstart site but is not, itself, transcribed.
- **The RNA-coding region,** a sequence of DNA nucleotidesthat is copied into an RNA molecule.
- **Terminator,** a sequence ofnucleotides that signals where transcription is to end.Terminators are usually part of the coding sequence; that is,transcription stops only after the terminator has beencopied into RNA.

The terms ***upstream*** **a**nd***downstream***to refer to the direction of transcription andthe location of nucleotide sequences surrounding the RNAcoding sequence.



Figure(2):A transcription unit includes a promoter, an RNA-coding and a terminator.

The Substrate for Transcription

RNA is synthesized from **ribonucleosidetriphosphosphates**(rNTPs). In synthesis, nucleotidesare added one at a time to the 3_-OH group of the growingRNA molecule. Two phosphates are cleaved from the incoming ribonucleoside triphosphate; the remaining phosphateparticipates in a phosphodiester bond that connectsthe nucleotide to the growing RNA molecule. Nucleotides are always added to the 3′end of the RNAmolecule, and the direction of transcription is therefore5′−3′, the same as the direction of DNAsynthesis during replication. RNA is made complementaryand antiparallel to one of the DNA strands (the templatestrand).

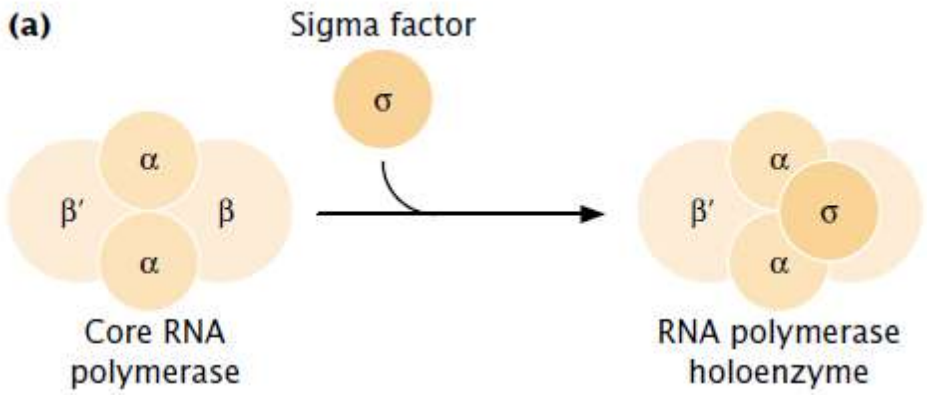
$$RNA_n + rNTP \longrightarrow RNA_{n+1} + PP_i$$

The Transcription Apparatus

A single enzyme—**RNA polymerase**—carries out all the required steps of transcription.The action ofRNA polymerase is enhanced by a number of accessoryproteins that join and leave the polymerase at different stages of the process. Each accessory protein is responsiblefor providing or regulating a special function.

- Bacterial RNA polymerase Bacterial cells typically possessonly one type of RNA polymerase, which catalyzes thesynthesis of all classes of bacterial RNA: mRNA, tRNA, andrRNA.
- Bacterial RNA polymerase is a large, multimericenzyme (meaning that it consists of several polypeptidechains).
- Bacterial RNA polymerase are foursubunits (individual polypeptide chains) that make up the**Core Enzyme:(α2β β’): ** two copies of a subunit called alpha ,single copy of beta and single copy of beta prime .

- The core enzyme catalyzes the elongationof the RNA molecule by the addition of RNA nucleotides.Other functional subunits join and leave thecore enzyme at particular stages of the transcriptionprocess.
- The **sigma factor** controls the binding of theRNA polymerase to the promoter. Without sigma, RNApolymerase will initiate transcription at a random pointalong the DNA. After sigma has associated with the coreenzyme forming a **Holoenzyme: α2ββ'σ**.



The Process of Transcription

Transcription can be conveniently divided intothree stages:

1. Initiation, in which the transcription apparatusassembles on the promoter and begins the synthesisof RNA.

2.Elongation, in which RNA polymerase moves along theDNA, unwinding it and adding new nucleotides, one ata time, to the 3′ end of the growing RNA strand.

3. Termination, the recognition of the end of thetranscription unit and the separation of the RNA molecule from the DNA template.

1. Initiation

Initiation includes all the steps necessary to begin RNA synthesis,including;(1) promoter recognition, (2) formation ofthe transcription bubble, (3) creation of the first bonds between rNTPs, and (4) escape of the transcription apparatusfrom the promoter.

Bacterial promoters Essential information for the transcription Unitwhere it will start transcribing, whichstrand is to be read, and in what direction the RNA polymerase will move is imbedded in the nucleotide sequenceof the promoter. **Promoters are sequences in the DNA thatare recognized by the transcription apparatus and arerequired for transcription to take place.** In bacterial cells,promoters are usually adjacent to an RNA coding sequence(is located just upstream of the start site)(figure-3).

a) Promoter recognition: closed complex

- **Promoters possess a consensus sequence**

• **Typical *E. coli* promoter sequence ,**some times we call it the **Pribnow box**

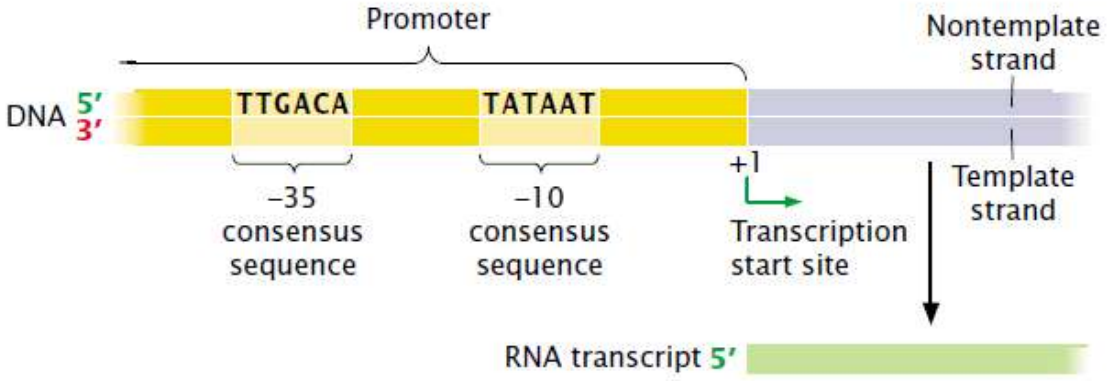
-10 box: TATAAT
-35 box: TTGACA

• **σsubunit**

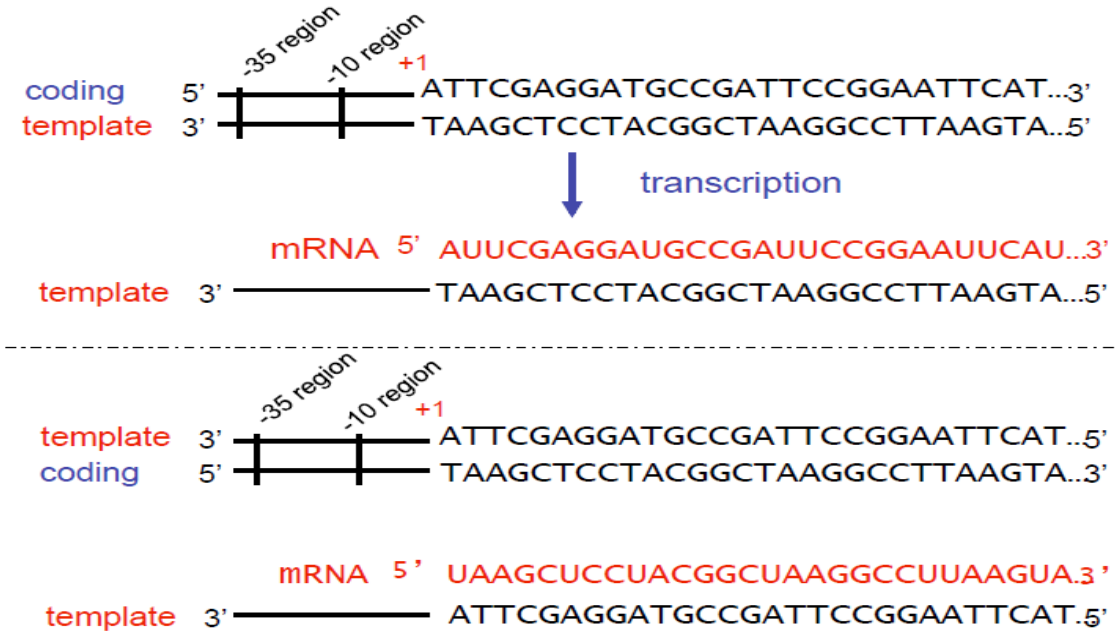
b) Unwinding ofpromoter region:open complex

begins atthe +1 site
• σ subunit I released oncetranscriptsynthesis isinitiated

To begin the synthesis of an RNA molecule, RNApolymerase pairs the base on a ribo nucleoside tri phosphate with its complementary base at the start site on theDNA template strand .No primer is requiredto initiate the synthesis of the 5′end of the RNAmolecule. Two of the three phosphates are cleaved from the ribo nucleoside tri phosphate as the nucleotide is addedto the 3′ end of the growing RNA molecule. However, becausethe 5′end of the first ribo nucleoside tri phosphate does not take part in the formation of a phosphodiesterbond, all three of its phosphates remain. An RNA moleculetherefore possesses, at least initially, three phosphates at its 5′end (Figure -4).



Figure(3): In bacterial promoters, consensus sequences are found upstream of the start site,approximately at positions_10 and _35.



.

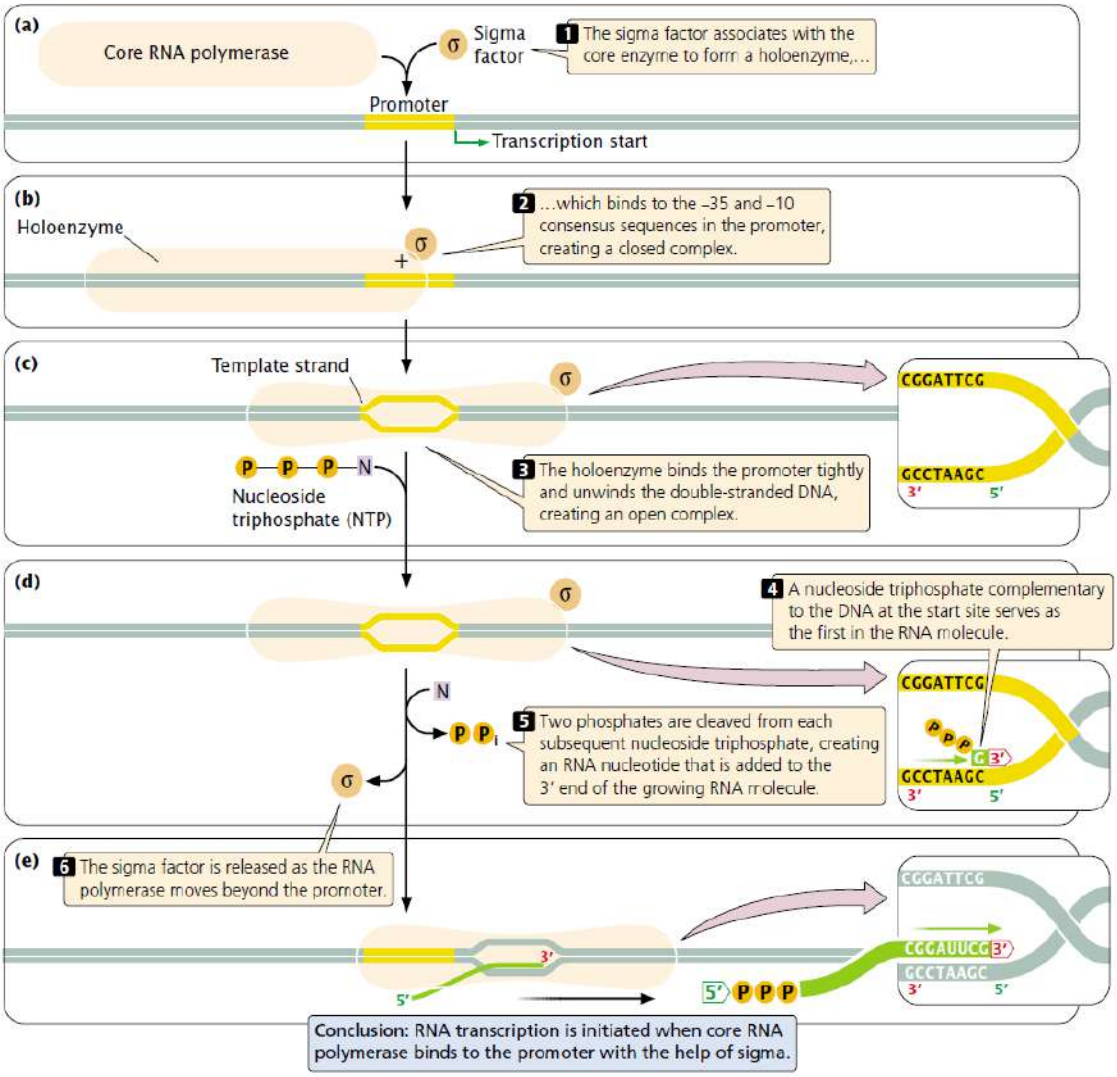


Figure-4: Transcription in bacteria is carried out by RNA polymerase,which must bind to the sigma factor to initiate transcription.

2-Elongation

After initiation, RNA polymerase moves downstreamalong the template, progressively unwinding the DNA atthe leading (downstream) edge of the transcription bubble,joining nucleotides to the RNA molecule according to

the sequence on the template, and rewinding the DNA atthe trailing (upstream) edge of the bubble. In bacterial cells at 37°C, about 40 nucleotides are added per second.This rate of RNA synthesis is much lower than that ofDNA synthesis, which is more than 1500 nucleotides per second in bacterial cells Transcription takes place within a short stretch

ofabout 18 nucleotides of unwound DNA—the transcriptionbubble. Within this region, RNA is continuously synthesizedwith single-stranded DNA used as a template. About8 nucleotides of newly synthesized RNA are paired with theDNA-template nucleotides at any one time. As the transcription apparatus moves down the DNA template, it generates positive supercoiling ahead of the transcriptionbubble and negative supercoiling behind it. Topoisomeraseenzymes probably relieve the stress associated with theunwinding and rewinding of DN in transcription, as they do in DNA replication.

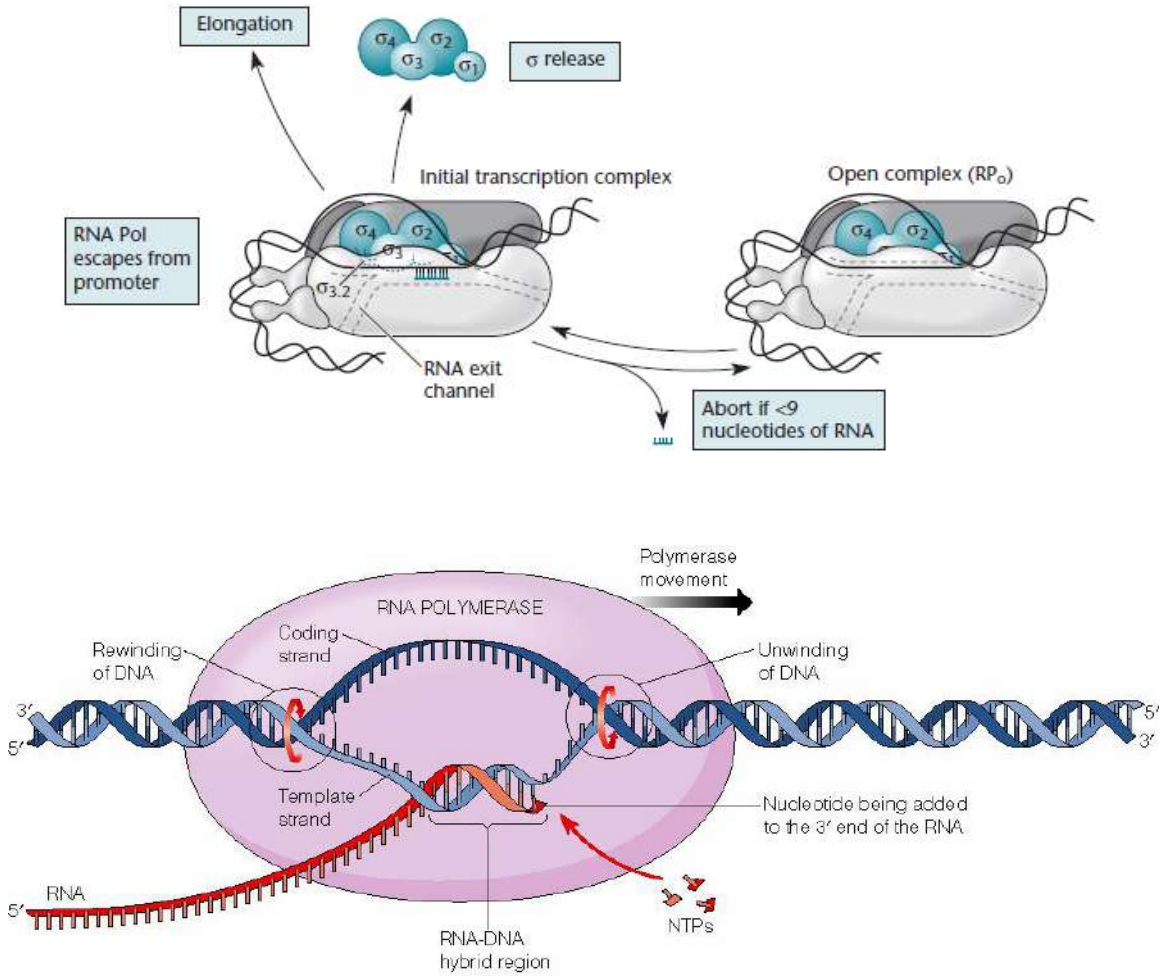


Figure-5: Elongation stage

3-Termination

RNA polymerase moves along the template, adding nucleotidesto the 3′end of the growing RNA molecule until ittranscribes a terminator. Most terminators are foundupstream of the point of termination. Rather,transcription ends after the terminator has been transcribed. At the terminator, RNA polymerasemust stop synthesizing RNA, the RNA molecule must bereleased from RNA polymerase, the newly made RNA moleculemust dissociate fully from the DNA, and RNA polymerasemust detach from the DNA template.

Bacterial cells possess two major types of terminators:

A-Rho-independent terminatorsare able to cause the end of transcription in the absenceof rho.

B-Rho-dependent terminators are able to cause the termination of transcription only in the presence of an ancillary protein called the rho factor.

A-Rho-independent terminators have two common features.

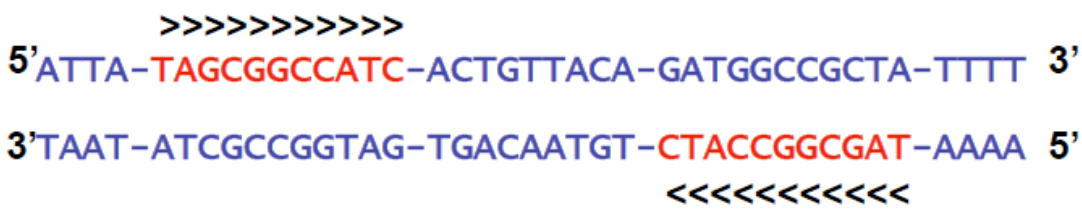
1-First, they contain **inverted repeats** (sequences ofnucleotides on one strand that are inverted and complementary).When inverted repeats have been transcribed intoRNA, a hairpin secondary structure forms .

2-Second, in rho-independent terminators, a string ofapproximately six adenine nucleotides follows the secondinverted repeat in the template DNA. Their transcriptionproduces a string of uracil nucleotides after the hairpin inthe transcribed RNA.

So Transcription stops at terminator sequences

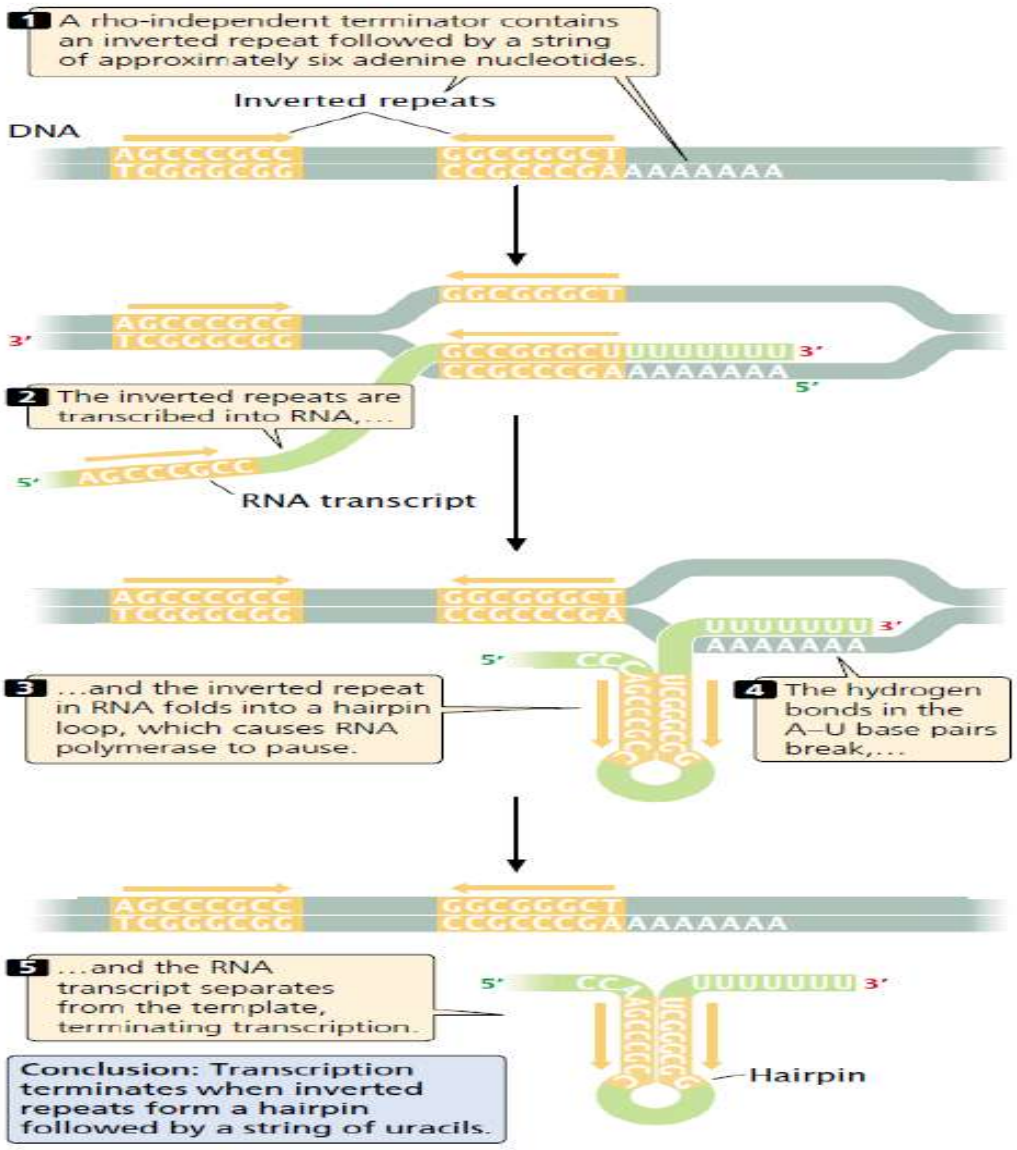
• Inverted repeats:

•Stretch of AAAAAAA…. on template strand



The presence of a hairpin in an RNA transcript causesRNA polymerase to slow down or pause, which creates anopportunity for termination. The adenine–uracil base

pairingsdownstream of the hairpin are relatively unstable comparedwith other base pairings, and the formation of thehairpin may itself destablize the DNA–RNA pairing, causingthe RNA molecule to separate from its DNA template.When the RNA transcript has separated from the template,RNA synthesis can no longer continue (Figure-6).



Figure(6): Termination by bacterial rho-independent terminators is a multistep process.

B- Rho-dependent terminators have two features:

(1) DNAsequences that produce a pause in transcription;

(2) A DNA sequence that encodes a stretch of RNA upstream ofthe terminator that is devoid of any secondary structures. This unstructured RNA serves as binding site for the rhoprotein, which binds the RNA and moves toward its 3_ end,following the RNA polymerase (Figure-7) .When RNApolymerase encounters the terminator, it pauses, allowing rho to catch up. The rho protein has helicase activity, whichit uses to unwind the RNA–DNA hybrid in the transcription bubble, bringing an end to transcription.

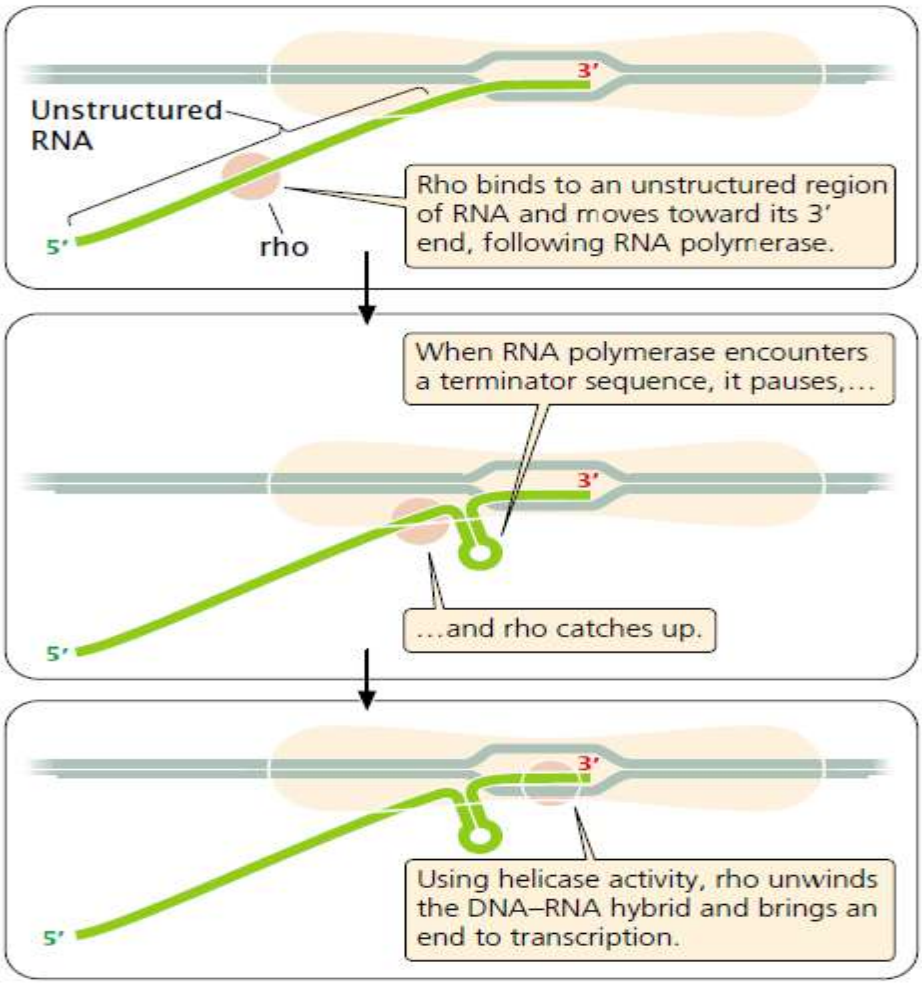


Figure-7: The termination of transcription in some bacterial genes requires the presence of the rho protein.

The Process of Eukaryotic Transcription

The process of eukaryotic transcription is similar to that of bacterial transcription. Eukaryotic transcription also includes initiation, elongation, and termination, and the basic principles of transcription already outlined apply to eukaryotic transcription. However, there are some important differences. Bacterial cells possess a single type of RNA polymerase, consisting of a core enzyme and other subunits that participate in various stages of transcription. Eukaryotic cells possess three different types of RNA polymerases, each of which transcribe a different class of RNA and recognize a different type of promoter. And this included RNA polymeras I transcribes rRNA; RNA polymerase II transcribes pre-mRNA, snoRNAs, and some snRNAs; and RNA polymerase III transcribes tRNAs, small rRNAs, an some snRNAs

Table 13.3 Eukaryotic RNA polymerases

Type	Transcribes
RNA polymerase I	Large rRNAs
RNA polymerase II	Pre-mRNA, some snRNAs, snoRNAs
RNA polymerase III	tRNAs, small rRNA, snRNAs

Characteristics of Eukaryotic Promoters and Transcription Factorsidentify some general principles of eukaryotic promotersand transcription factors:

1. Several types of DNA sequences take part in the initiation of transcription in eukaryotic cells. These sequences generally serve as the binding sites for proteins that interact with RNA polymerase and influence the initiation of transcription.
2. Some sequences that affect transcription, called promoters, are adjacent to or within the RNA coding region and are relatively fixed with regard to the start site of transcription. Promoters consist of a core promoter located adjacent to the gene and a regulatory promoter located farther upstream.
3. Other sequences, called enhancers, are distant from the gene and function independently of position and direction. Enhancers stimulate transcription.
4. General transcription factors bind to the core promoter near the start site and, with RNA polymerase, assemble into a basal transcription apparatus. The TATA-binding protein (TBP) is a critical transcription factor that positions the active site of RNA polymerase over the start site.

5. Transcriptional activator proteins bind to sequences inthe regulatory promoter and enhancers and affecttranscription by interacting with the basal transcription apparatus.
6. Proteins binding to enhancers interact with the basaltranscription apparatus by causing the DNA betweenthe promoter and the enhancer to loop out, bringingthe enhancer into close proximity to the promoter.

The different eukaryotic RNA polymerases utilizedifferent mechanisms of termination.

Concepts

General transcription factors assemble into thebasal transcription apparatus, which binds to DNAnear the start site and is necessary fortranscription to take place at minimal levels.Additional proteins called transcriptional activatorsbind to other consensus sequences in promotersand enhancers, and affect the rate of transcription.

RNA polymerase I promoters have two keycomponents: (1) the core element, whichsurrounds the start site and is sufficient toinitiate transcription, and (2) the upstream controlsequence, which increases the efficiency of thecore promoter.

Concepts

Two classes of DNA sequences in eukaryotic cellsaffect transcription: enhancers and promoters.A promoter is near the gene and has a fixedposition relative to the start site of transcription.An enhancer can be distant from the gene andvariable in location.
The three types of RNA polymerase in eukaryotic cellsrecognize different types of promoters, all of which have consensus sequences that serve as binding sites for transcription factors.

Promoters possess a consensus sequence
• **promoter sequence ,**sometimes we call it the **Hokneesbox or Goldberg box**

-25box: 5'– TATAAAT- 3'
-75 box: 5' - GGNNCAATCT-3'

Lecture 10:

The modification and splicing of pre –mature RNA:

The Structure of Messenger RNA

Messenger RNA functions as the template for protein synthesis; it carries genetic information from DNA to

a ribosome and helps to assemble amino acids in their correct order. Each amino acid in a protein is specified by a set of three nucleotides in the mRNA, called a **codon.** Both prokaryotic and eukaryotic mRNAs contain three primary regions **(Figure -1)**. The **5_ untranslated region** (5_ UTR; sometimes call the leader) is a sequence of nucleotides at the 5_ end of the mRNA that does not code for the amino acid sequence of a protein. In bacterial mRNA, this region contains a consensus sequence called the **Shine-Dalgarno sequence,** which serves as the ribosome binding site during translation; it is found approximately

seven nucleotides upstream of the first codon translated into an amino acid (called the start codon). Eukaryotic mRNA has no equivalent consensus sequence in its 5_ untranslated region. In eukaryotic cells, ribosomes bind to a modified 5_ end of mRNA, **region,** which comprises the codons that specify the amino acid sequence of the protein. The protein-coding region begins with a start codon and ends with a stop codon. The last region of mRNA is the **3_ untranslated region** (3_ UTR), a sequence of nucleotides at the 3_ end of mRNA that is not translated into protein. The 3_ untranslated region affects the stability of mRNA and the translation of the mRNA protein-coding sequence.

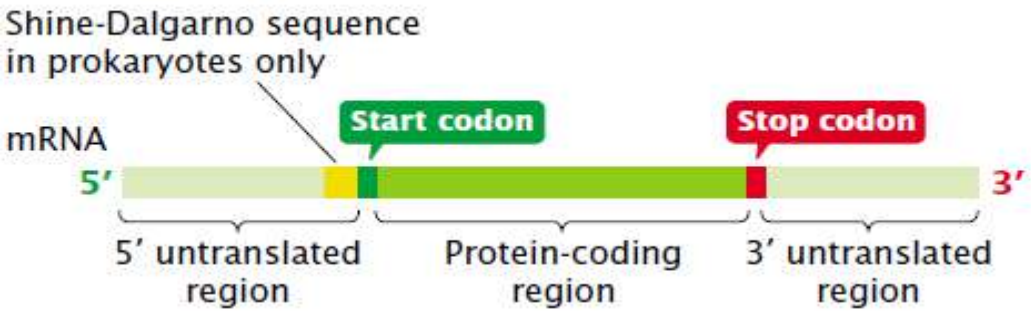


Figure – 1: Three primary regions of mature mRNA are the 5_ untranslated region, the protein-coding region, and the 3_ untranslated region.

Pre-mRNA Processing

In bacterial cells, transcription and translation take place simultaneously; while the 3_ end of an mRNA is undergoing transcription, ribosomes attach to the Shine-Dalgarno sequence near the 5_ end and begin translation. Because transcription and translation are coupled, there is little opportunity for the bacterial mRNA to be modified before protein synthesis. In contrast, transcription and translation are separated in both time and space in eukaryotic cells. Transcription takes place in the nucleus, whereas most translation takes place in the cytoplasm; this separation provides an opportunity for eukaryotic RNA to be modified before it is translated. Indeed, eukaryotic mRNA is extensively altered after transcription. Changes are made to the 5_ end, the 3_ end, and the protein-coding section of the RNA molecule. The initial transcript of protein-encoding genes of eukaryotic cells is called pre-mRNA, whereas the mature, processed transcript is mRNA. We will reserve the term mRNA for RNA molecules that have been completely processed and are ready to undergo translation.

The process of modification and splicing of pre-mature RNA in Eukaryotic Cells

The process of modification of pre-mature RNA:

1-The Addition of the 5ʿ Cap

Almost all eukaryotic pre-mRNAs are modified at their 5ʿ ends by the addition of a structure called a 5ʿ **cap.** This capping consists of the addition

of an extra nucleotide at the 5ʿ end of the mRNA and methylation by the addition of a methyl group (CH3) to the base in the newly added neucleotide and to the 2ʿ OH group of the sugar of one or more nucleotides at the 5ʿ end **(figure – 2)**.

The function of 5ʿcap :

1-Capping takes place rapidly after the initiation of transcription

2-the 5ʿcap functions in the initiation of translation. Cap binding proteins recognize the cap and attach to it; a ribosome then binds to these proteins and moves downstream along the mRNA until the start codon is reached and translation begins.

3-The presence of a 5ʿ cap also increases the stability of mRNA and influences the removal of introns.

three phosphates are present at the 5ʿ end of all RNA molecules, because phosphates are not cleaved.

from the first ribonucleoside triphosphate in the transcription reaction. The 5′end of pre-mRNA can be represented as 5′–pppNpNpN . . . , in which the letter N represents a ribonucleotide and p represents a phosphate. Shortly after the initiation of transcription, one of these phosphates is removed and a guanine nucleotide is added (Figure 2). This guanine nucleotide is attached to the pre-mRNA by a unique 5ʿ–5ʿ bond, which is quite different from the usual 5ʿ–3ʿ phosphodiester bond that joins all the other nucleotides in RNA. One or more methyl groups are then added to the 5ʿ end; the first of

these methyl groups is added to position 7 of the base of the terminal guanine nucleotide, making the base 7- methylguanine. Next, a methyl group may be added to the 2ʿ position of the sugar in the second and third nucleotides, Rarely, additional methyl groups may be attached to the bases of the second and third nucleotides of the pre-mRNA.

CONCEPT :

The cap consists of a nucleotide with 7-methyl guanine attached to the pre-mRNA by a unique 5ʿ–5ʿ bond . The cap is added shortly after the initiation of transcription. A methyl group is added to position 7 of the guanine base of the newly added (now the terminal) nucleotide and to the 2ʿposition of each sugar of the next two nucleotides

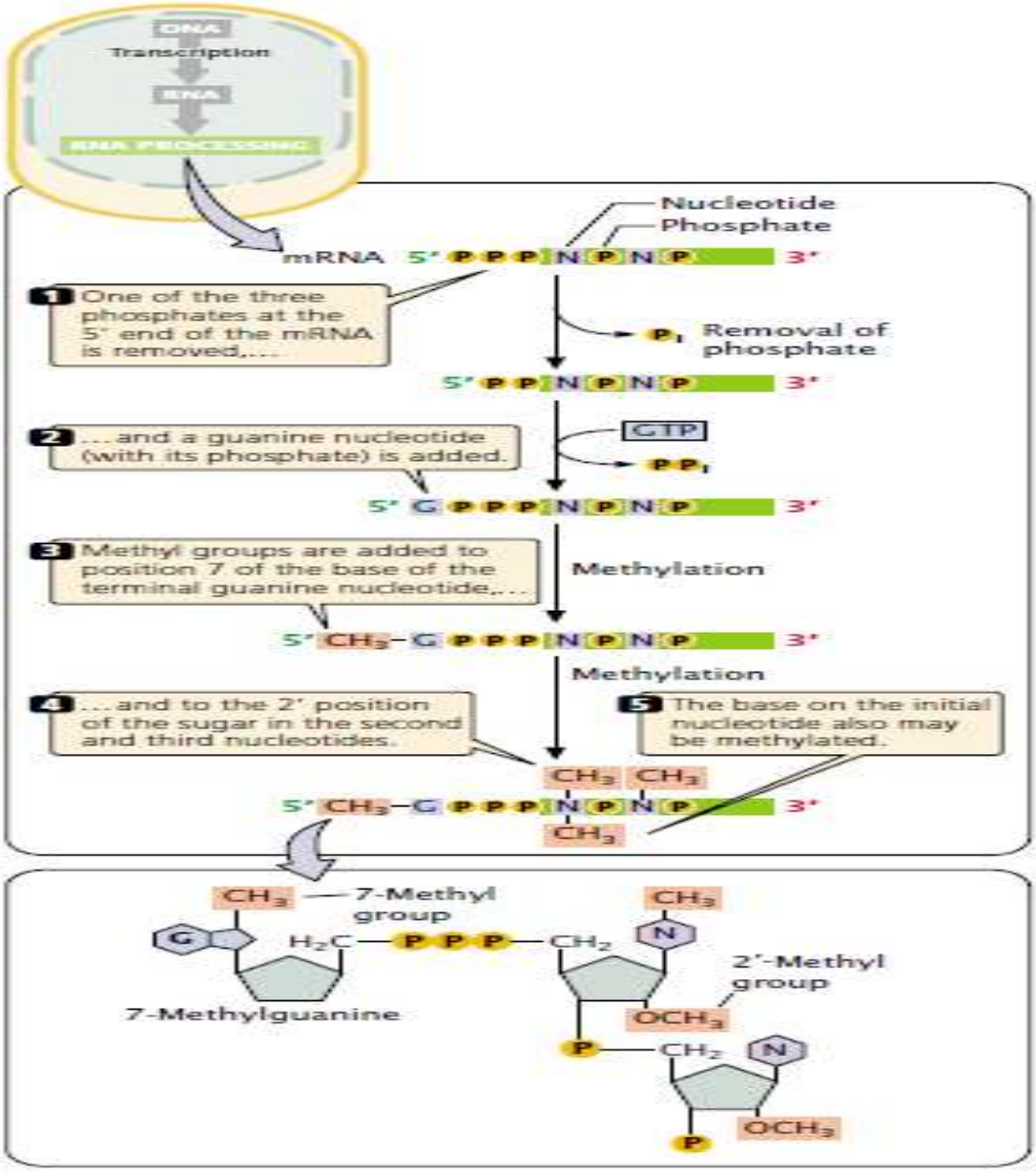


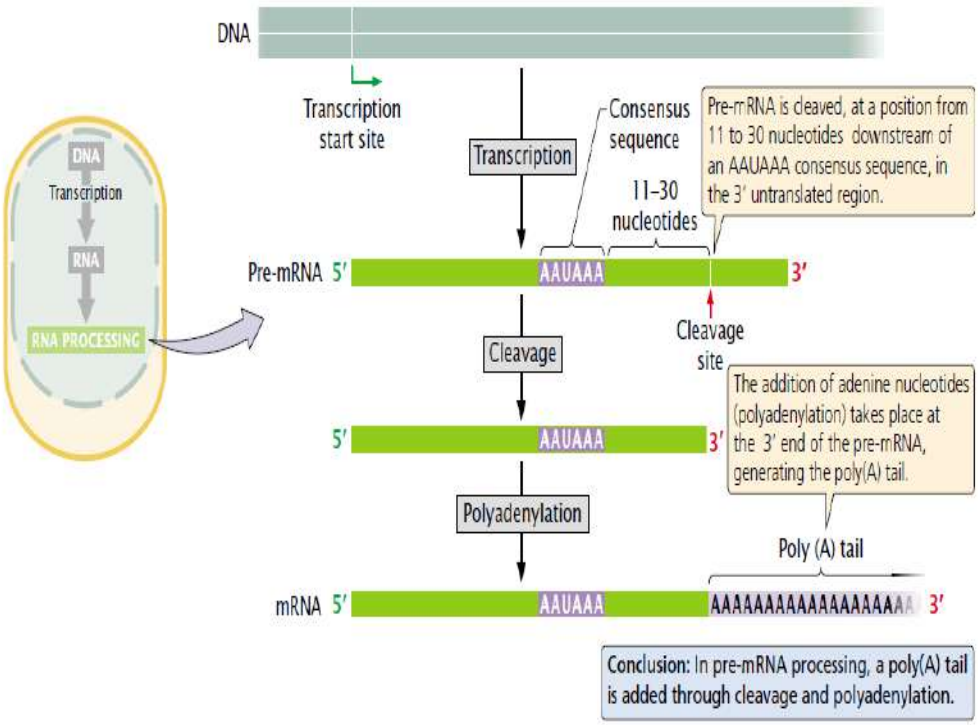
Figure -2: Most eukaryotic mRNAs have a 5ˋ cap.

2-The Addition of the Poly(A) Tail

Most mature eukaryotic mRNAs have from 50 to 250 adenine nucleotides at the 3‘ end (a **poly(A) tail**). These nucleotides are not encoded in the DNA but are added after transcription **(figure-3)** in a process termed polyadenylation.

Many eukaryotic genes transcribed by RNA polymerase II are transcribed well beyond the end of the coding sequence , the extra material at the 3‘ end is then cleaved and the poly(A) tail is added. For some pre-mRNA molecules, more than 1000 nucleotides may be cleaved from the 3‘ end. Processing of the 3‘ end of pre-mRNA requires sequences both upstream and downstream of the cleavage site .

The consensus sequence AAUAAA is usually from 11 to 30 nucleotides upstream of the cleavage site and determines the point at which cleavage will take place. A sequence rich in Us (or Gs and Us) is typically downstream of the cleavage site.



The process of splicing :

Different types of methods for splicing of pre mRNA ,tRNA and rRNA.
In this Lecture we focus on one method for splicing per mRNA.
Before splicing takes place, an upstream exon (exon 1) and a downstream exon (exon 2) are separated by an intron (Figure-3).

Exon: coding region
Intron :non coding region

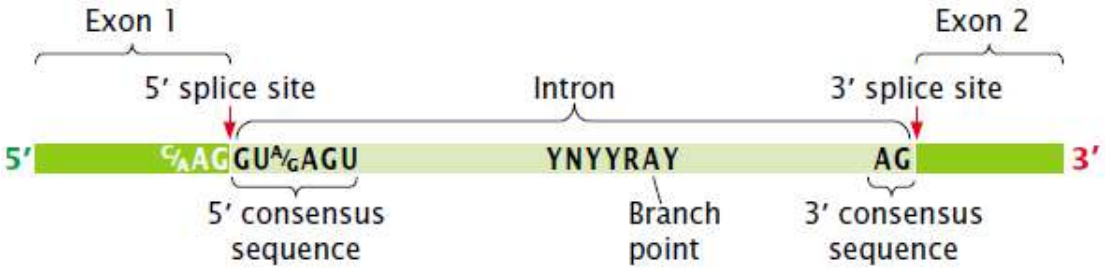


Figure-3:Splicing of pre-mRNA requires consensus sequences. In the consensus sequence surrounding the branch point (YNYYRAY) Y is any pyrimidine, R is any purine, A is adenine, and N is any base.

Pre-mRNA is spliced in two distinct steps.

1- In the first step, the pre-mRNA is cut at the 5‘ splice site. This cut frees exon 1 from the intron, and the 5‘ end of the intron attaches to the branch point; that is, the intron folds back on itself, forming a structure called a **lariat.** The guanine nucleotide in the consensus sequence at the 5‘ splice site bonds with the adenine nucleotide at the branch point. This bonding is accomplished through **transesterification,** a chemical reaction in which the OH group on the 2‘carbon atom of the adenine nucleotide at the branch point attacks the 5‘ phosphodiester bond of the guanine nucleotide at the 5‘splice site, cleaving it and forming a new 5‘–2‘ phosphodiester bond between the guanine andadenine nucleotides.

2-In the second step of RNA splicing, a cut is made at th3‘splice site and, simultaneously, the 3‘ end of exon 1 becomes covalently attached (spliced) to the 5‘ end of exon 2. This bond also forms through a transesterification reaction, in which the 3‘-OH group attached to the end of exon 1 attacks the phosphodiester bond at the 3‘ splice site, cleaving it and forming a new phosphodiester bond between the 3‘ end of exon 1 and the 5‘ end of exon 2; the intron is released as a lariat. The intron becomes linear when the bond breaks at the branch point and is then rapidly degraded by nuclear enzymes. The mature mRNA consisting of the exons spliced together is exported to the cytoplasm where it is translated.(Figure-4).

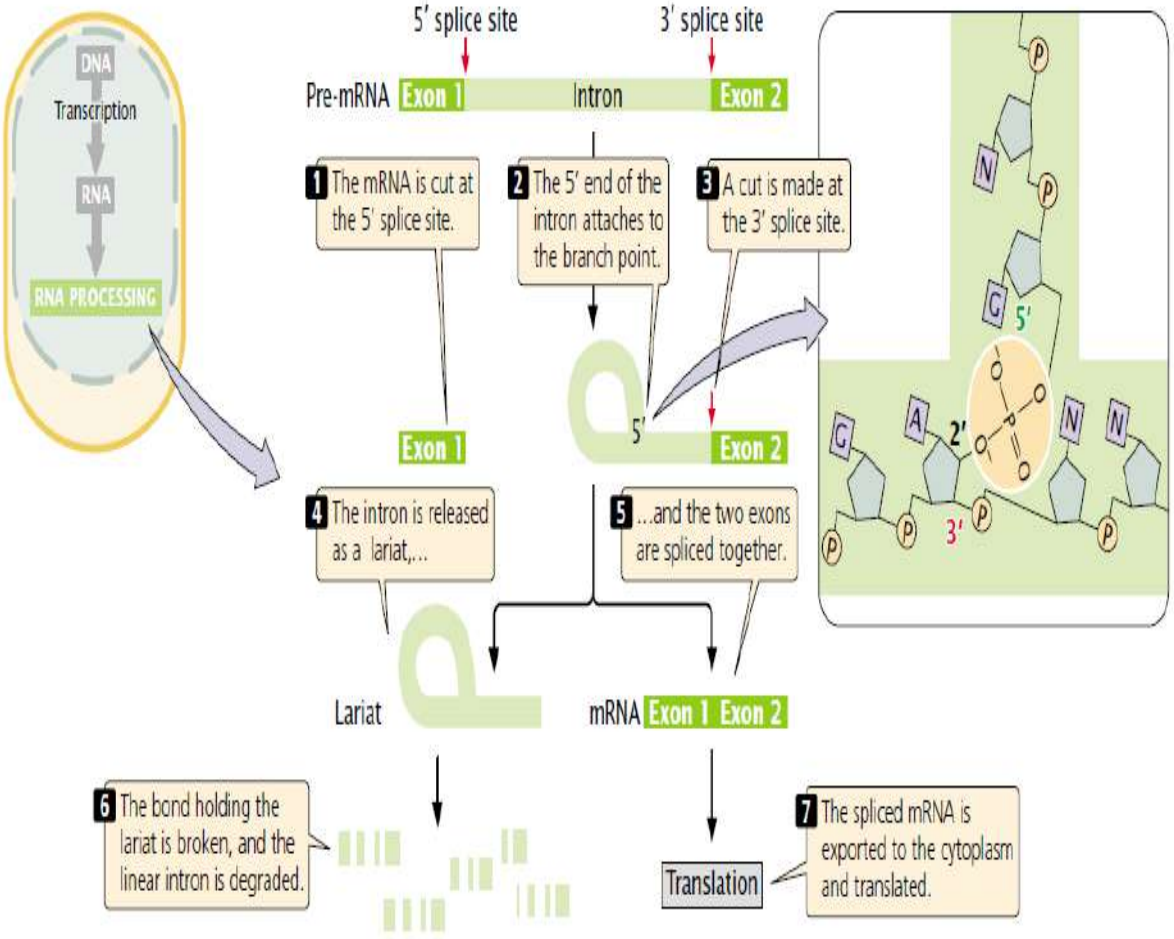
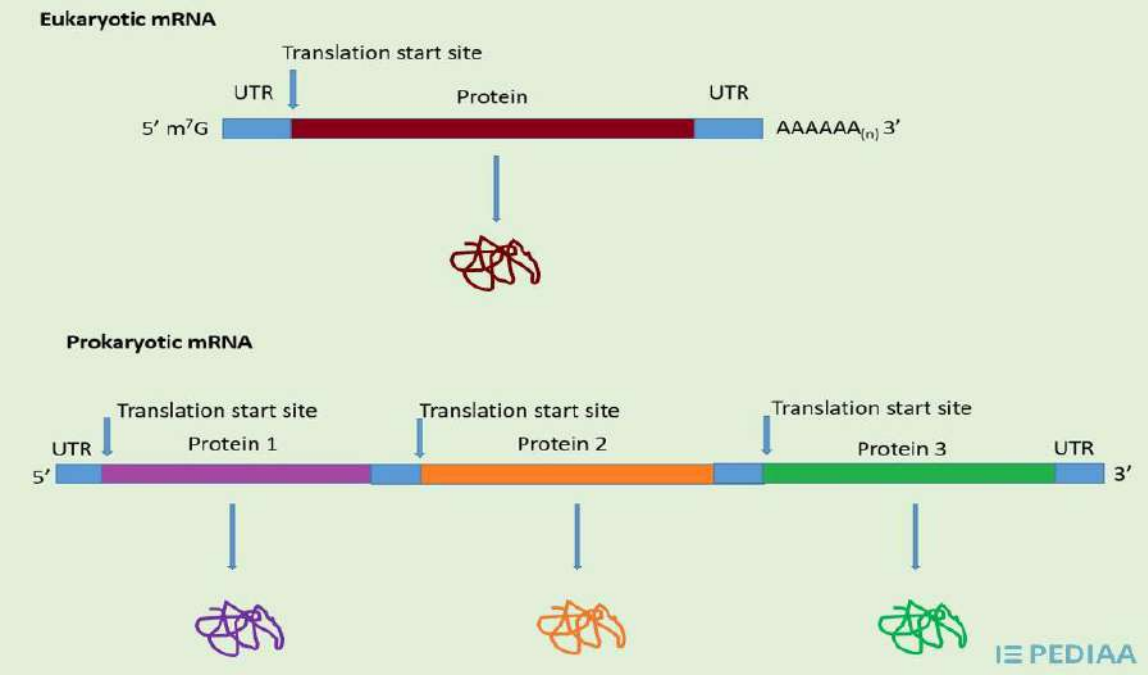
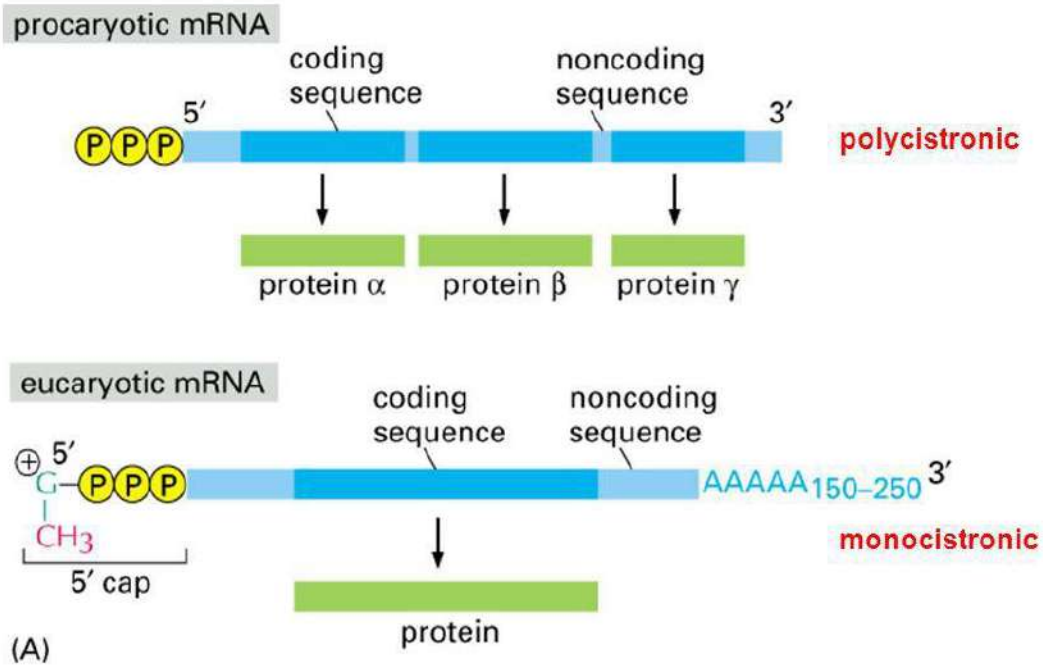


Figure-4:The splicing of nuclear introns requires a two-step process.



Comparison between procaryotic and eucaryotic mRNA

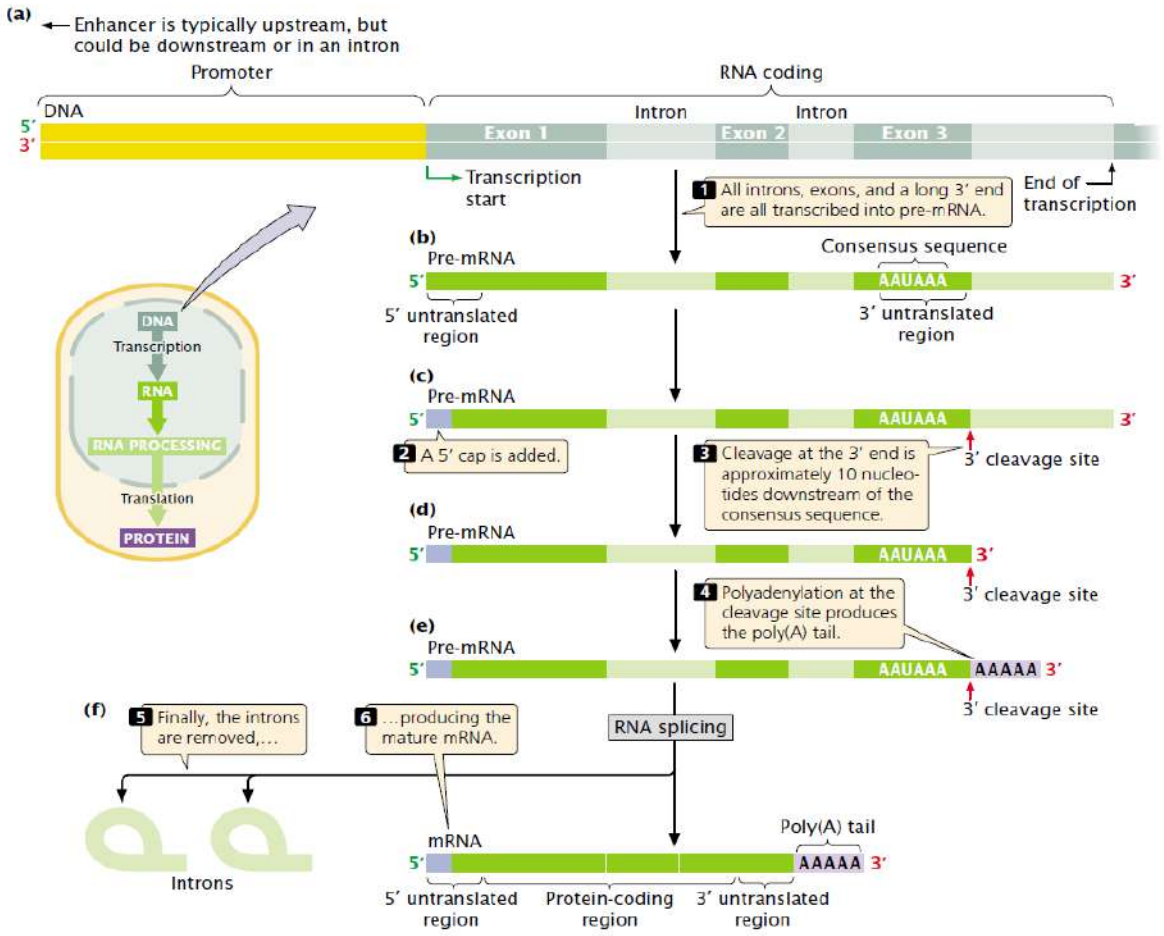


cistron Segment of DNA (or RNA) that encode a single polypeptide chain

monocistronic mRNA :mRNA carrying the information of a single cistron, that is a coding sequence for only a single protein

polycistronic Mrna: mRNA carrying the information of multiple cistrons, that is coding sequences for several proteins

Why add cap and poly A:
1-protect mRNA from degradation by nuclease enzymes
2-Facilitate translation and binds proteins important for function



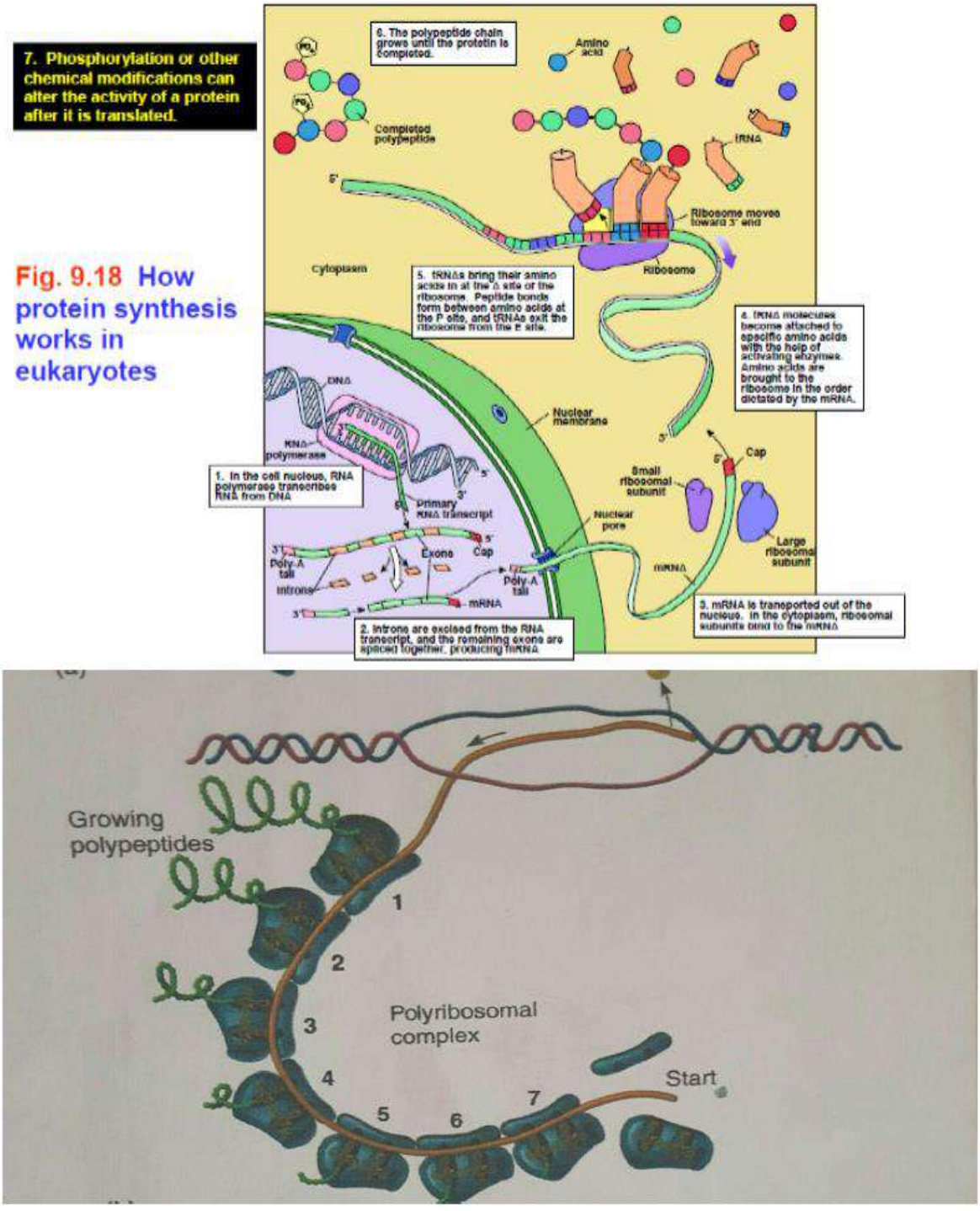
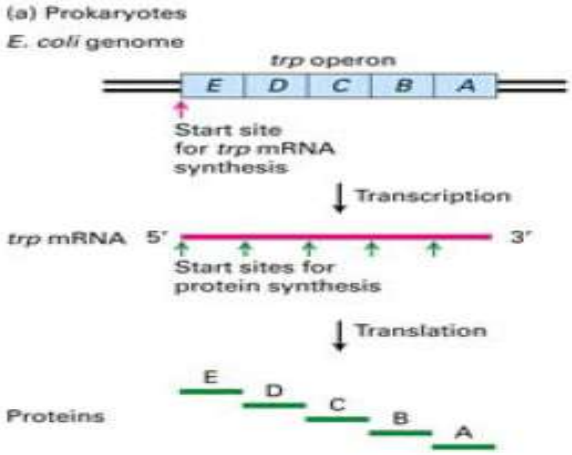


Fig. 9.18 How protein synthesis works in eukaryotes

Figure: Describe that transcription and translation separated in eukaryotes (figure above). transcription and translation coupled in prokaryotes(figure below)

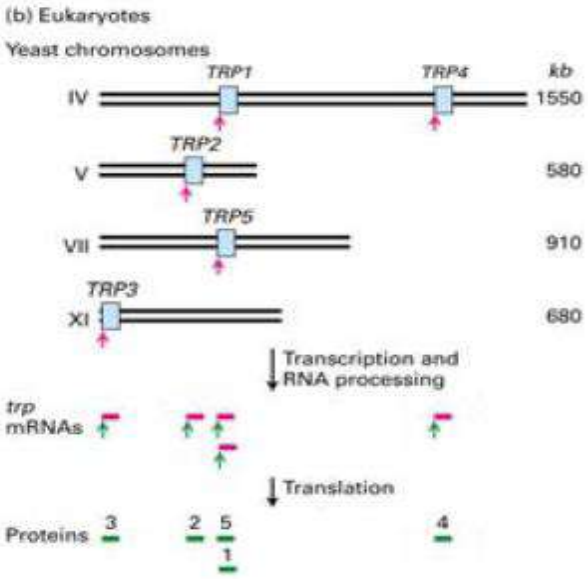
Prokaryotic Transcription

- Polycistronic
- No cap
- No poly(A) tail
- No introns
- Coupled transcription-translation



Eukaryotic Transcription

- Monocistronic (*with interesting exceptions!*)
- Capped
- Polyadenylated
- Spliced
- Compartmentalized



Lecture : 11

The Translation process

RNA - Translation - Protein

The Translation:

Mean the translation of the sequence of nucleotides in mRNA into the sequence of amino acids in a protein occurs on the ribosome.
Ribosomes as moving protein-synthesizing machines.
we will focus primarily on bacterial translation Because more is known about translation in bacteria, eukaryotic translation is similar, although there are some significant differences that will be noted as we proceed through the stages of translation.
Four Components of the Translation System included:

A-Ribosomes:

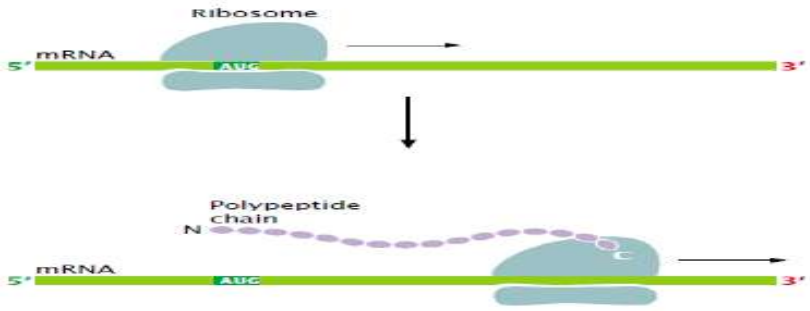
- Proteins
- rRNAs

B- tRNAs and tRNA synthetases:

C-mRNAs

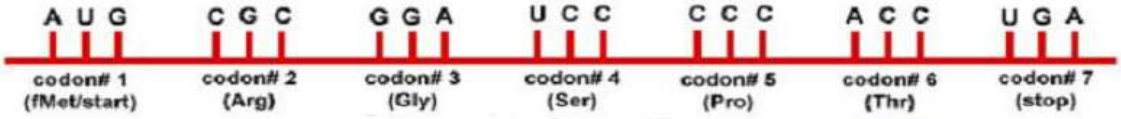
D-Accessory proteins:

- Initiation factors
- Elongation factors
- Release factors

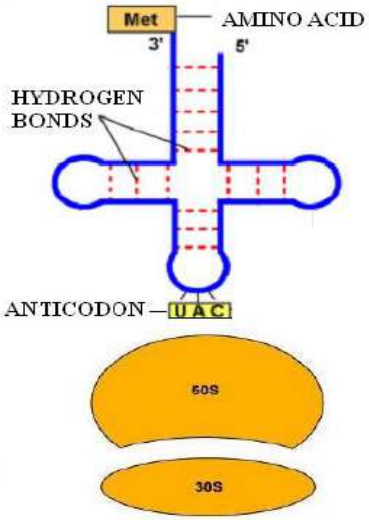


There are three types of RNA :

TYPES OF RNA



- **mRNA (MESSENGER RNA)** – SINGLE HELIX, COPY OF A DNA GENE, CODONS CALL FOR AMINO ACID IN PROTEIN
- **tRNA (TRANSFER RNA)** – "CLOVER LEAF-SHAPED", ANTICODON ON ONE END (COMPLEMENT TO CODON), AMINO ACID ON OTHER END, ADDS AMINO ACID TO PROTEIN. (tRNA IS RECYCLED BY AMINOACYL-TRNA SYNTHETASE)
- **RIOBOSOME** (50-SUBUNIT & 30-SUBUNIT) ORGANELLE THAT MAKES PROTEIN



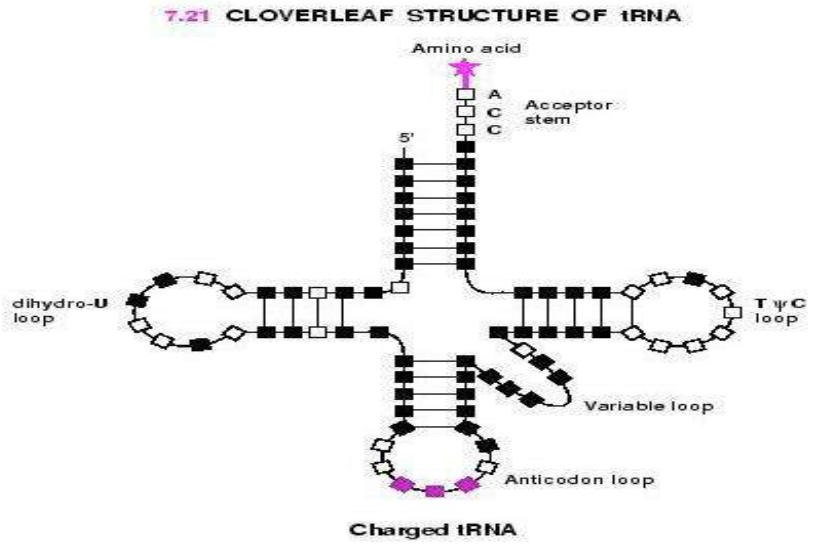


Table 14.3 Composition of ribosomes in bacterial and eukaryotic cells

Cell Type	Ribosome Size	Subunit	rRNA Component	Proteins
Bacterial	70S	Large (50S)	23S (2900 nucleotides) 5S (120 nucleotides)	31
		Small (30S)	16S (1500 nucleotides)	21
Eukaryotic	80S	Large (60S)	28S (4700 nucleotides) 5.8S (160 nucleotides) 5S (120 nucleotides)	49
		Small (40S)	18S (1900 nucleotides)	33

Major Steps in Translation

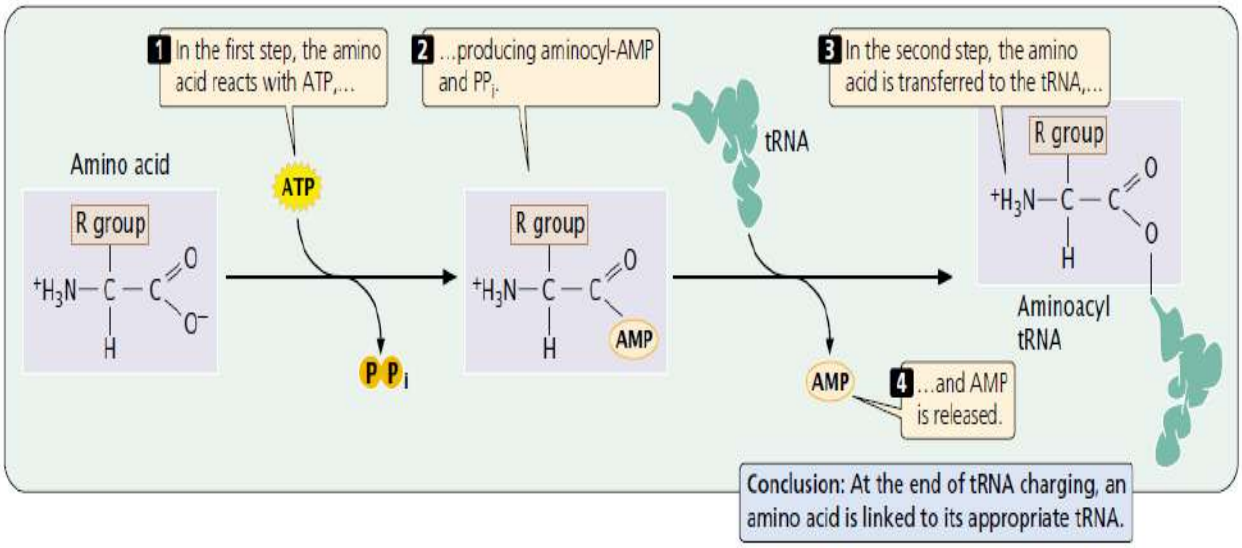
Protein synthesis can be divided into four stages:
(1) The binding of amino acids to the tRNAs;
(2) Initiation, in which the components necessary for translation are assembled at the ribosome;
(3) Elongation, in which amino acids are joined, one at a time, to the growing polypeptide chain.
(4) Termination, in which protein synthesis halts at the termination codon and the translation components are released from the ribosome.

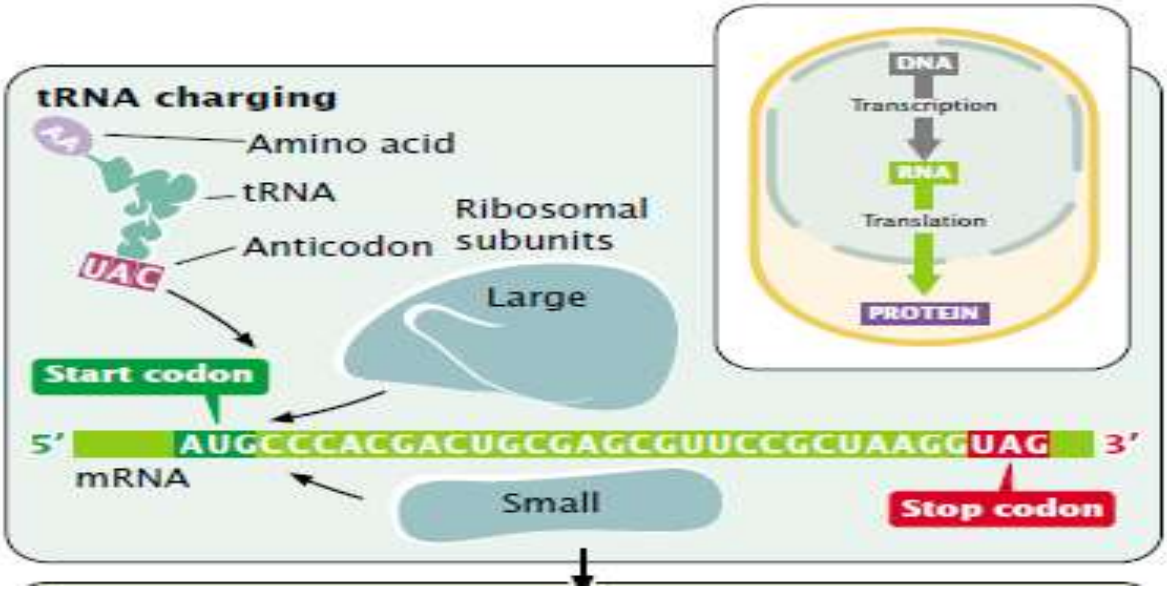
Charging the tRNA with the Amino Acid

For each tRNA there is a specific enzyme that recognizes both the tRNA and the corresponding amino acid.These enzymes, known as **aminoacyl tRNA synthetases**, attach the amino acid to the tRNA.This is called charging the tRNA. Empty tRNA is known as **uncharged tRNA** while tRNA with its amino acid is **charged tRNA**. Charging occurs in two steps (Fig. 8.06). First the amino acid reacts with ATP to form aminoacyl-AMP (also known as aminoacyl-adenylate). Next the aminoacyl-group is transferred to the 3¢-end of the tRNA.

a) amino acid + ATP → aminoacyl-AMP + PPi

b) aminoacyl-AMP + tRNA → aminoacyl-tRNA + AMP





Initiation.

The second stage in the process of protein synthesis is initiation. During initiation, all the components necessary for protein synthesis assemble: (1) mRNA; (2) the small and large subunits of the ribosome; (3) a set of three proteins called initiation factors(IF-1, IF-2, and IF-3); (4) initiator tRNA with *N*-formylmethionine attached (fMet-tRNAfMet); and (5) guanosine triphosphate (GTP). Initiation comprises three major steps. First, mRNA binds to the small subunit of the ribosome. Second, initiator tRNA binds to the mRNA through base pairing between the codon and anticodon. Third, the large ribosome joins the initiation complex

At this point, the initiation complex consists of (1) the small subunit of the ribosome; (2) the mRNA; (3) the initiator tRNA with its amino acid (fMet-tRNAfMet); (4) one molecule of GTP; and (5) IF-3, IF-2, and IF-1. These components are collectively known as the **30S initiation complex** . In the final step of initiation, IF-3 dissociates from the small subunit, allowing the large subunit of the ribosome to join the initiation complex. The molecule of GTP (provided by IF-2) is hydrolyzed to guanosine diphosphate (GDP), and IF-1 and IF-2 depart . When the large subunit has joined the initiation complex, it is called the **70S initiation complex.** Similar events take place in the initiation of translation
in eukaryotic cells, but there are some important differences.
In bacterial cells, sequences in 16S rRNA of the small subunit of the ribosome bind to the Shine-Dalgarno sequence in mRNA; this binding positions the ribosome over the start codon. No analogous consensus sequence exists in eukaryotic mRNA. Instead, the cap at the 5_ end of
eukaryotic mRNA plays a critical role in the initiation of translation. The small subunit of the eukaryotic ribosome, with the help of initiation factors, recognizes the cap and binds there; the small subunit then migrates along (scans) the mRNA until it locates the first AUG codon. The identification of the start codon is facilitated by the presence of a consensus sequence (called the Kozak sequence) that surrounds the start codon:

Another important difference is that eukaryotic initiation requires more initiation factors. Some factors keep the ribosomal subunits separated, just as IF-3 does in bacterial cells. Others recognize the 5_ cap on mRNA and allow th small subunit of the ribosome to bind there. Still others possess RNA helicase activity, which is used to unwind secondary structures that may

exist in the 5_ untranslated region of mRNA, allowing the small subunit to move down the mRNA until the initiation codon is reached. Other initiation factors help bring the initiator tRNA and methionine (Met-tRNAfMet) to the initiation complex. The poly(A) tail at the 3_ end of eukaryotic mRNA also plays a role in the initiation of translation. Proteins that attach to the poly(A) tail interact with proteins that bind to the 5_ cap, enhancing the binding of the small subunit of the ribosome to the 5_ end of the mRNA. This interaction between the 5_ cap and the 3_ tail suggests that the mRNA bends backward during the initiation of translation, forming a circular structure **(FIGURE 15.21)**. A few eukaryotic mRNAs contain internal ribosome entry sites, where ribosomes can bind directly without first attaching to the 5_ cap.

.

Concept:

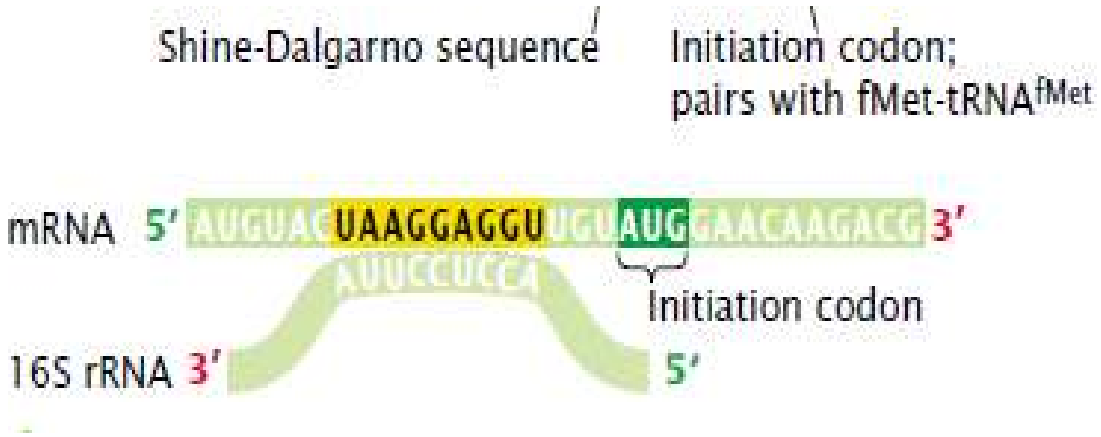
In the initiation of translation in bacterial cells, the small ribosomal subunit attaches to mRNA, and initiator tRNA attaches to the initiation codon. This process requires several initiation factors (IF-1, IF-2, and IF-3) and GTP. In the final step, the large ribosomal subunit joins the initiation complex

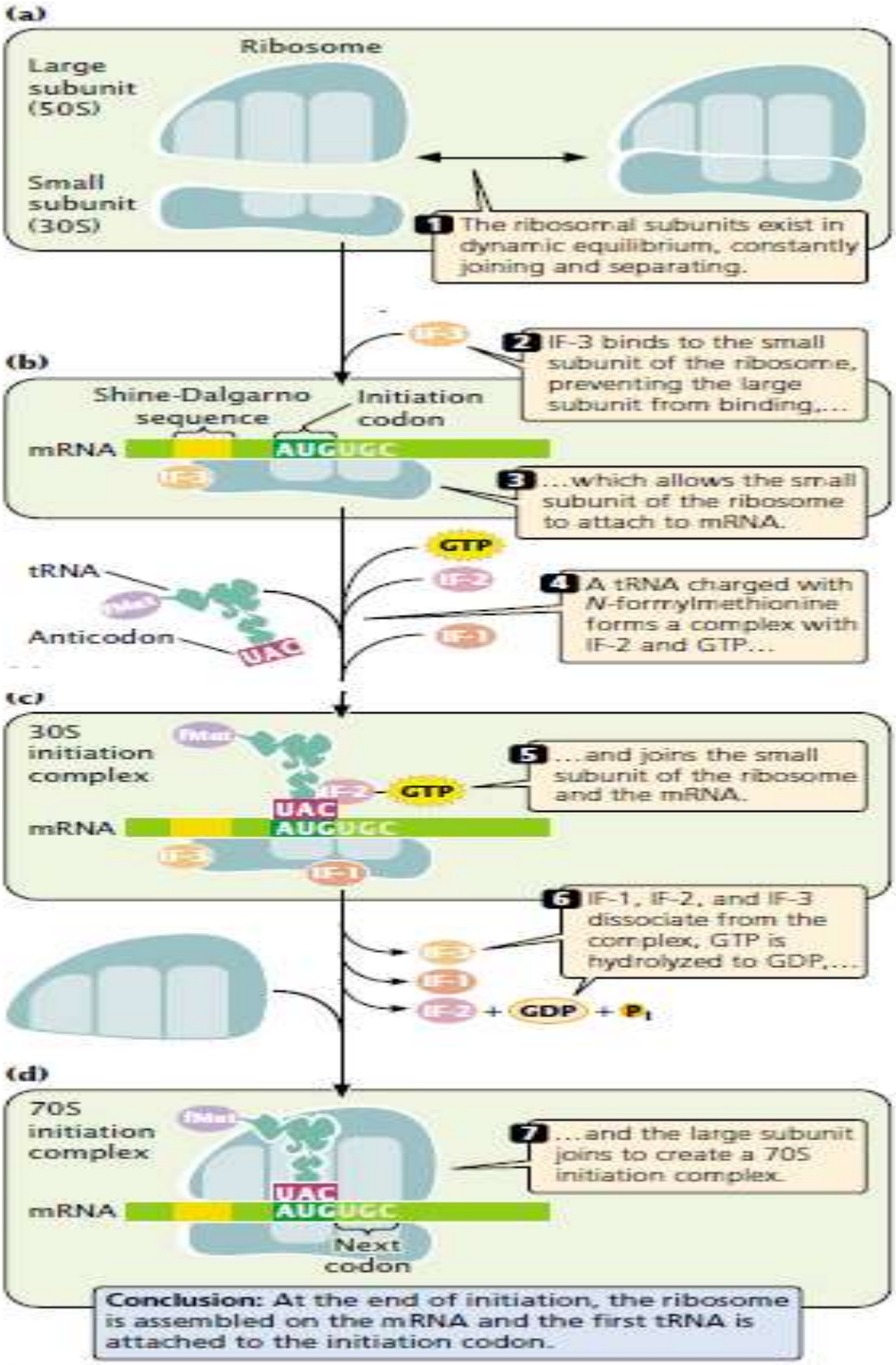
Terms:

Shine-Dalgarno (S-D) sequence Same as RBS; sequence close to the front of mRNA that is recognized by the ribosome; only found in prokaryotic Cells.

anti-Shine-Dalgarno sequence Sequence on 16S rRNA that is complementary to the Shine-Dalgarno sequence of mRNA

initiator tRNA The tRNA that brings the first amino acid to the ribosome when starting a new polypeptide chain
N-formyl-methionine or fMet Modified methionine used as the first amino acid during protein synthesis in bacteria
ribosome binding site (RBS) Same as Shine-Dalgarno sequence; sequence close to the front of mRNA that is recognized by the ribosome; only found in prokaryotic cells
Shine-Dalgarno (S-D) sequence Same as RBS; sequence close to the front of mRNA that is recognized by the ribosome; only found in prokaryotic cells





15.19 The initiation of translation in bacterial cells requires several initiation factors and GTP.

Elongation

The next stage in protein synthesis is elongation, in which amino acids are joined to create a polypeptide chain.
Elongation requires
(1) the 70S complex just described;
(2) tRNAs charged with their amino acids;
(3) several elongation factors (EF-Ts, EF-Tu, and EF-G); and (4) GTP.

A ribosome has three sites that can be occupied by tRNAs; the **aminoacyl,** or **A, site,** the **peptidyl,** or **P, site,** and the **exit,** or **E, site** . The initiator tRNA immediately occupies the P site (the only site to which the fMet-tRNAfMet is capable of binding), but all other tRNAs first enter the A site. After initiation, the ribosome is attached to the mRNA, and fMet-tRNAfMet is positioned over the AUG start codon in the P site; the adjacent
A site is unoccupied .

Elongation occurs in three steps.
1- The first step is the delivery of a charged tRNA (tRNA with its amino acid attached) to the A site. This requires the
presence of **elongation factor Tu** (EF-Tu), **elongation factor Ts** (EF-Ts), and GTP. EF-Tu first joins with GTP and then binds to a charged tRNA to form a three-part complex. This three-part complex enters the A site of the ribosome, where the anticodon on the tRNA pairs with the codon on the mRNA. After the charged tRNA is in the A site, GTP is cleaved to GDP, and the EF-Tu–GDP complex is released **.** Factor EF-Ts regenerates EF-Tu–GDP to EF-Tu–GTP. In eukaryotic cells, a similar set of reactions delivers the charged tRNA to the A site.
2-The second step of elongation is the creation of a peptide bond between the amino acids that are attached to tRNAs in the P and A sites . The formation
of this peptide bond releases the amino acid in the P site from its tRNA. The activity responsible for peptidebond formation in the ribosome is referred to as **peptidyl transferase.** For many years, the assumption was that this
activity is carried out by one of the proteins in the large subunit of the ribosome. Evidence, however, now indicates that the catalytic activity is a property of the rRNA in the large subunit of the ribosome; this rRNA acts as a ribozyme .

3-The third step in elongation is **translocation,** the movement of the ribosome down the mRNA in the 5_:3_ direction. This step positions the ribosome over the next codon and requires **elongation factor G** (EF-G) and the hydrolysis of GTP to GDP.

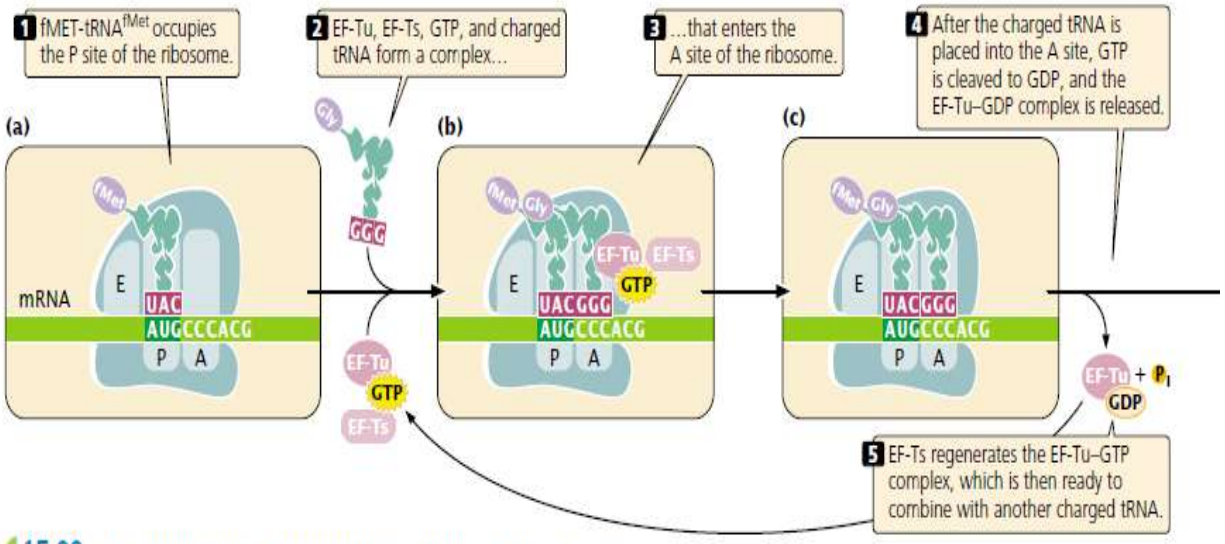
Because the tRNAs in the P and A site are still attached to the mRNA through codon– anticodon pairing, they do not move with the ribosome as it translocates. Consequently, the ribosome shifts so that the tRNA that previously occupied the P site now occupies the E site, from which it moves into the cytoplasm where it may be recharged with another amino acid. Translocation also causes the tRNA that occupied the A site (which is attached to the growing polypeptide chain) to be in the P site, leaving the A site open. Thus, the progress of each tRNA through the ribosome during elongation can be summarized as follows: cytoplasm : A site : P site : E site : cytoplasm. As discussed earlier, the initiator tRNA is an exception: it attaches directly to the P site and never occupies the A site.

After translocation, the A site of the ribosome is empty and ready to receive the tRNA specified by the next codon.

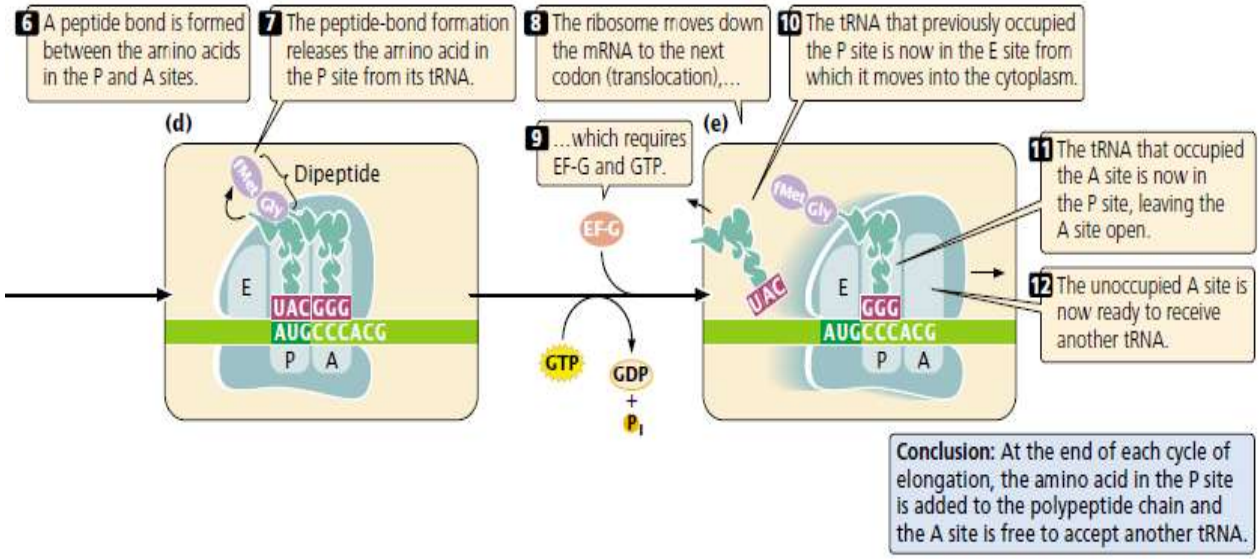
The elongation cycle repeats itself: a charged tRNA and its amino acid occupy the A site, a peptide bond is formed between the amino acids in the A and P sites, and the ribosome translocates to the next codon. Throughout the cycle, the polypeptide chain remains attached to the tRNA in the P site. The ribosome moves down the mRNA in the 5 _:3_ direction, adding amino acids one at a time according to the order specified by the mRNA's codon sequence. Elongation in eukaryotic cells takes place in a similar manner.

Concept:

Elongation consists of three steps: (1) a charged tRNA enters the A site, (2) a peptide bond is created between amino acids in the A and P sites, and (3) the ribosome translocates to the next codon. Elongation requires several elongation factors (EF-Tu, EF-Ts, and EF-G) and GTP.



15.22 **The elongation of translation comprises three steps.**



Termination

Protein synthesis terminates when the ribosome translocates to a termination codon. Because there are no tRNAs with anticodons complementary to the termination codons, no tRNA enters the A site of the ribosome when a termination codon is encountered .

Instead, proteins called **release factors** bind to the ribosome . *E. coli* has three release factors—RF1, RF2, and RF3. Release factor 1 recognizes the termination codons UAA and UAG, and RF2 recognizes UGA and UAA.

Release factor 3 forms a complex with GTP and binds to the ribosome. The release factors then promote the cleavage of the tRNA in the P site from the polypeptide chain; in the process, the GTP that is complexed to RF3 is hydrolyzed to GDP. Additional factors help bring about the release of the tRNA from the P site, the release of the mRNA from the ribosome, and the dissociation of the ribosome . Translation in eukaryotic cells terminates in a similar way, except that there are two release factors:

eRF1, which recognizes all three termination codons, and eRF2, which binds GTP and stimulates the release of the polypeptide from the ribosome. Findings from recent studies suggest that the release factors bring about the termination of translation by completing a final elongation cycle of protein synthesis. In this model, RF1 and RF2 are similar in size and shape to tRNAs and occupy the A site of the ribosome, just as the amino acid–tRNA–EF–Tu–GTP

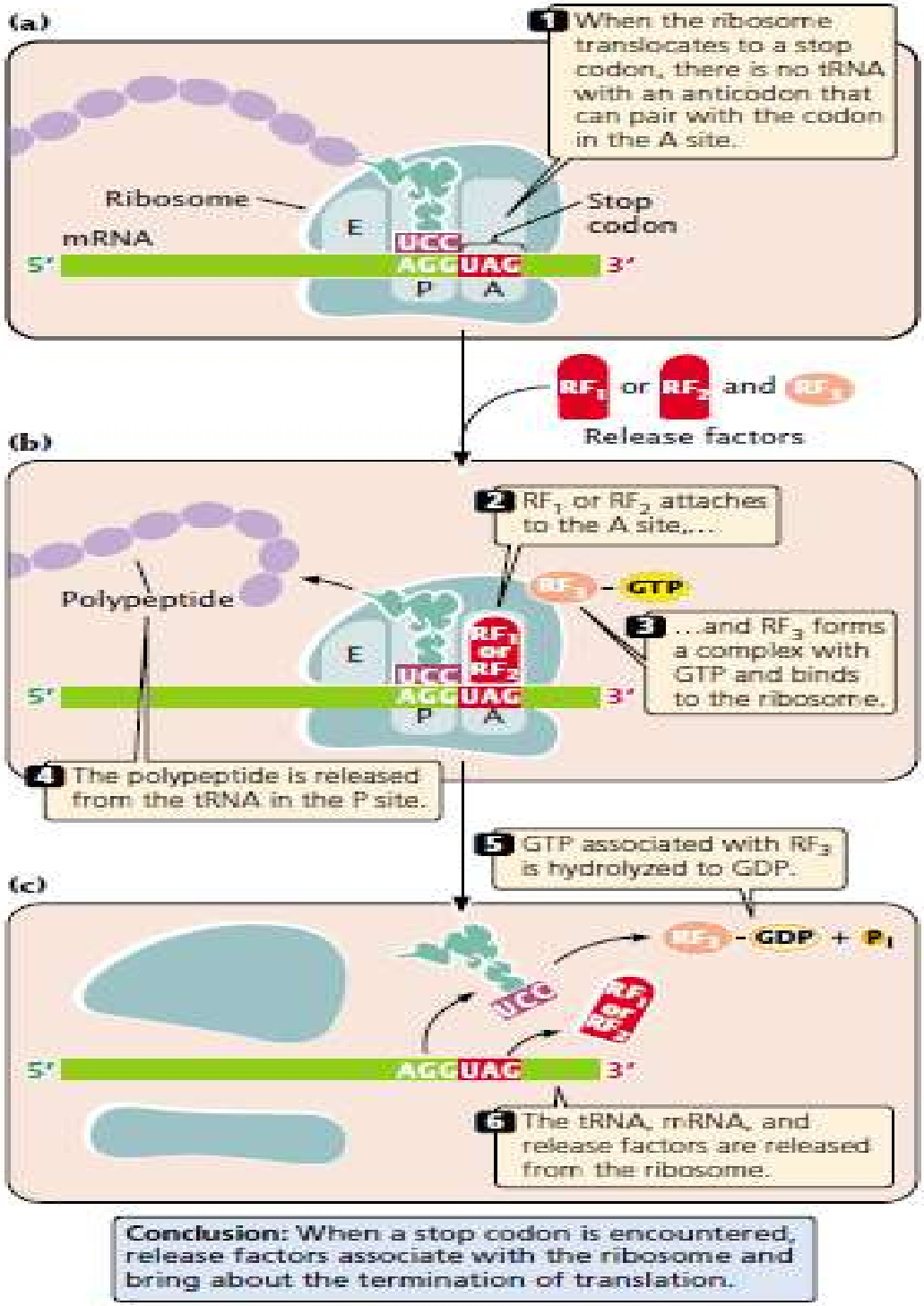
complex does during an elongation cycle. Release factor 3 is structurally similar to EF-G; it then translocates RF1 and RF2 to the P site, as well as the last tRNA to the E site, in a way similar to that in which EF-G brings about translocation. When both the A site and the P site of the ribosome ar cleared of tRNAs, the ribosome can dissociate. Research finding also indicate that some of the sequences in the rRN play a role in the recognition of termination codons.

Concept:

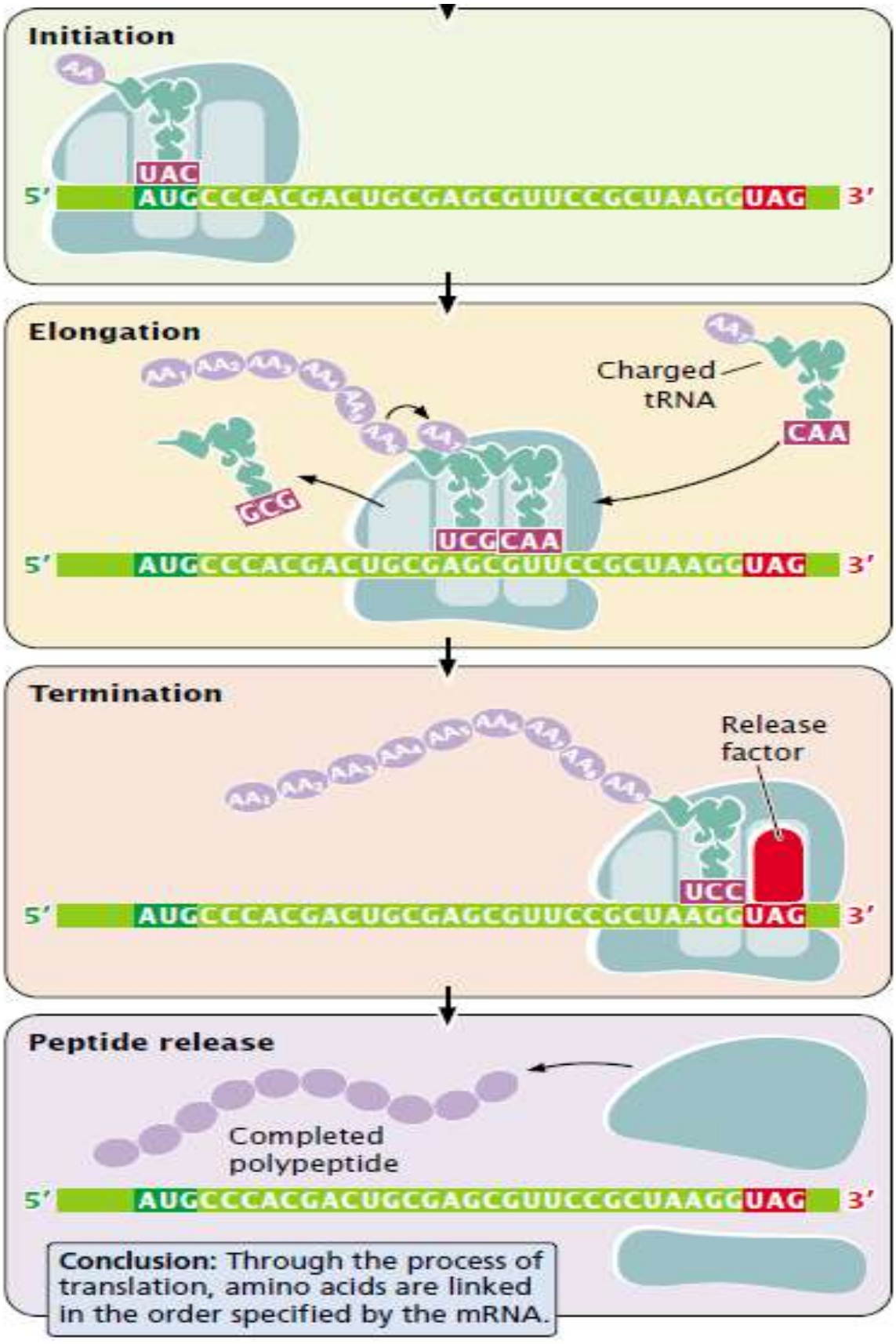
Termination takes place when the ribosome reaches a termination codon. Release factors bind to the termination codon, causing the release of the polypeptide from the last tRNA, the tRNA from the ribosome, and the mRNA from the ribosome.

Concepts

In both prokaryotic and eukaryotic cells, multiple ribosomes may be attached to a single mRNA, generating a structure called a polyribosom

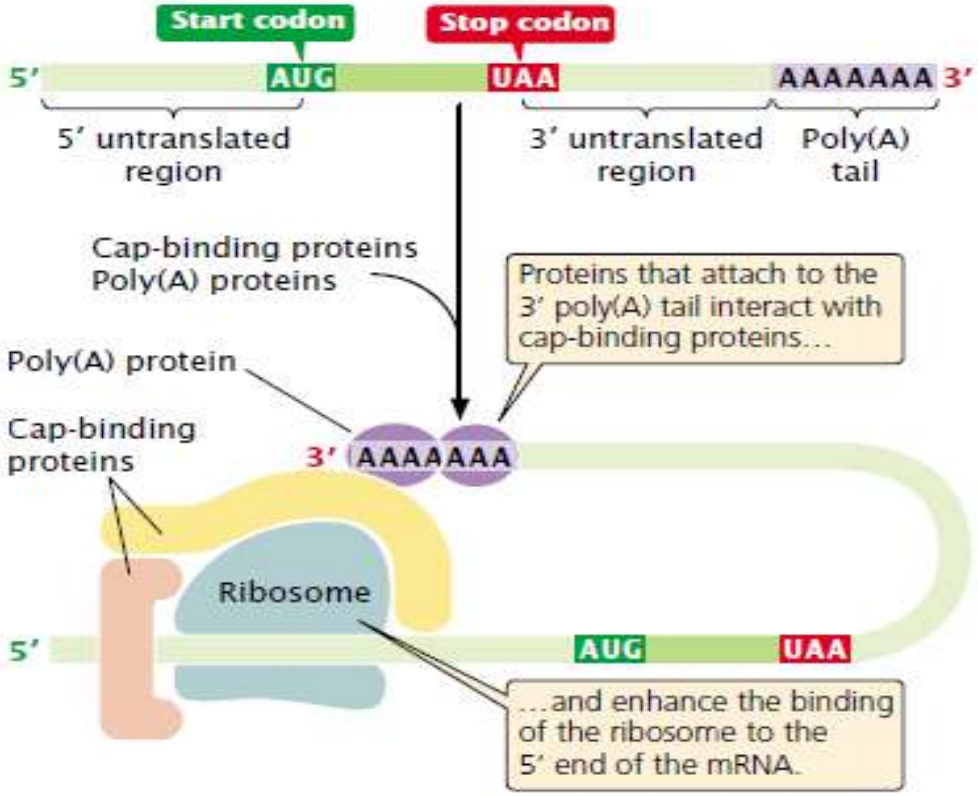


15.23 Translation ends when a stop codon is encountered.


Initiation
UAC
5′ AUGCCCACGACUGCGAGCGUUCCGCUAAGGUAG 3′
Elongation
Charged tRNA
CAA
GCG
UCGCAA
5′ AUGCCCACGACUGCGAGCGUUCCGCUAAGGUAG 3′
Termination
Release factor
UCC
5′ AUGCCCACGACUGCGAGCGUUCCGCUAAGGUAG 3′
Peptide release
Completed polypeptide
5′ AUGCCCACGACUGCGAGCGUUCCGCUAAGGUAG 3′
Conclusion: Through the process of translation, amino acids are linked in the order specified by the mRNA.

Comparison of Protein Synthesis

Prokaryotes	Eukaryotes (cytoplasm)
Polycistronic mRNA	Monocistronic mRNA
Coupled transcription and translation	No coupled transcription and translation for nuclear genes
Linear polyribosomes	Circular polyribosomes
No cap on mRNA	5′-End of mRNA is recognized by cap
Start codon is next AUG after ribosome binding site	No ribosome binding site so first AUG in mRNA is used
First amino acid is formyl-Met	First Met is unmodified
70S ribosomes made of:	80S ribosomes made of:
30S and 50S subunits	40S and 60S subunits
Small 30S subunit: 16S rRNA 21 proteins	Small 40S subunit: 18S rRNA 33 proteins
Large 50S subunit: 23S and 5S rRNA 31 proteins	Large 60S subunit: 28S, 5.8S and 5S rRNA 49 proteins
Elongation factors EF-T (2 subunits) and EF-G	Elongation factors eEF1 (3 subunits) and eEF2
Three initiation factors IF1, IF2 and IF3	Multiple initiation factors eIF2 (3 subunits), eIF3, eIF4 (4 subunits), eIF5
Shut-off by dimerization of ribosomes in non-growing cells	Control via IF sequestration



15.21 The poly(A) tail at the 3′ end of eukaryotic mRNA plays a role in the initiation of translation.