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

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**

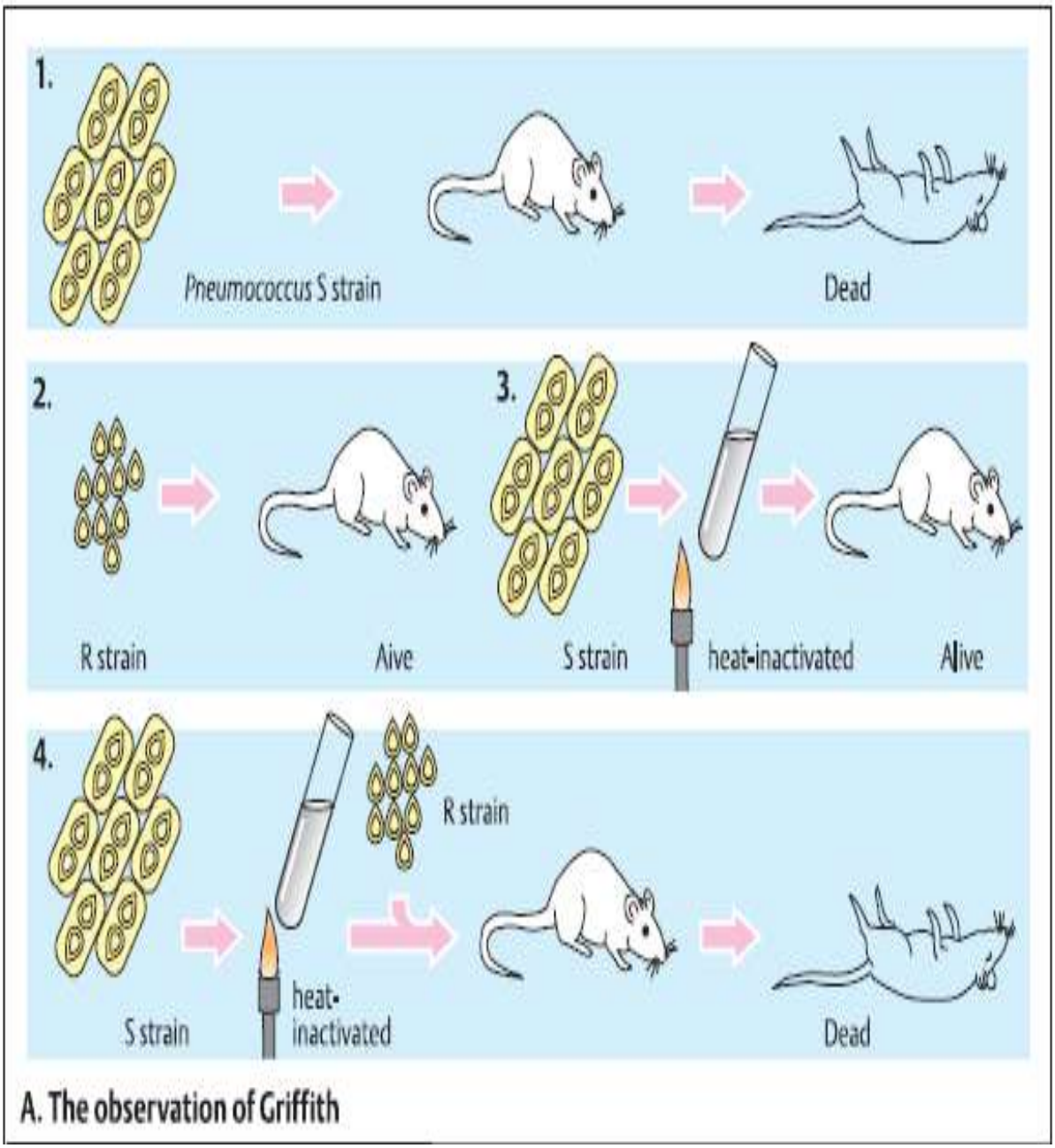
In this case when the bacteria were recovered from the cold lifeless mouse they were smooth virulent pneumococcus (i.e. indistinguishable from wild type).

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

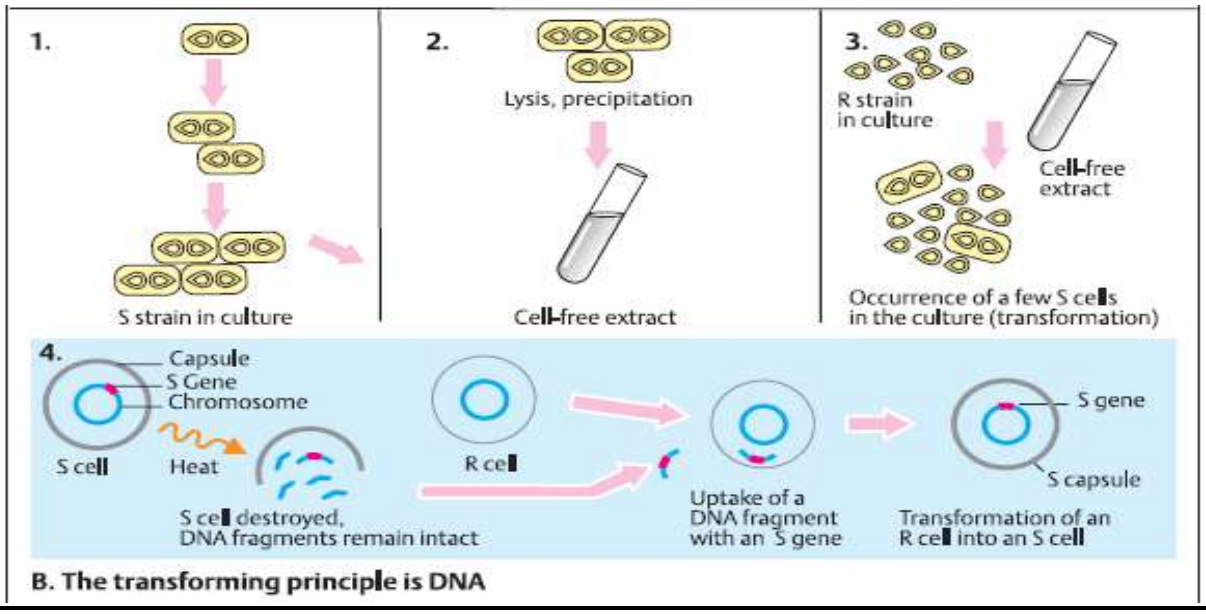
Isolation of bacteria from the dead mouse showed that they were type I . Wild Type *(smooth) bacteria*

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:

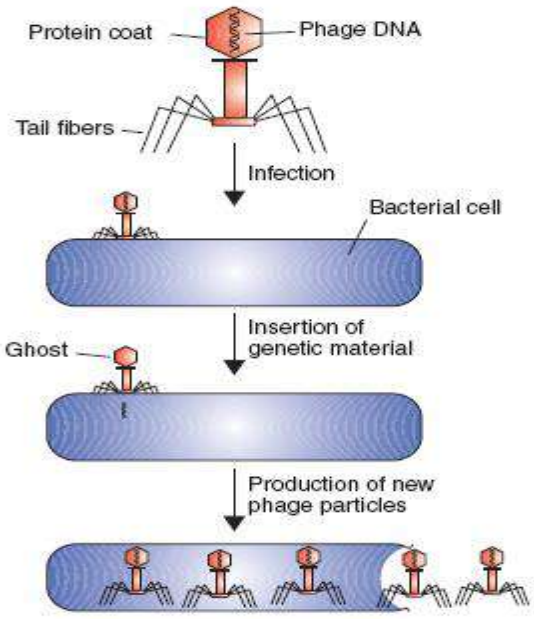
The work of Avery provided strong evidence that the "transforming agent" was in fact DNA (and not protein). However, not everyone was convinced. Some people felt that a residual amount of protein might remain in the purified DNA, even after Trypsin treatment, and could be the "transforming agent".



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 .

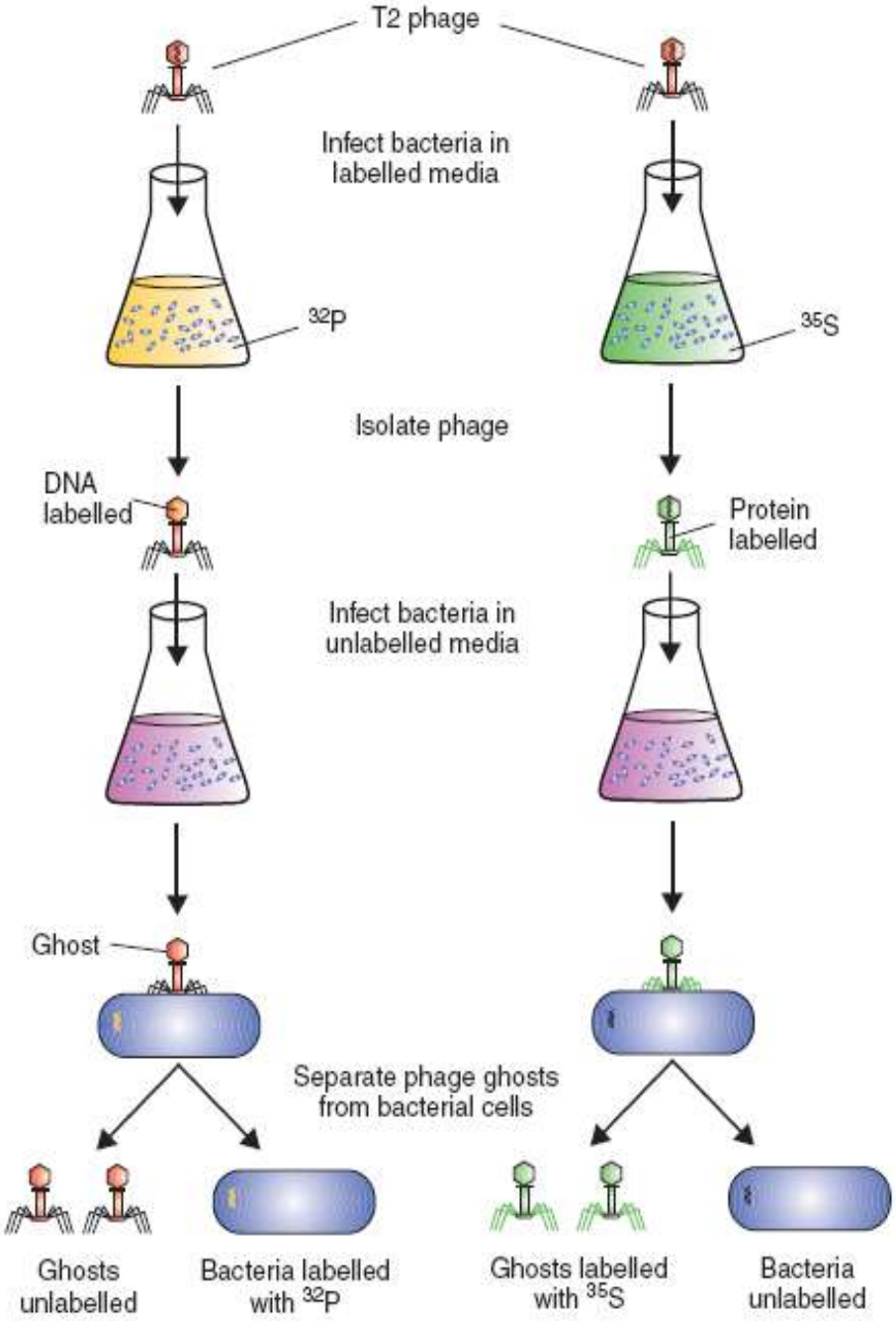
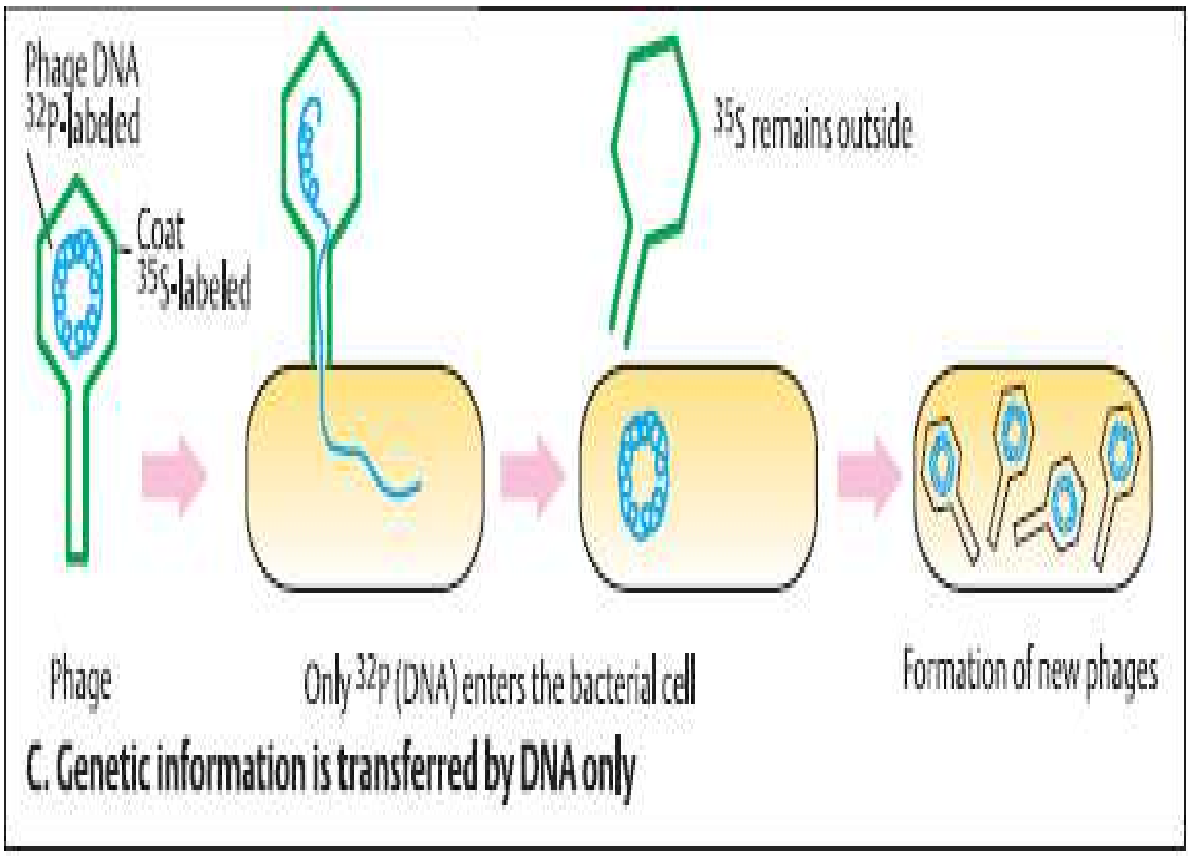


Figure 1.4. The Hershey–Chase blender experiment to show that nucleic acid was the genetic material. Hershey and Chase grew T2 bacteriophages on bacteria whose media contained either ^{32}P (to label the phosphorus of nucleic acid) or ^{35}S (to label the sulphur of proteins – the side chains of the 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



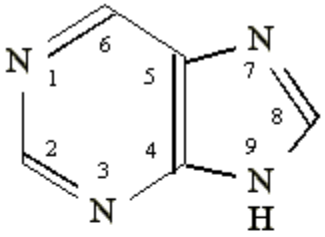
C. Genetic information is transferred by DNA only

Lecture 2:

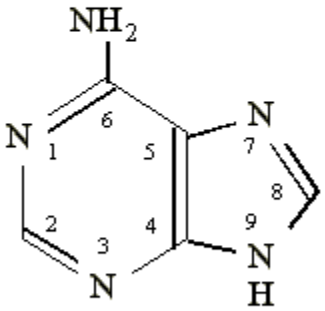
The chemical composition of nucleic acid

The structure of DNA and RNA.

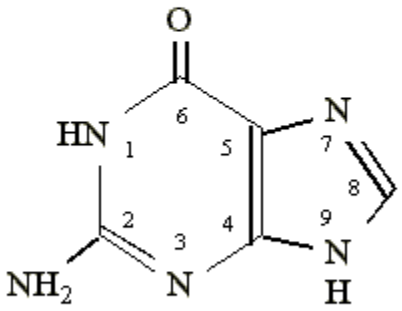
DNA (deoxyribonucleic acid) and RNA (ribonucleic acid) are composed of two different classes of nitrogen containing bases: the **purines** and **pyrimidines**. The most commonly occurring purines in DNA are **adenine** and **guanine**:



Purine

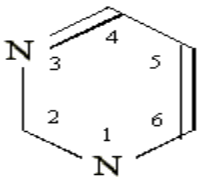


Adenine
135.1 gm/mol

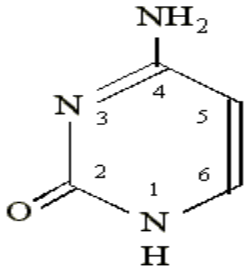


Guanine
151.1 gm/mol

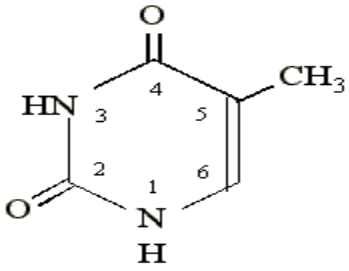
The most commonly occurring pyrimidines in DNA are **cytosine** and **thymine**:



Pyrimidine



Cytosine
111.1 gm/mol

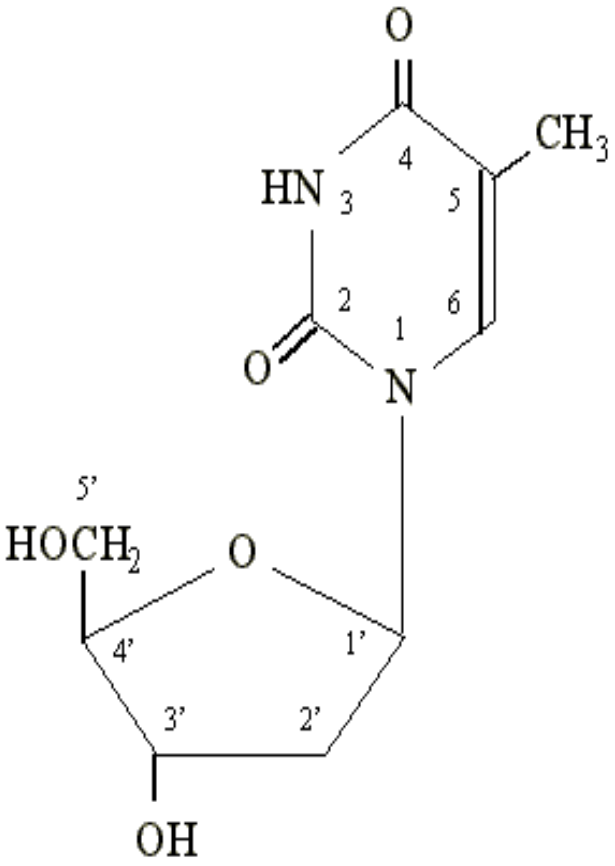


Thymine
126.1 gm/mol

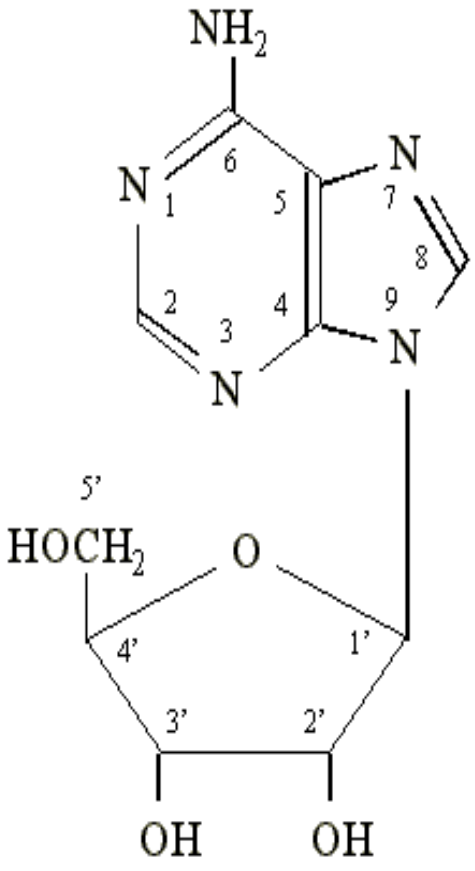
RNA contains the same bases as DNA with the exception of thymine. Instead, RNA contains the pyrimidine **uracil**:

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*.

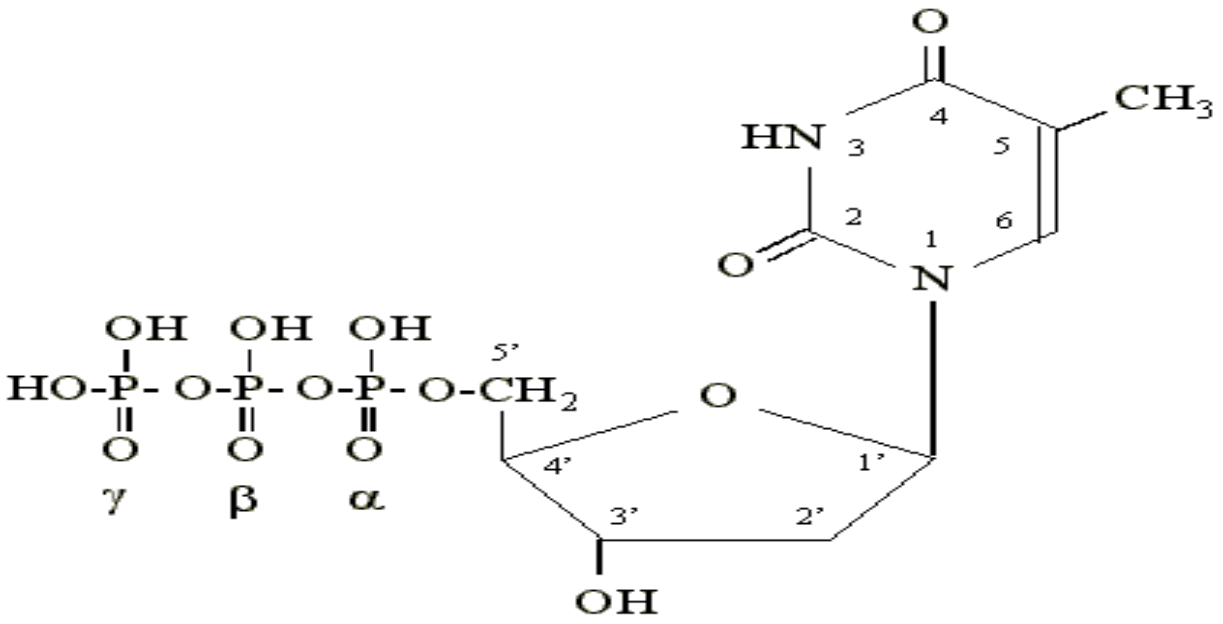


2'-deoxy Thymidine
(2'-deoxy ribose sugar
DNA precursor)



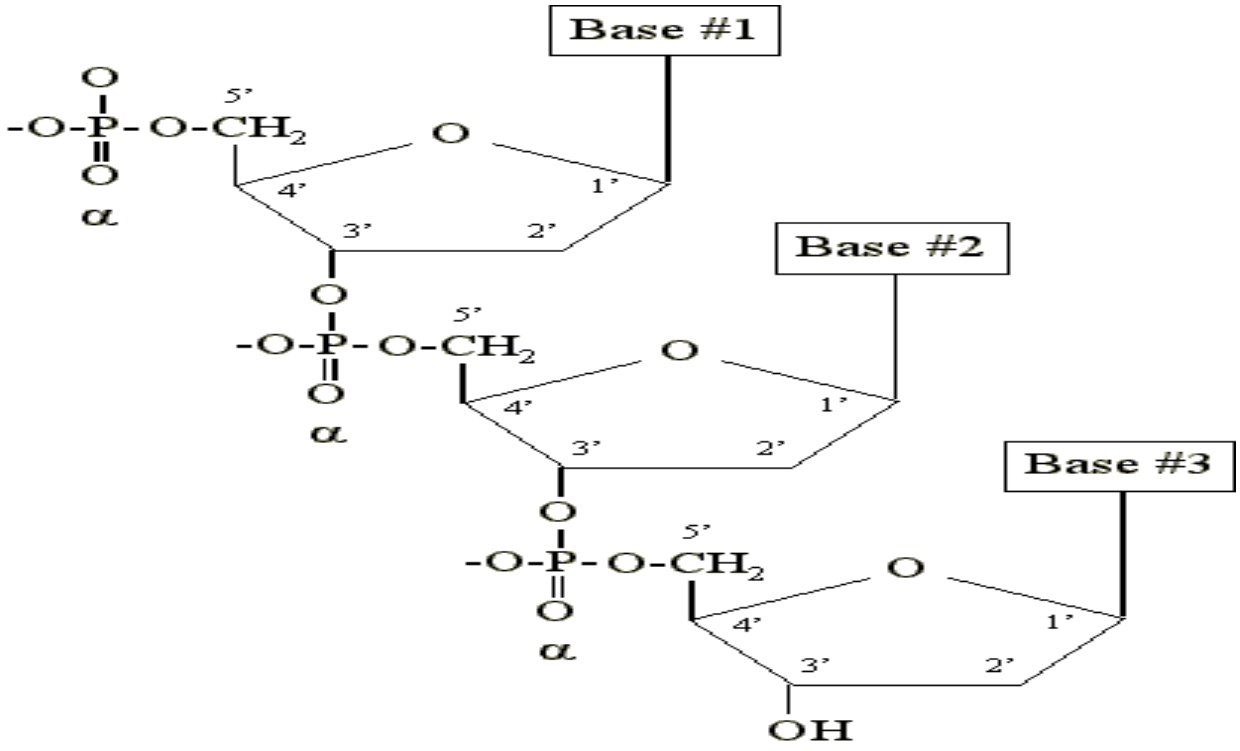
Adenosine
(ribose sugar
RNA precursor)

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.

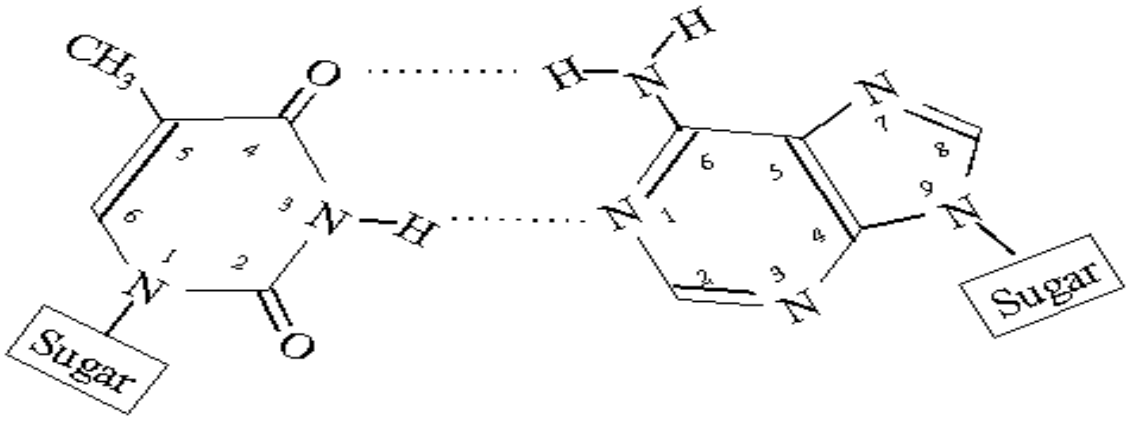
1953 L. Pauling and R.B. Corey

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

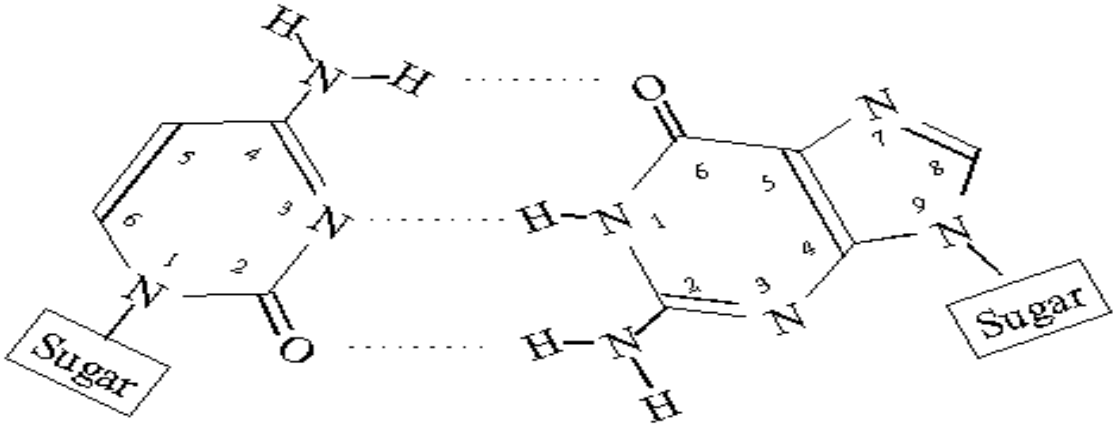
Lecture 3:

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:

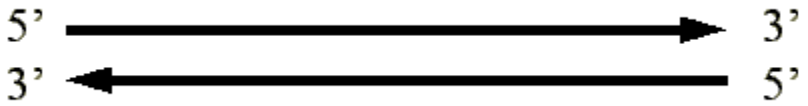
- 34 Å per helical repeat
- 10 base pairs per repeat (i.e. per turn of the helix)
- 3.4 Å inter-base stacking distance
- 20 Å diameter for the helical width

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.

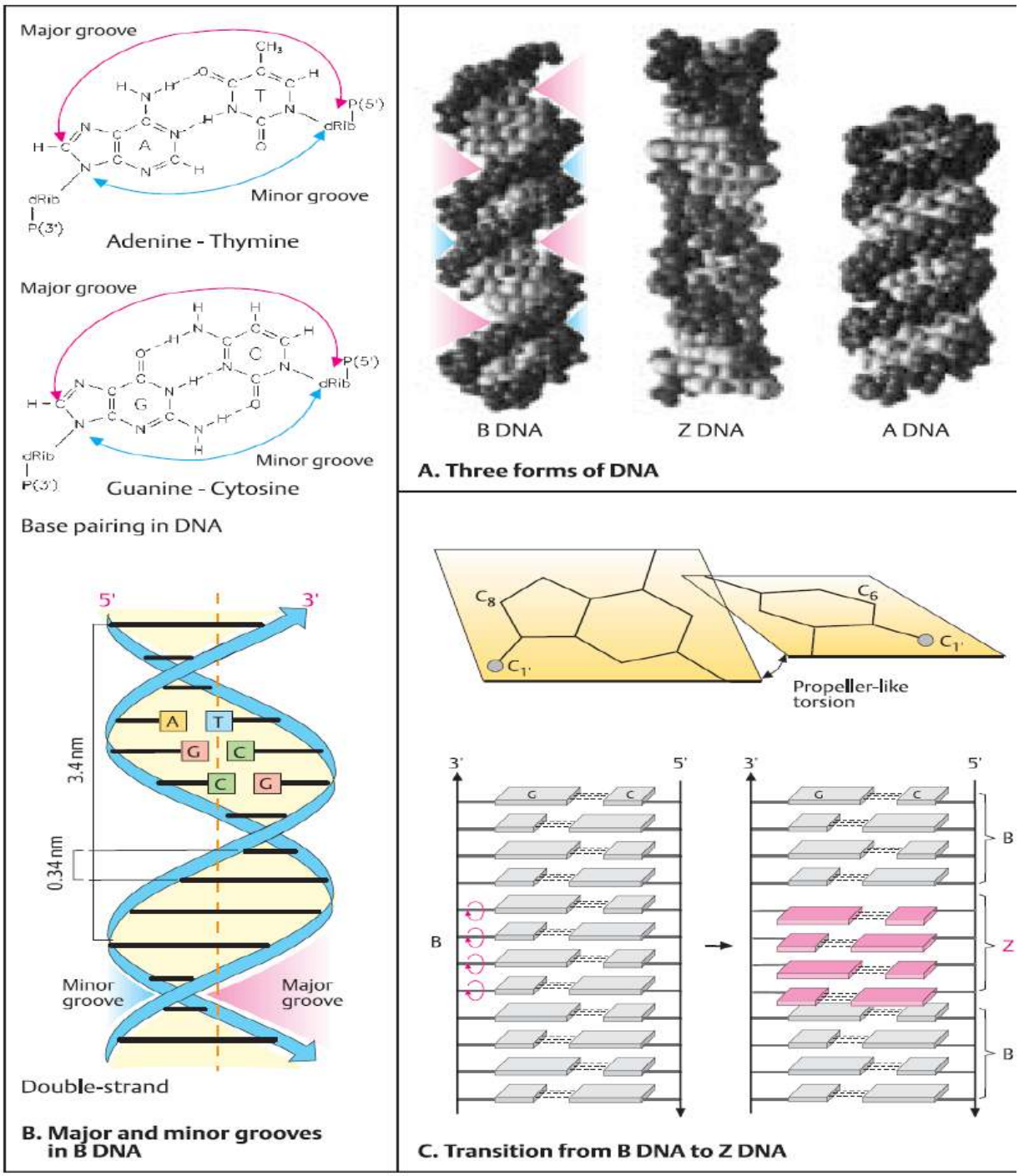
1-The original classic form, determined by Watson and Crick in 1953, is B-DNA. The essential structural characteristic of B-DNA is the formation of two grooves, one large (major groove) and one small (minor groove). the B-DNA forms a right-handed helix, (table -)

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

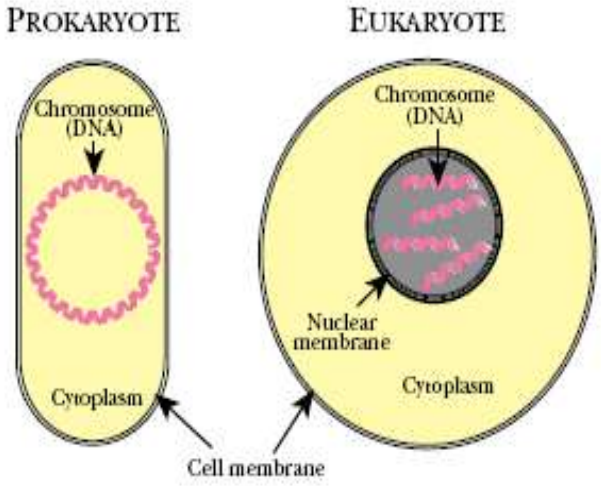
The base pairing in DNA (adenine–thymine and guanine–cytosine) leads to the formation of a large and a small groove because the glycosidic bonds to deoxyribose (dRib) are not **diametrically opposed**. In B-DNA, the purine and pyrimidine rings lie 0.34 nm apart. DNA has ten base pairs per turn of the 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



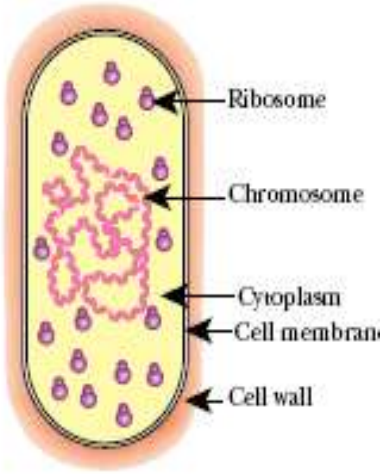
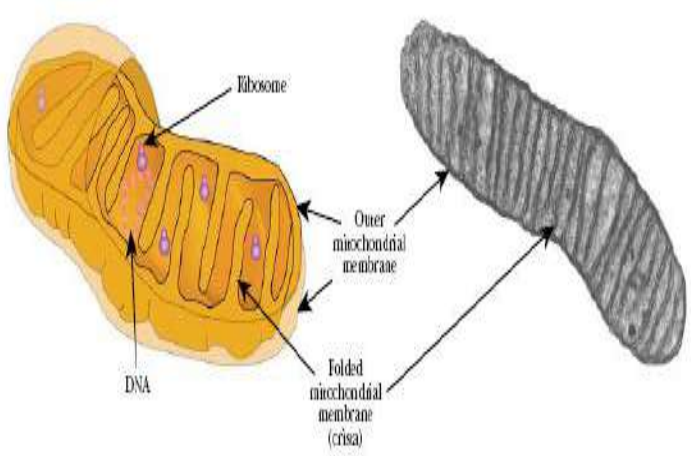
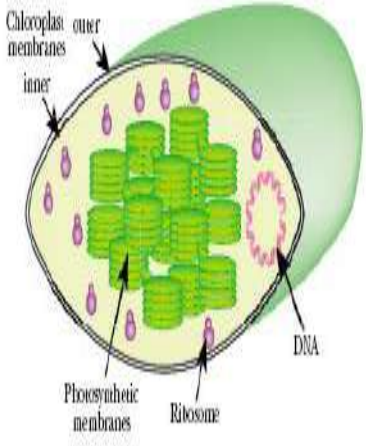


FIGURE 2.07 ***Typical Bacterium***

.





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**.

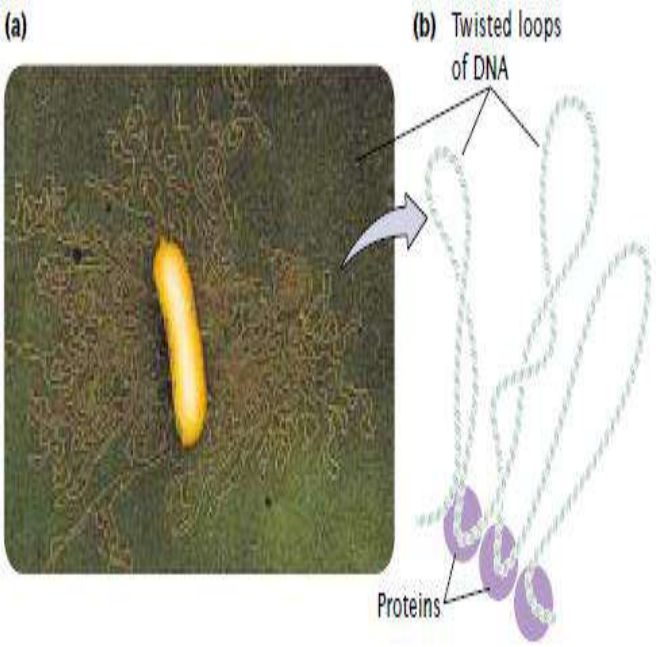
Supercoiling can only be introduced into or released from a closed-circular DNA molecule by breaking at least one of the phosphor diester backbones.

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.)

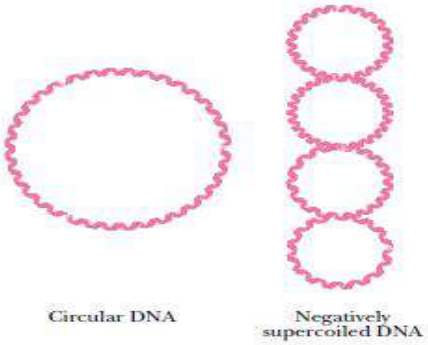


FIGURE 4.15 ***Supercoiling of DNA***

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 achieved.

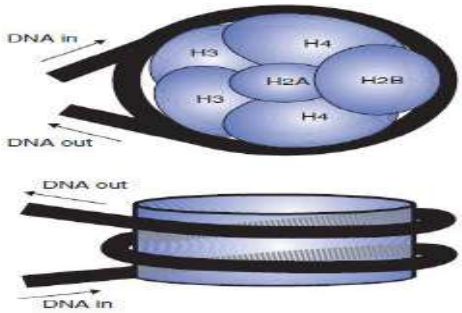
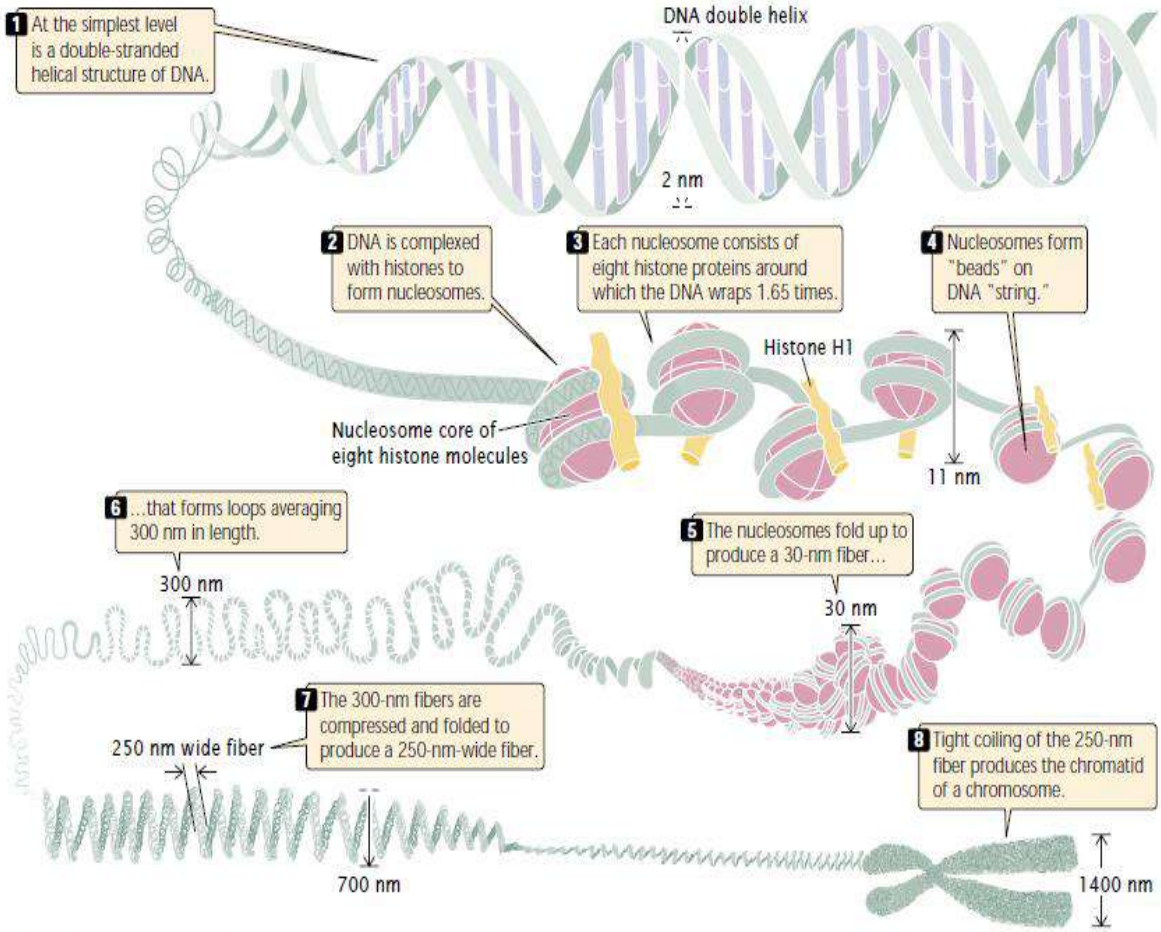


Figure 1.14. Wrapping of DNA around the nucleosome core. The nucleosome is composed of two molecules each of histones H4, H3, H2A and H2B. In the representation shown here only a monomer of H2A and H2B can be observed, the other monomers being located at the back of the octamer. Almost 150 bp of DNA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

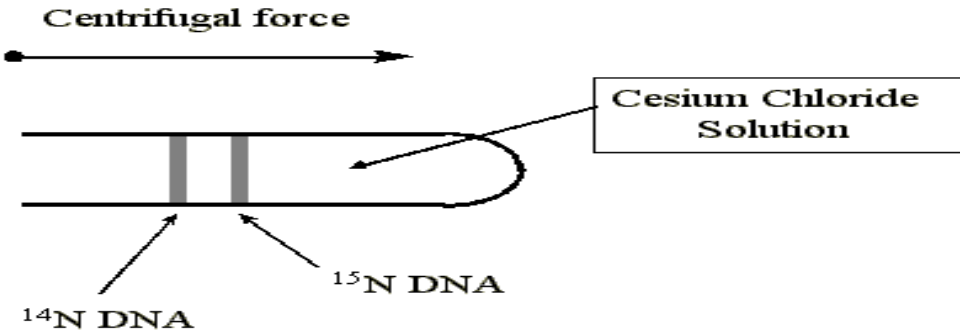
- The DNA model of Watson and Crick suggested how genetic information might be replicated: either strand of the duplex can be used as a template to replicate the sequence information.
- But, was the replication conservative (i.e. the original parental strands remain together after replication) or semi-conservative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

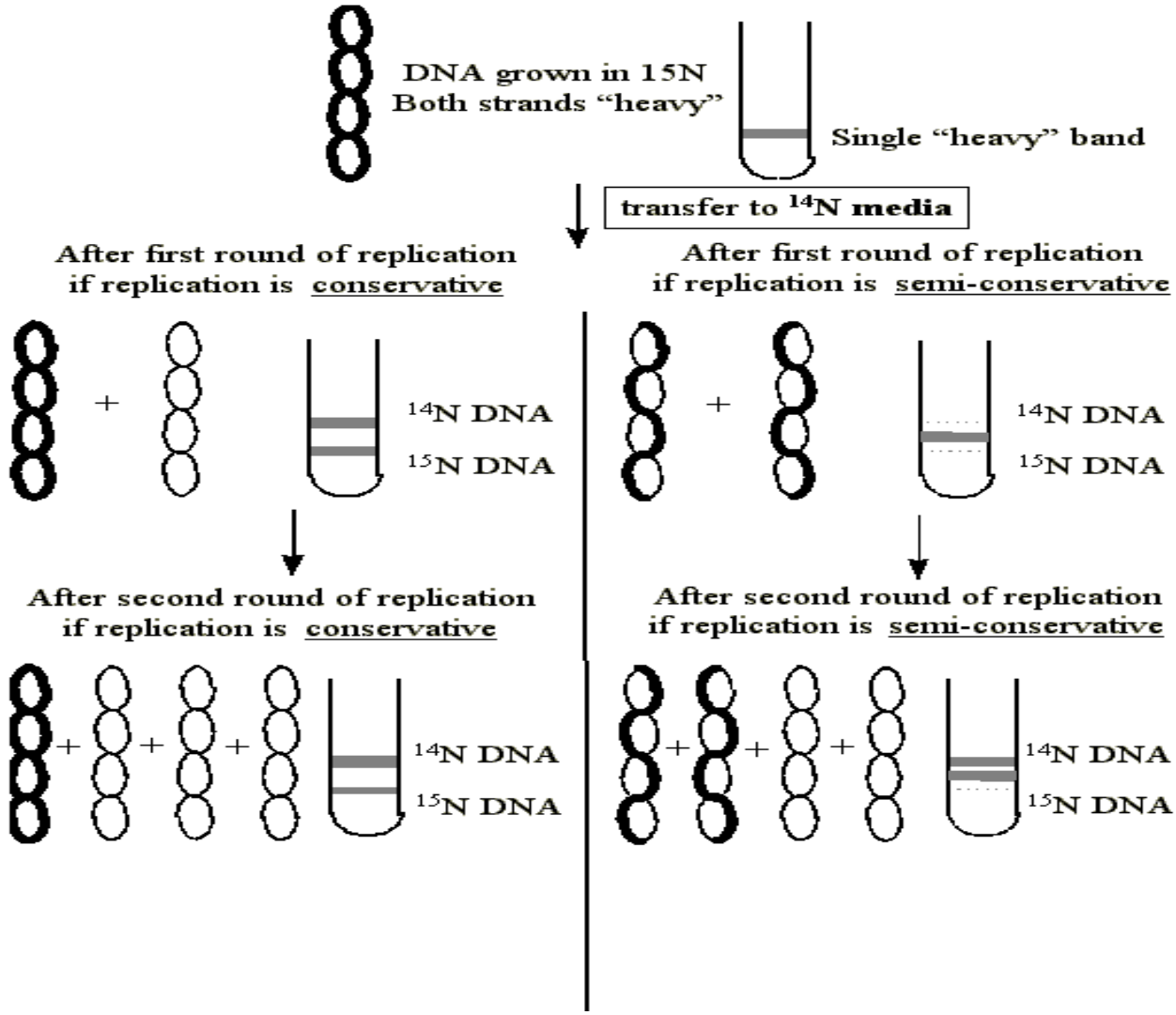
1958

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*).

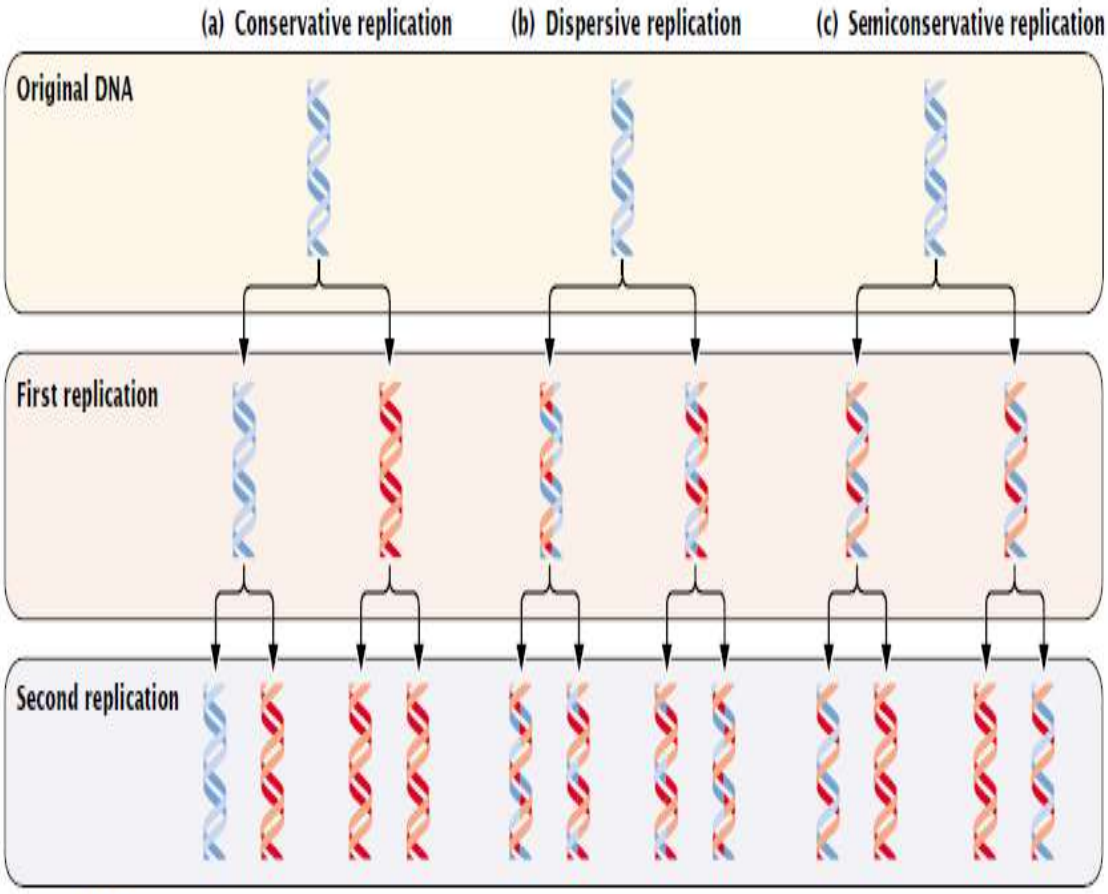


Meselson and Stahl reasoned that if they grew *E. coli* in ^{15}N salts then ***switched media to ^{14}N salts*** for additional rounds of replication, the <u>mode of replication</u> could be deduced from the <u>density</u> of the DN

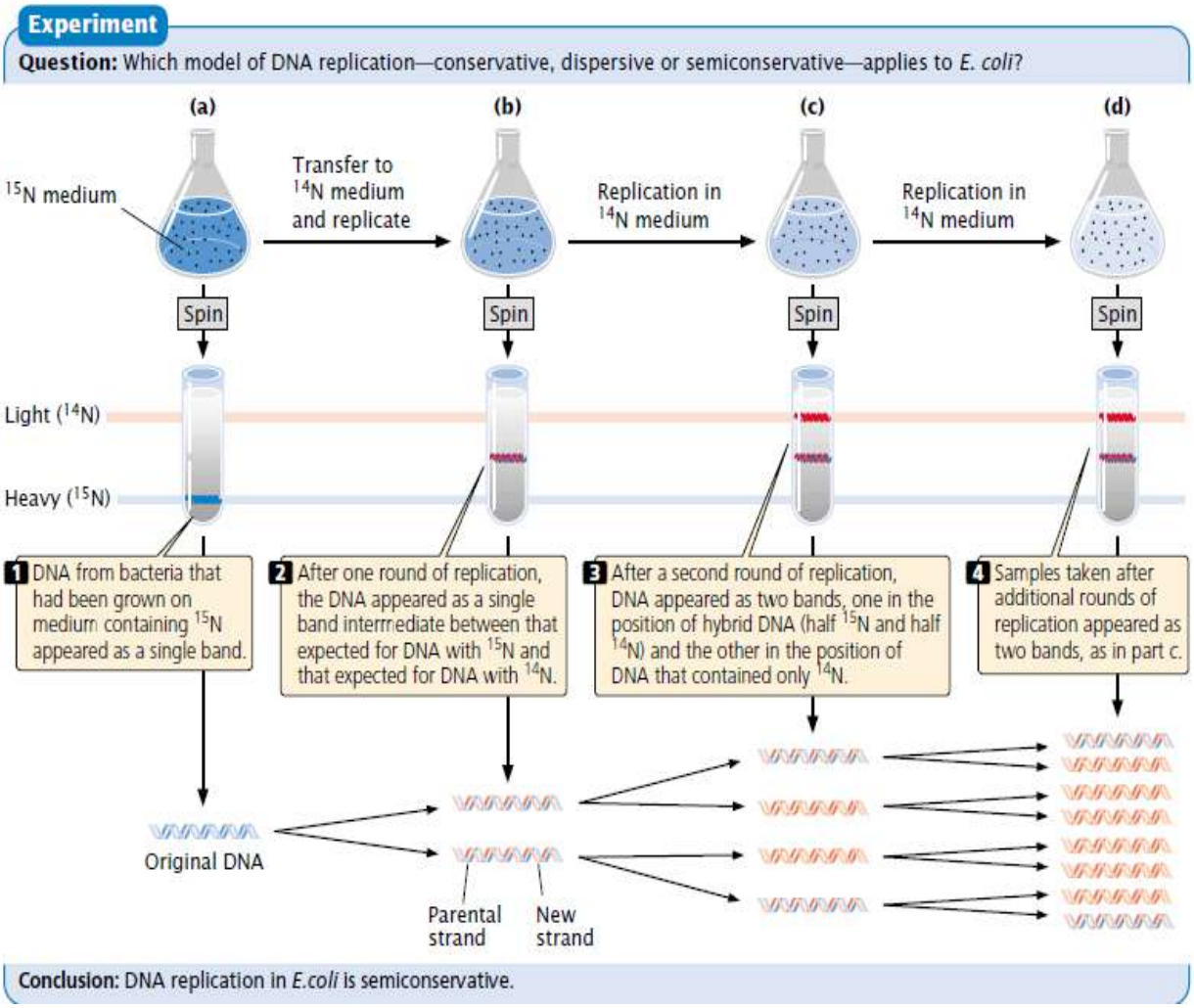


-
- After switching to the ^{14}N media and allowing the cells to go through a round of replication a single band of intermediate density was observed (i.e. between ^{14}N and ^{15}N control DNA samples).
- After a second round of replication in ^{14}N media two bands were present in approximately equimolar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

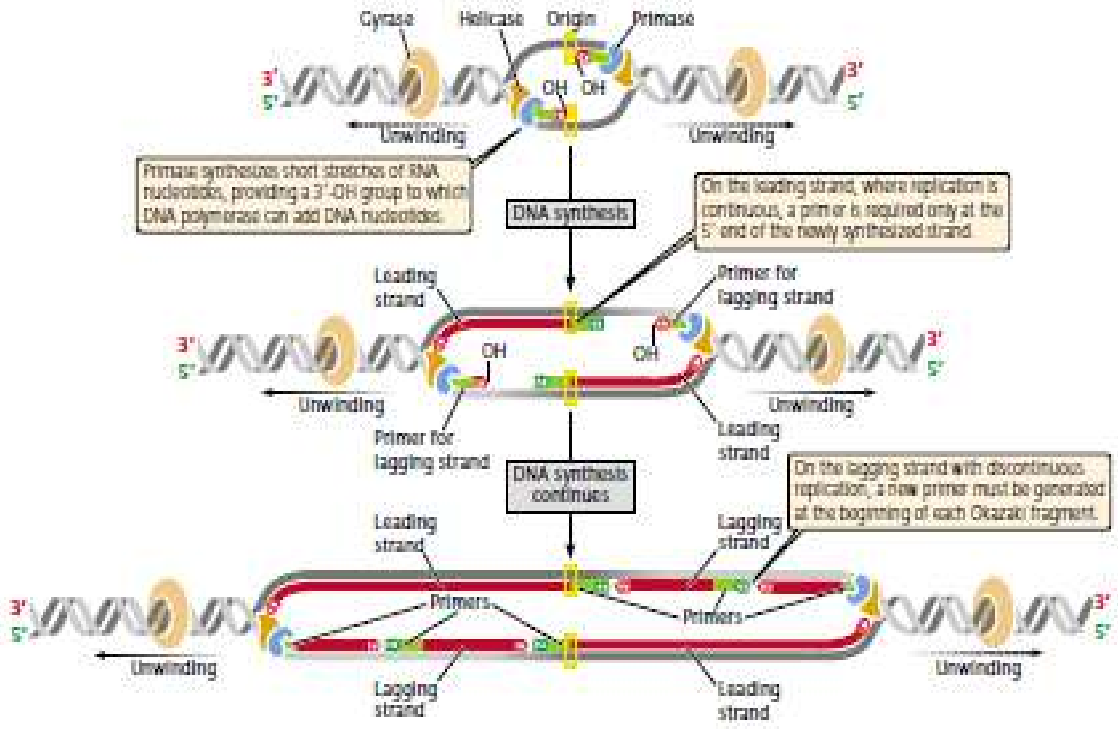
2. Polymerases cannot replicate a DNA template in the absence of a primer (either DNA or RNA).

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 a Enzyme that makes short segment of initiator DNA during replication of animal chromosomes
DNA polymerase d 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

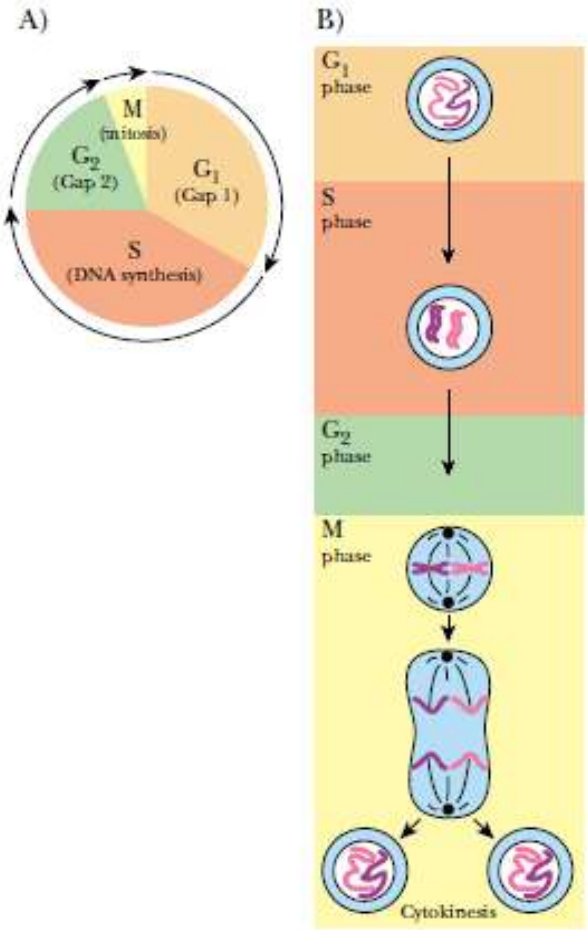
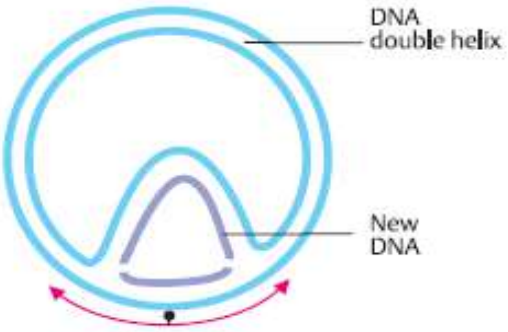
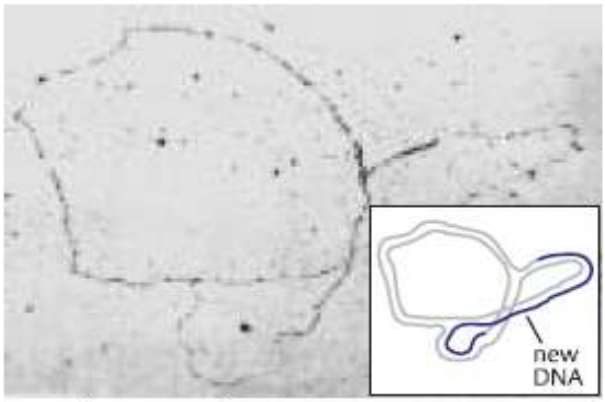


FIGURE 5.32 ***The Eukaryotic Cell Cycle***

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.

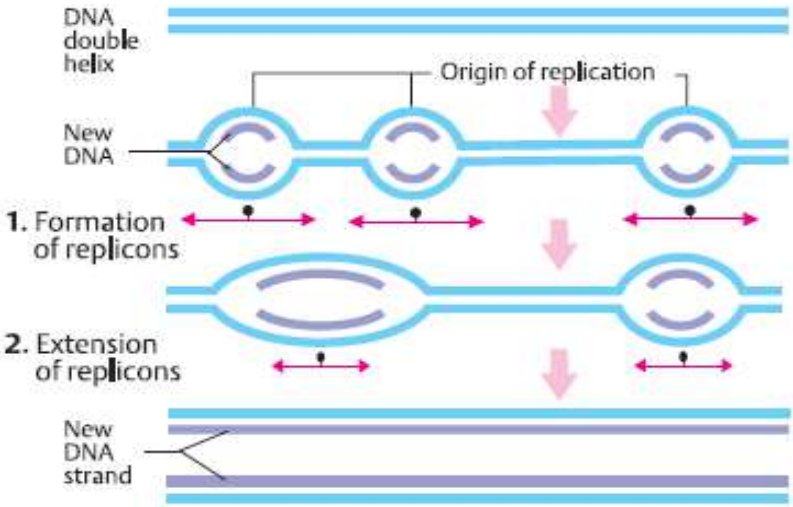


1. DNA replication in the bacterial chromosome

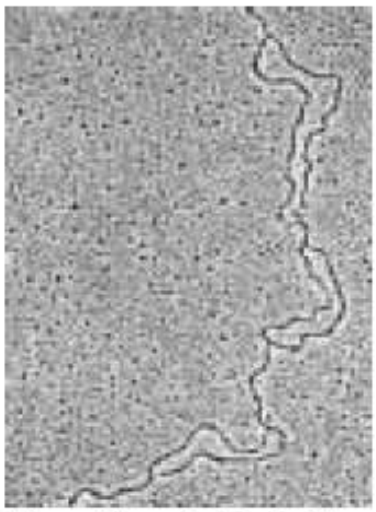


2. Prokaryotic replication in an autoradiogram in *E. coli* (J. Cairns)

A. Prokaryotic replication begins at one site

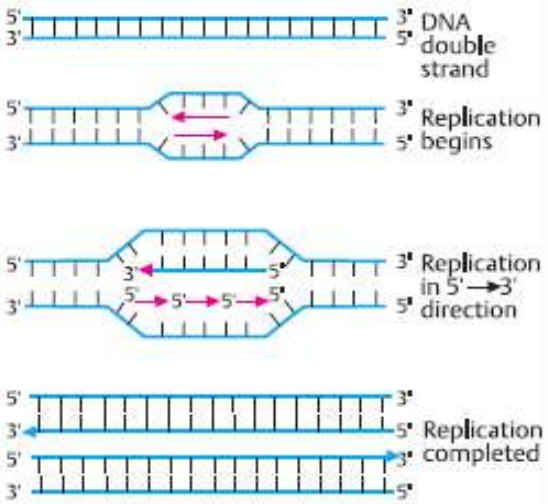


3. Replication completed

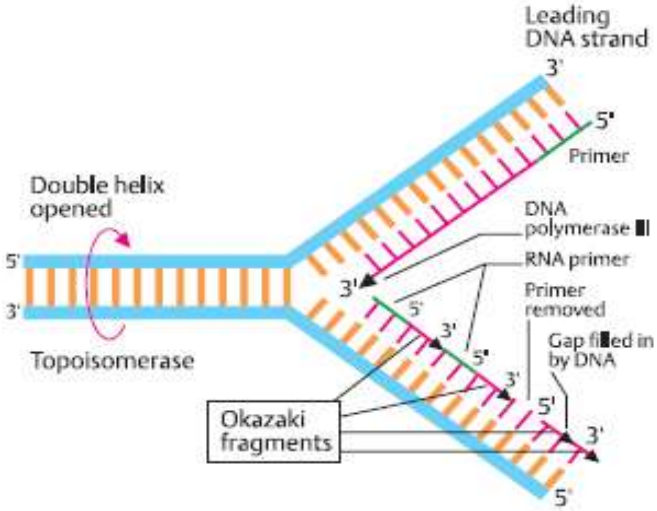


4. Eukaryotic replication in the EM (D. S. Hogness)

B. Eukaryotic replication begins at several sites



C. Scheme of replication



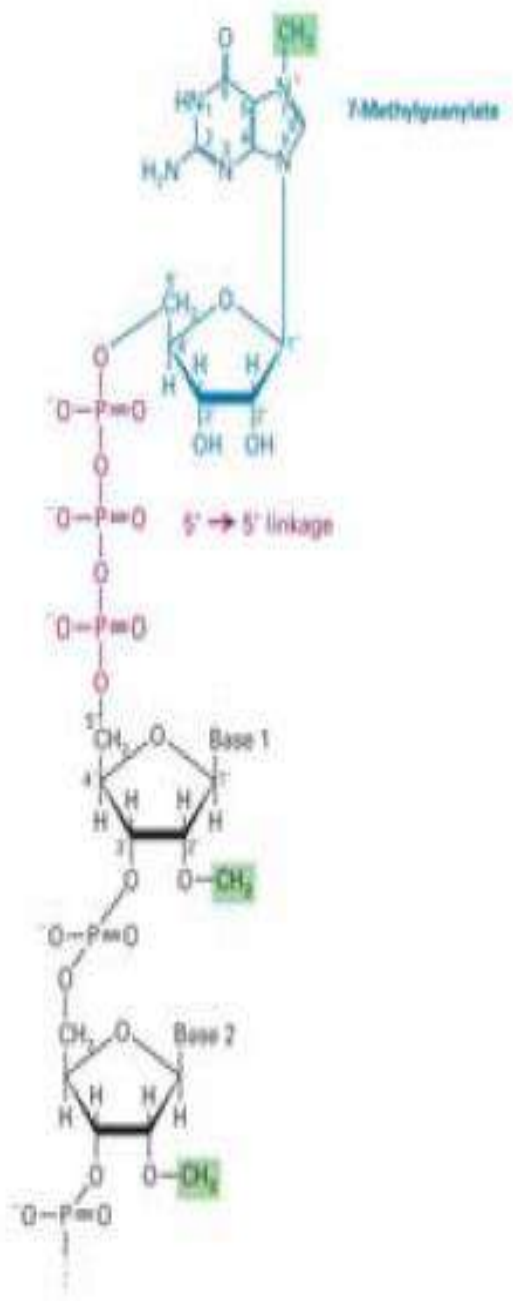
D. Replication fork

Lecture 7:

Splicing and modification of mRNA in eukaryotic cells

5' Cap Structure

- 5' to 5' triphosphate linkage
- 7-methylguanosine at terminus
- Protects mRNA from nucleases
- Binds proteins important for function



Splicing & Poly(A)

- Primary transcript 3' end cleaved and polyadenylated (no template!)
- Intervening sequences cut out of primary transcript
- Removed introns are very unstable (no cap!!)
- Only capped, polyadenylated, spliced RNAs exported from nucleus.

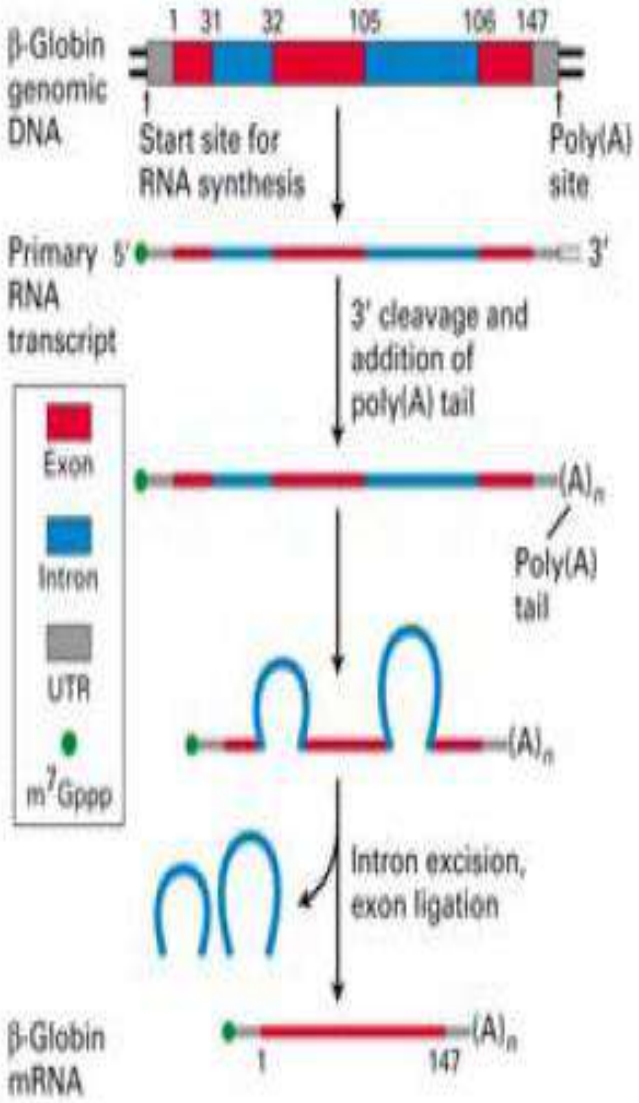
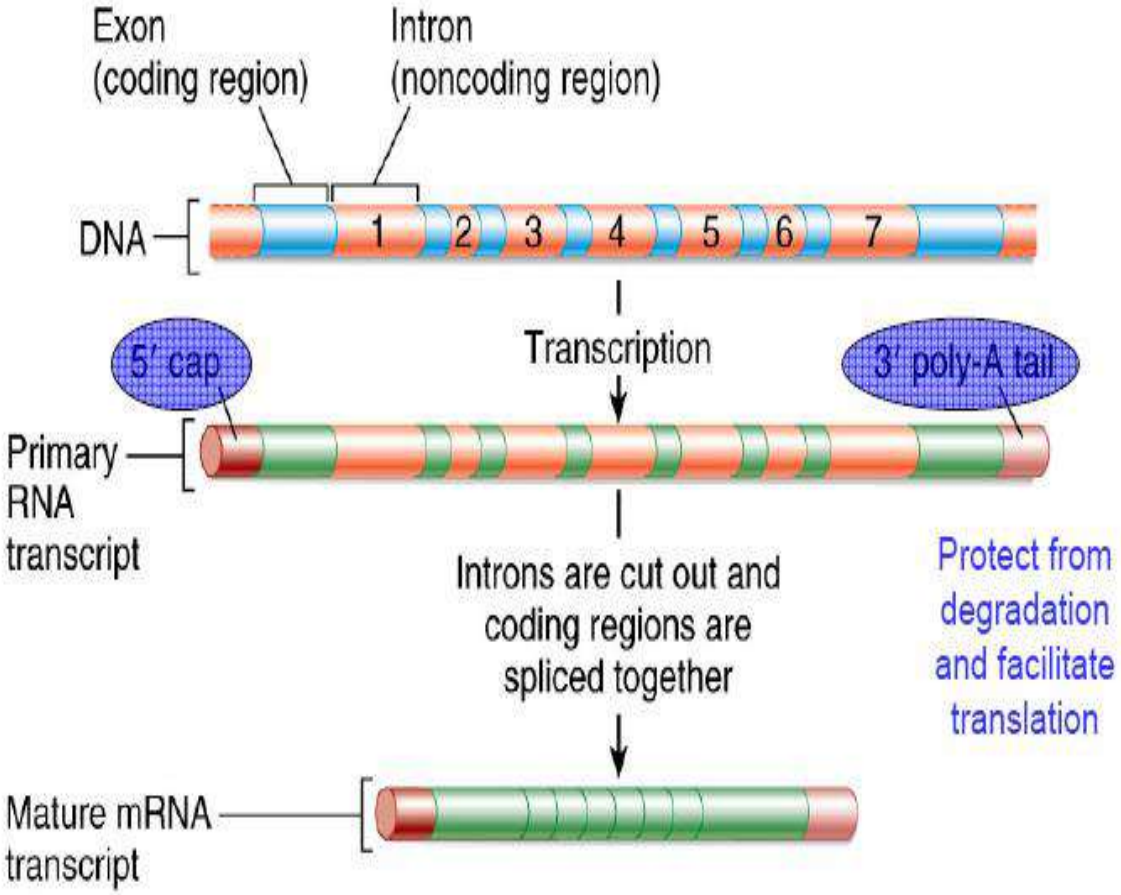


Fig. 9.17 Processing eukaryotic mRNA



Different combinations of exons can generate different polypeptides via alternative splicing

In eukaryotes, genes are fragmented They are composed of

Exons – Sequences that code for amino acids

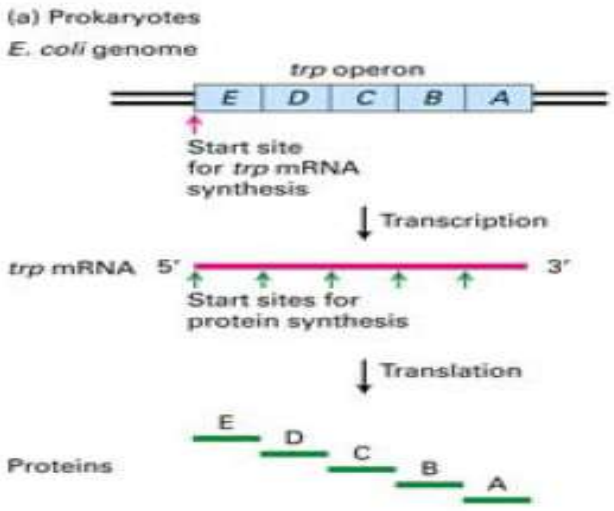
Introns – Sequences that don't code for amino acids

Eukaryotic cells transcribe the entire gene, producing a primary RNA transcript .This transcript is then heavily processed to produce the mature mRNA transcript ,this leaves the nucleus for the cytoplasm

Differiences in transcription between prokaryotes and eukaryotes

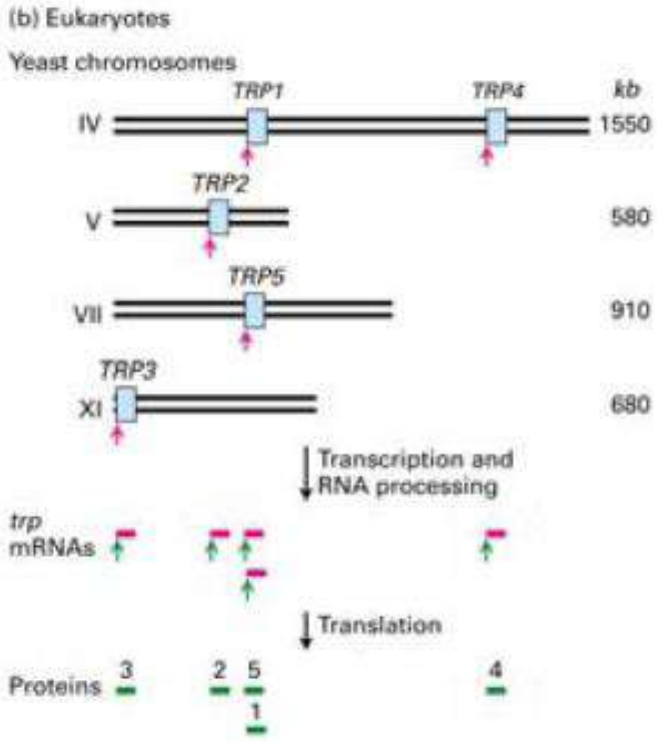
Prokaryotic Transcription

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



Eukaryotic Transcription

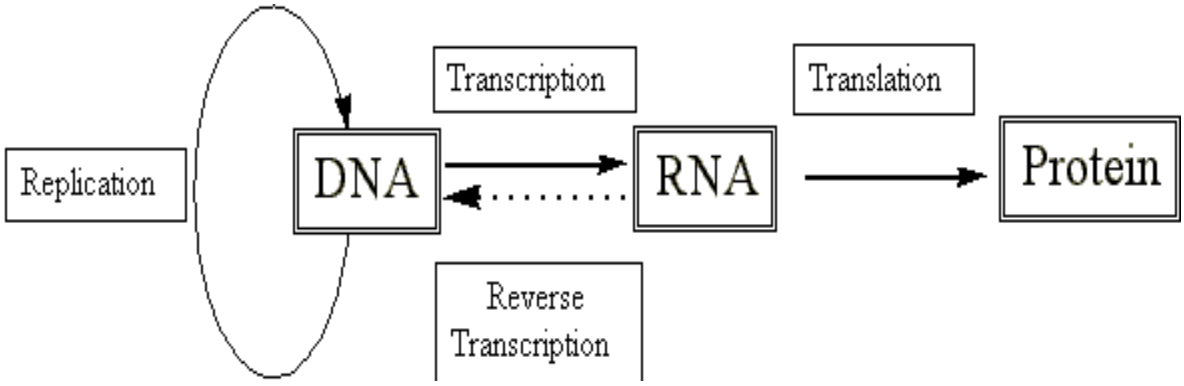
- Monocistronic (*with interesting exceptions!*)
- Capped
- Polyadenylated
- Spliced
- Compartmentalized



Lecture 8:

Introduction to Gene Expression :

The "central dogma"

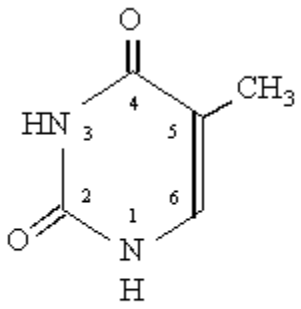


- We have seen how DNA, with the aid of specific polymerases and accessory proteins, is able to replicate.
- We have also seen how we can use this information to create autonomously replicating extra-chromosomal elements (i.e. **plasmids**).

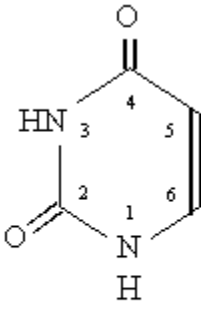
*However, the real utility of such systems arises when we use them to create **proteins** of interest. To get to proteins we have to go through **RNA** first.*

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 complexes with proteins. The most common example is ribosomes (involved in protein synthesis).

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

RNA polymerase:

- *Can initiate* 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. *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.

1. *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.

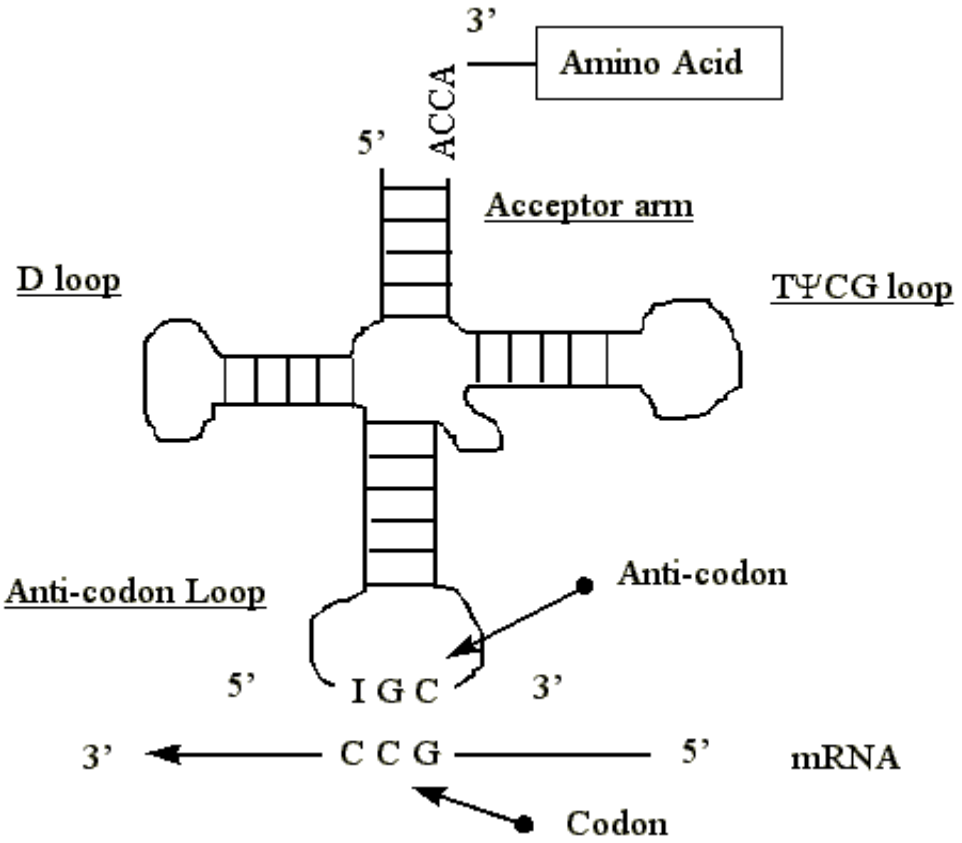
Structure and function of transfer RNA's

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**

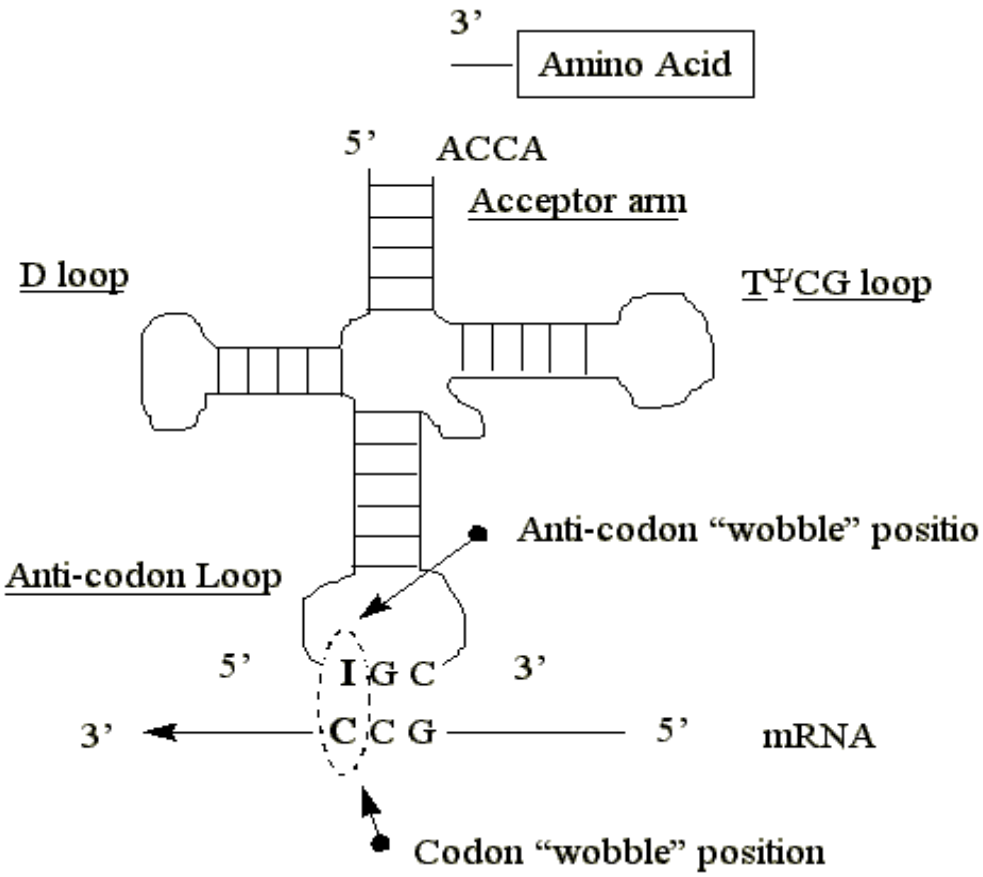
2. Uridylate may be **rearranged** to produce **pseudouridylate** (i.e. ribose attached to Carbon 5 instead of Nitrogen 1).
3. Guanidylate may be **methylated** at different positions.

- 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)



- 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:



-

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).

- **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)

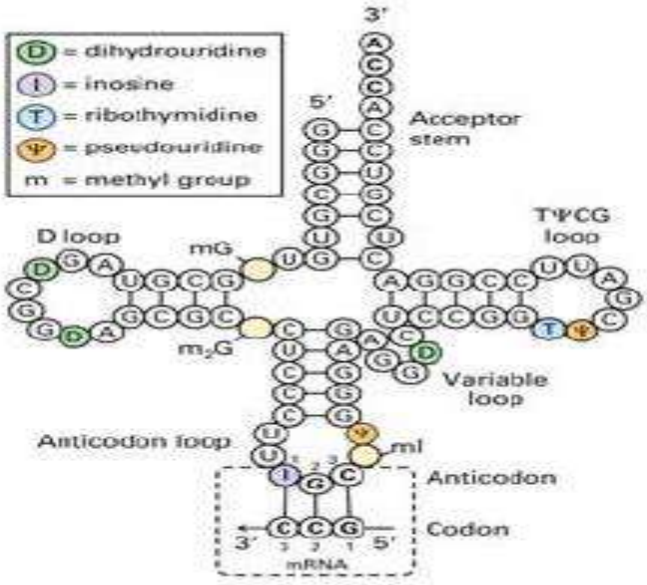
Conclusions:

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



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.

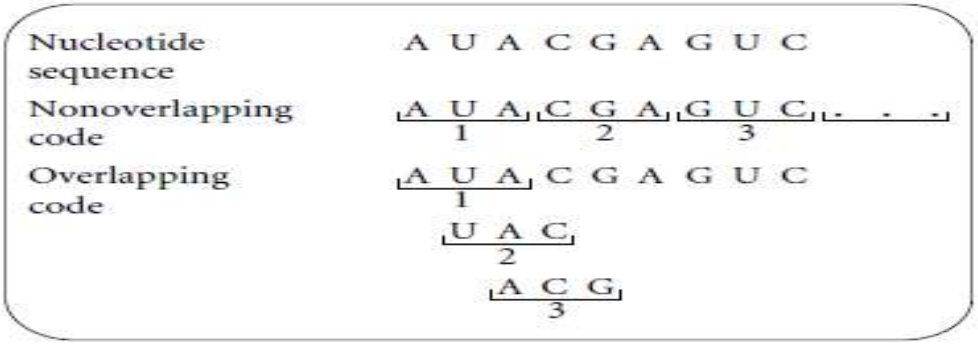
5. The code is generally nonoverlapping; each nucleotide in an mRNA sequence belongs to a single reading frame.

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.

Universal Genetic Code: 20 Amino Acids

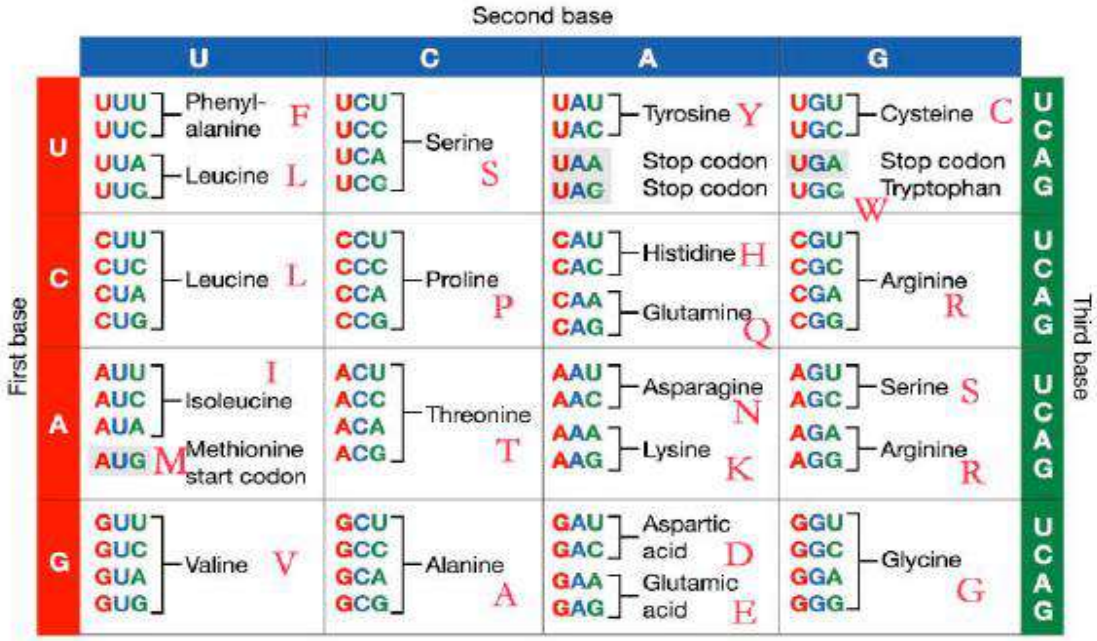


Table 14.3 Composition of ribosomes in bacterial and eukaryotic cells

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

Lecture : 9

Transcription

DNA(replication) transcription RNA translation Protein

Transcribe: to make anRNA that is complementary to a region within a strand of DNA

Types of RNA molecules:

1-rRNA

2-mRNA

3-tRNA

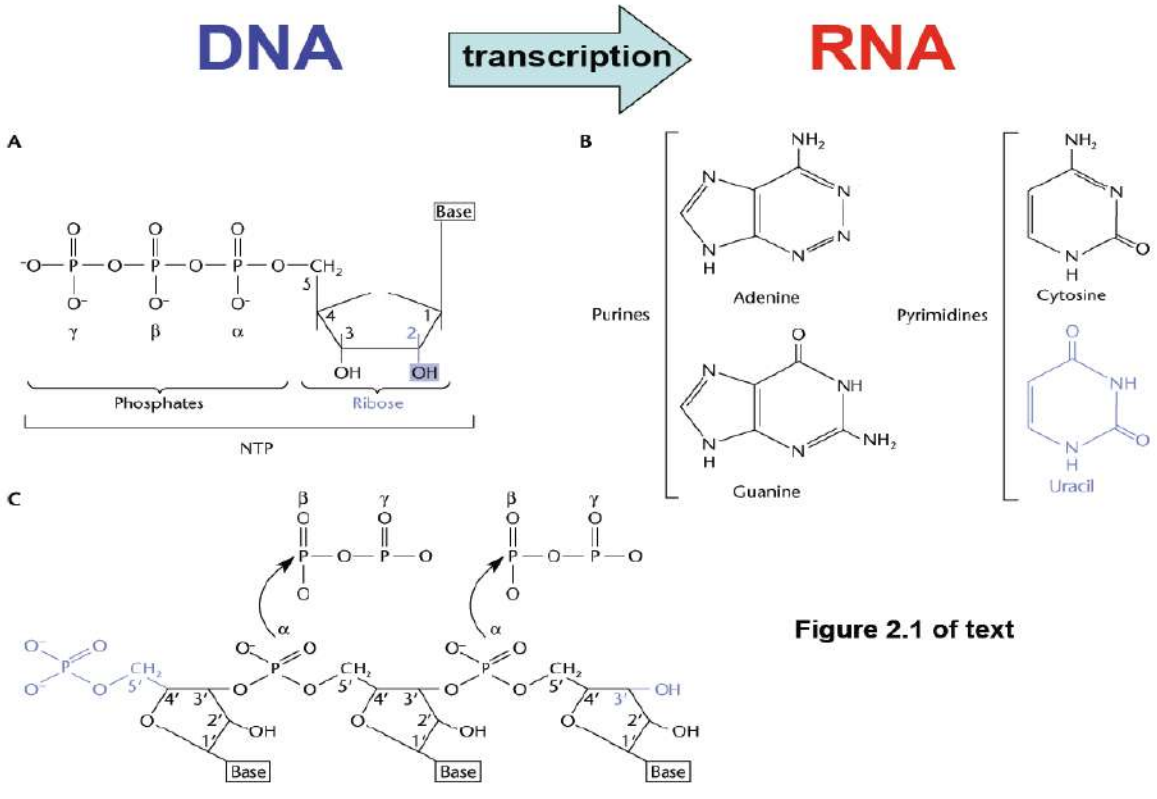


Figure 2.1 of text

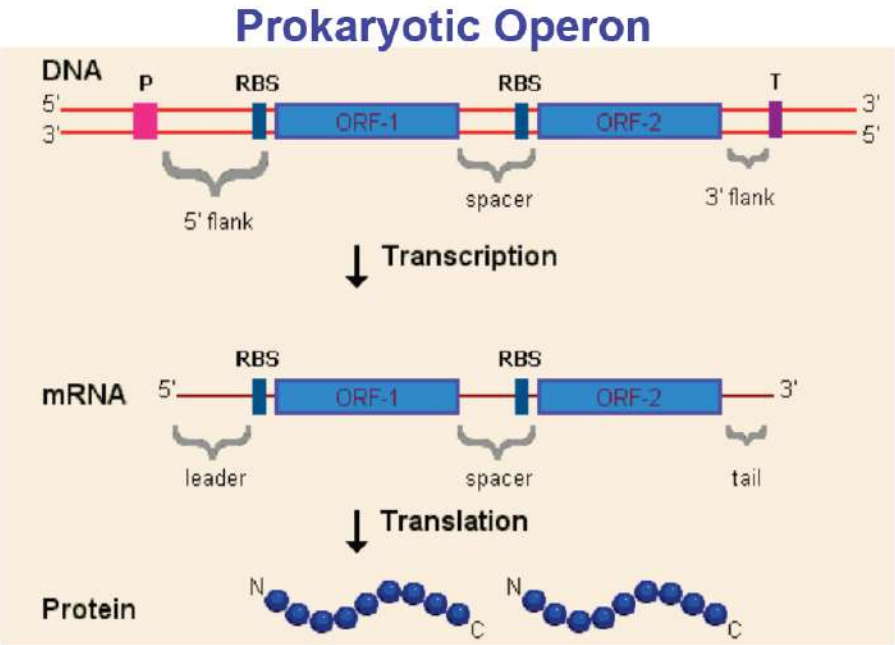
Differences between DNA and RNA

• contains ribose instead of deoxyribose

• single s

Transcriptional unit: region of DNA defined by a promoter and a terminator

• ORF: open reading frame- region of DNA defined by a start codon continuing to the stop codon that corresponds to a protein



Leader and tail are on primary transcript- play a role in the recognition and stability of mRNA in ribosome complex

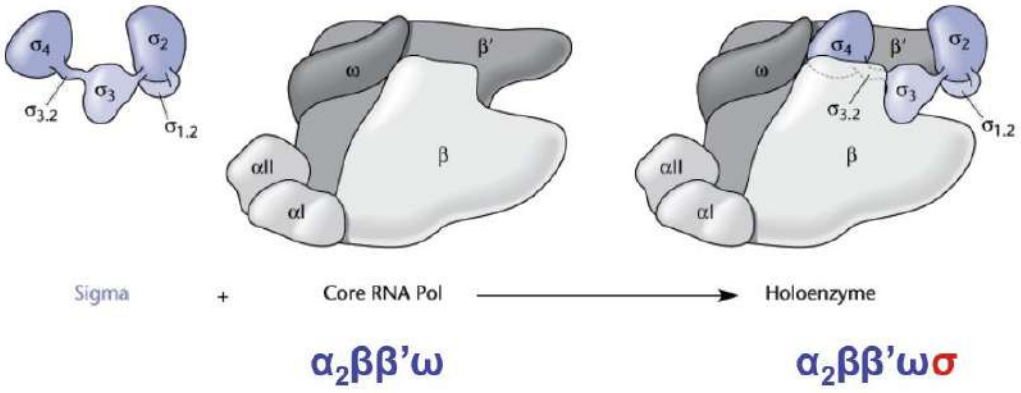
leader: upstream: in the 5' direction
tail: downstream: in the 3' direction

RNA Synthesis

RNA polymerase (RNAP):

• Core Enzyme: α2ββ'ω

• Holoenzyme :α2ββ'ωσ



3 Steps of Transcription

1) Initiation –

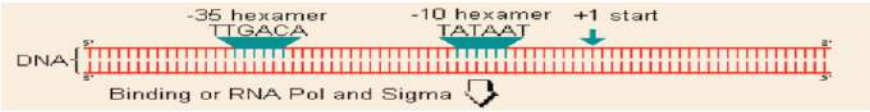
2) Elongation –

3) Termination – Transcription stops due to terminator sequence

Initiation

a) Promoter recognition: closed complex

- **Promoters possess a consensus sequence**



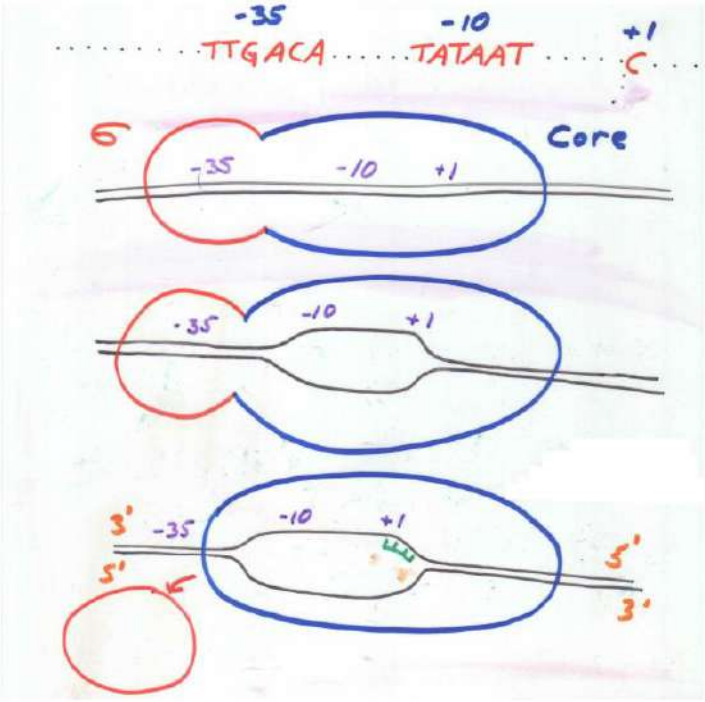
- **Typical *E. coli* promoter sequence**

  -10 box: TATAAT
  -35 box: TTGACA

- **σ subunit**

Initiation

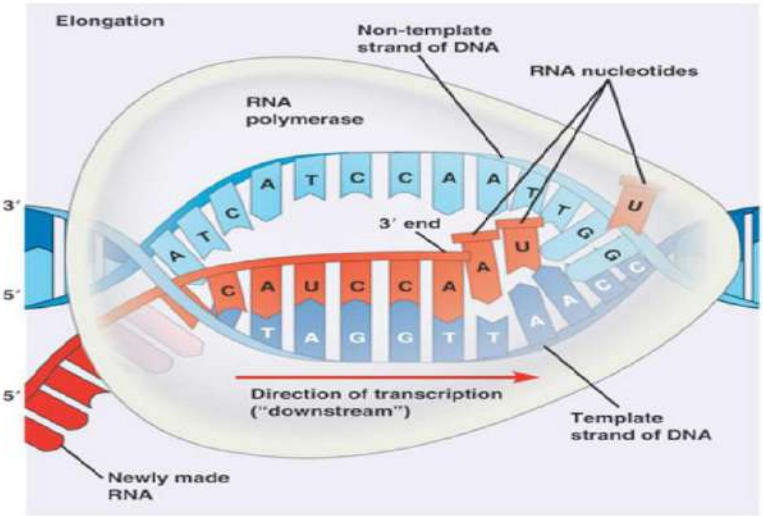
b) Unwinding of promoter region: open complex



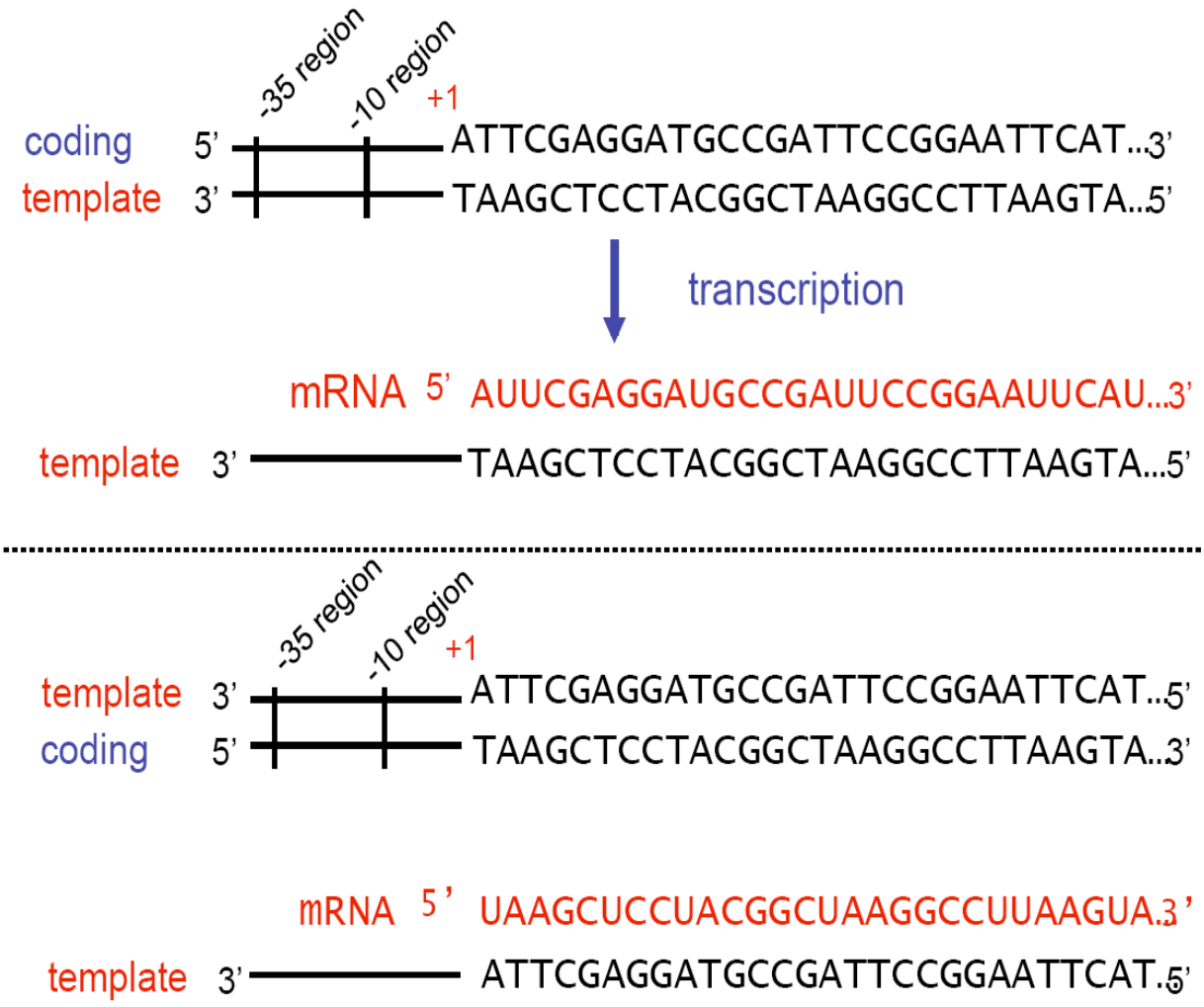
- Txn begins at the +1 site
- σ subunit is released once transcript synthesis is initiated

Elongation

- Template/Non-coding Non-template/Coding strand



http://fig.cox.miami.edu/~cmallery/150/gene/c7.17.7b.transcription.jpg



2 types of Termination

1) factor-independent

Transcription stops at terminator sequences

• Inverted repeats:

• Stretch of AAAAAAA…. on template strand

5'A TTA-TAGCGGCCATC-ACTGTTACA-GATGGCCGCTA-TTTT **3'**

3' TAAT-ATCGCCGGTAG-TGACAATGT-CTACCGGCGAT-AAAA **5'**

>>>>>>>>>>>

<<<<<<<<<<<

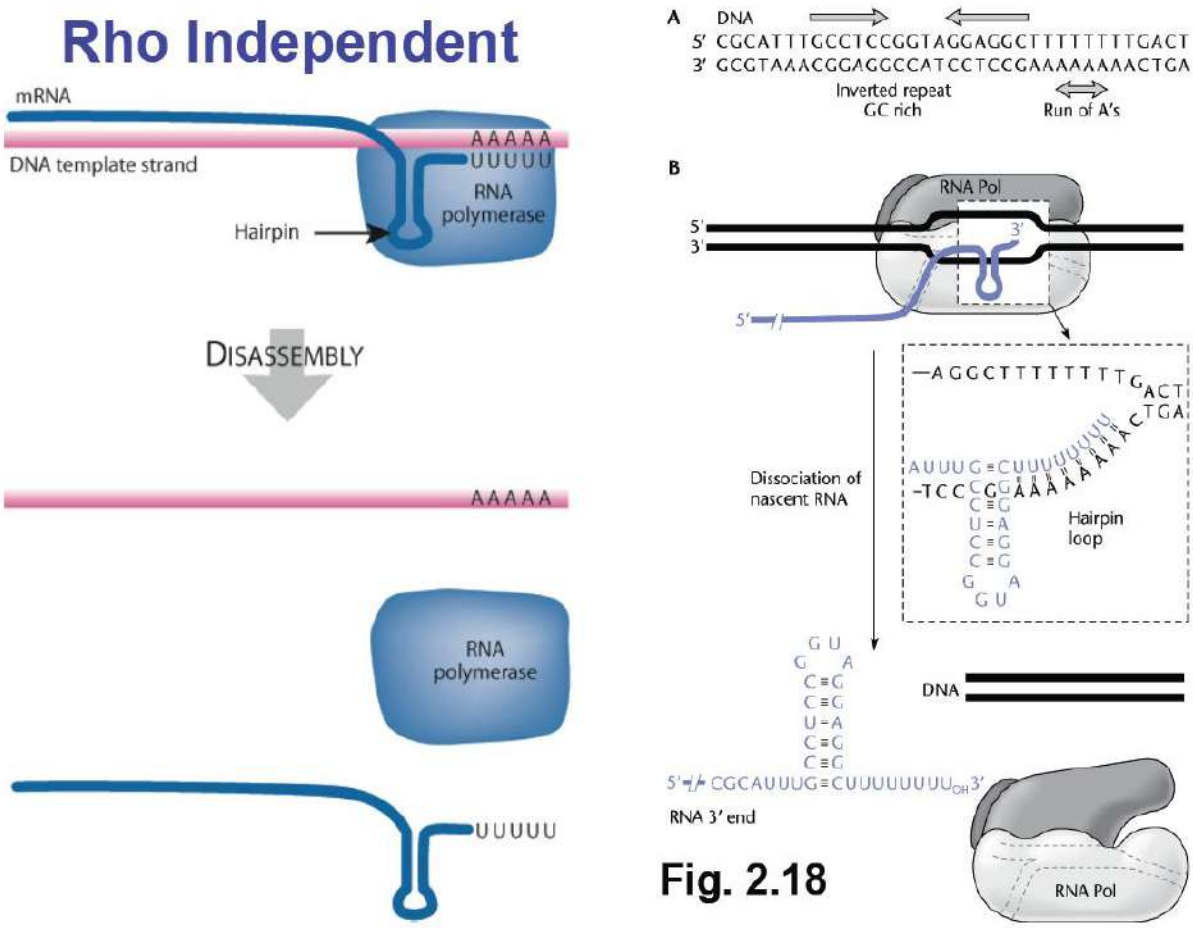


Fig. 2.18

Termination continued

2) Rho-dependent

• Rho protein binds to a *rut* site of the newly synthesized mRNA and chases the RNA polymerase

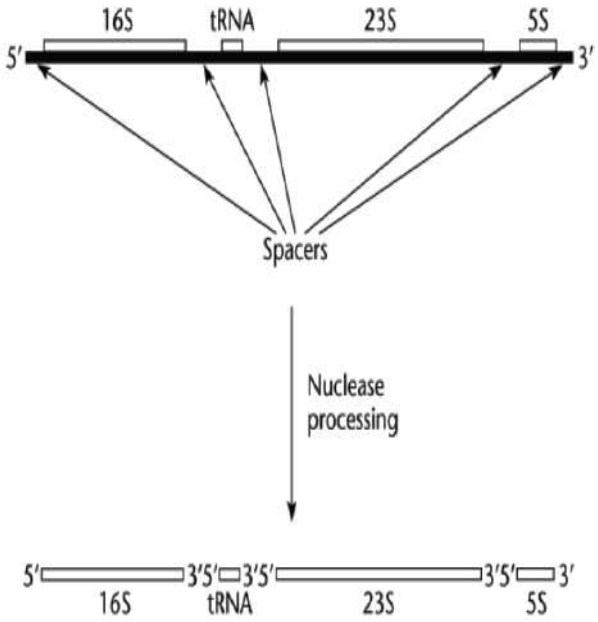
• When RNAP stalls at the terminator sequence, Rho uses a RNA:DNA helicase activity destroy the transcription bubble.

RNA Processing and Modification

Processing: involves forming and breaking phosphodiester bonds

Modification: involves modifying or adding specific bases to RNA

tRNA contains various modified bases

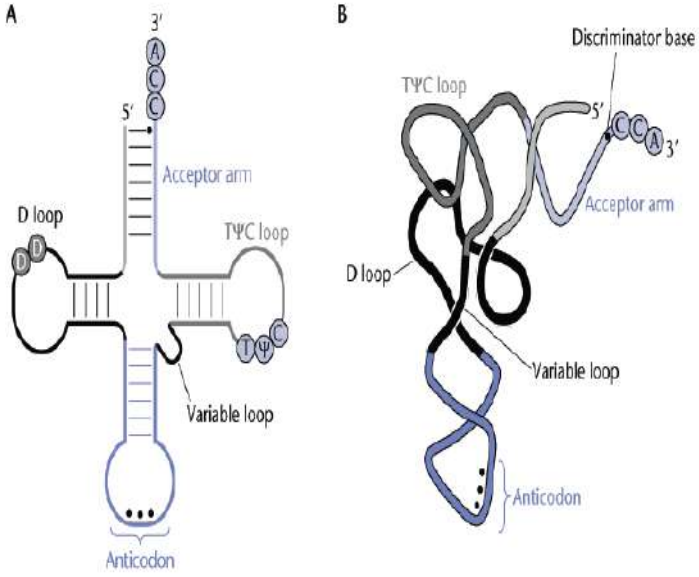


RNA processing

Figure 2.20

RNA modification

Figure 2.21



Rho Dependent

Figure 2.19 in text

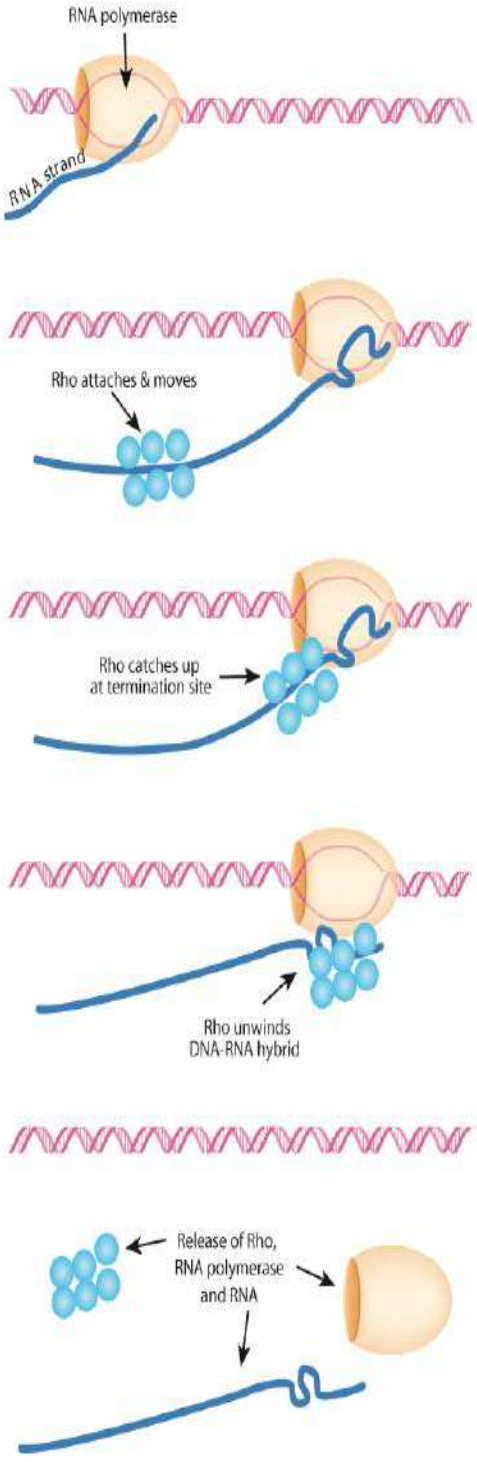
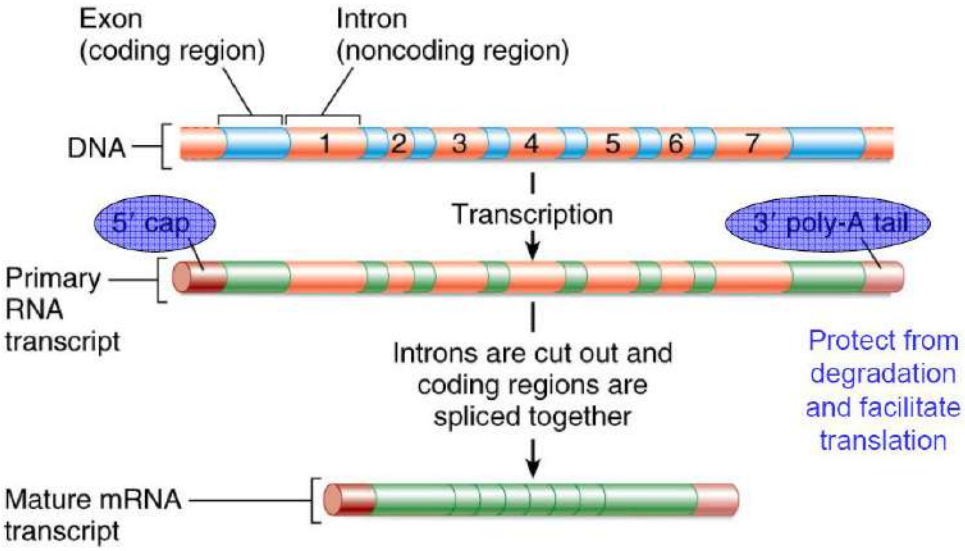


Fig. 9.17 Processing eukaryotic mRNA



- Different combinations of exons can generate different polypeptides via alternative splicing

Lecture : 10

Tranlation

RNA - Translation - Protein

The Process of 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:

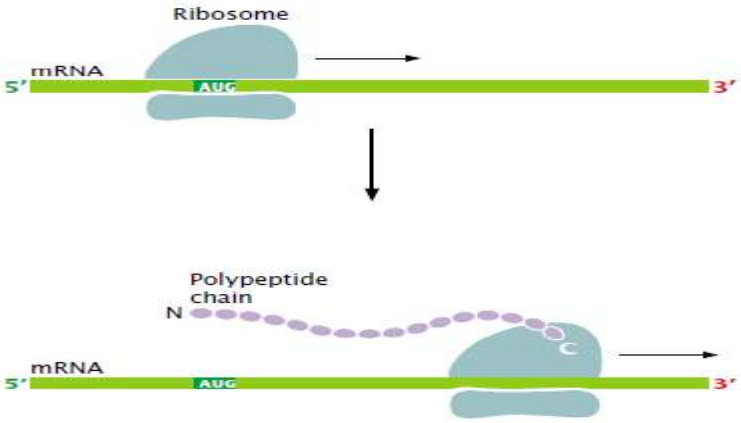
- Proteins
- rRNAs

B- tRNAs and tRNA synthetases:

C-mRNAs

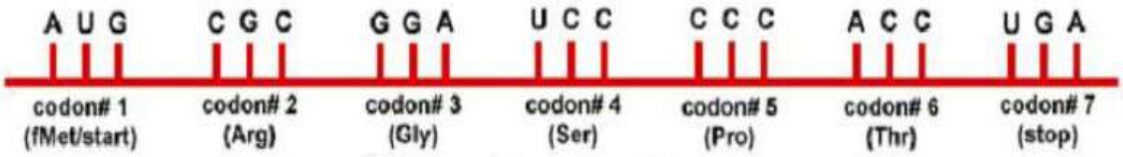
D-Accessory proteins:

- Initiation factors
- Elongation factors
- Release factors

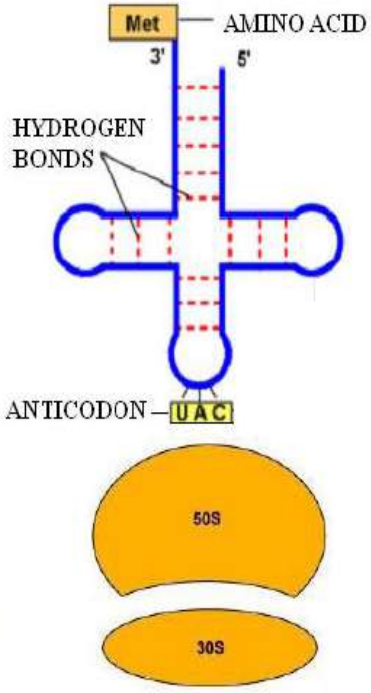


There are three types of RNA :

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



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.

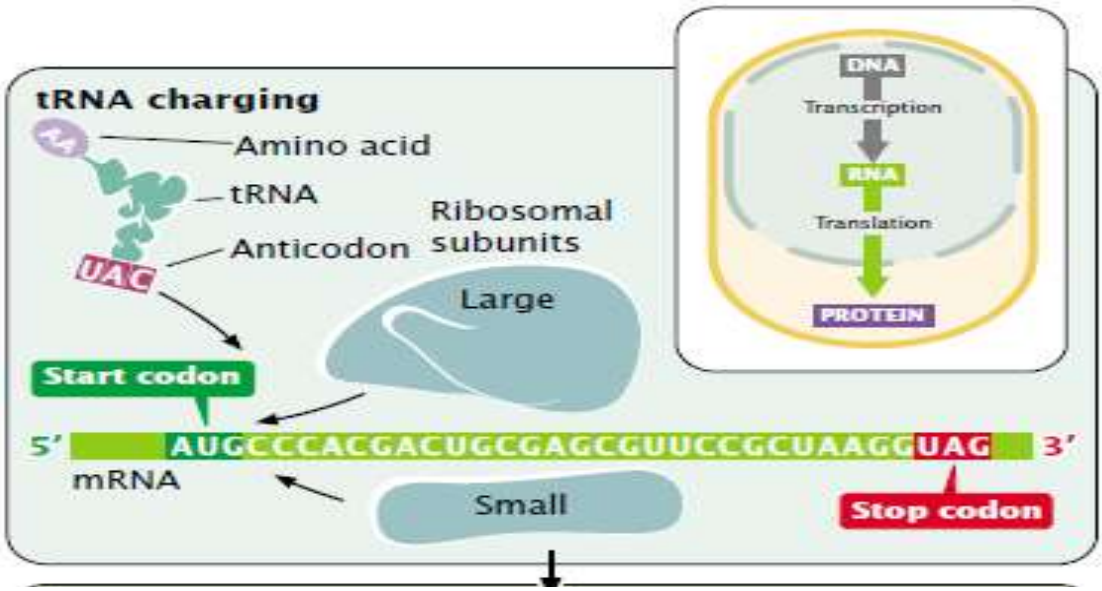
(4) Termination, in which protein synthesis halts at the termination codon and the translation components are released from the ribosome.

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 Stage :

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.

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.

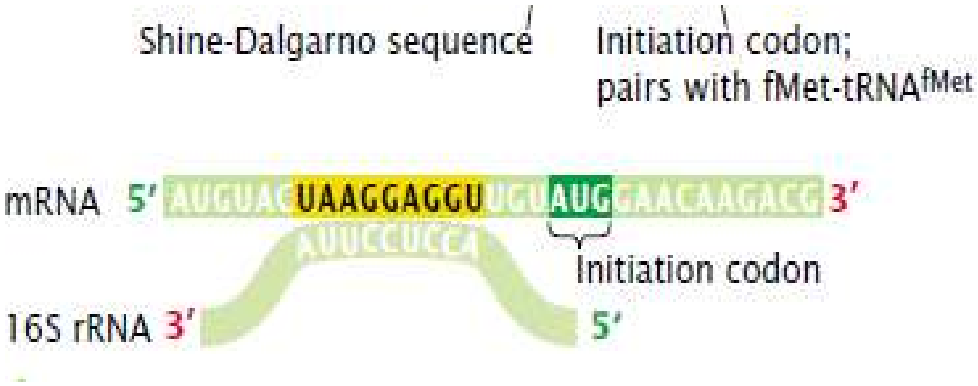
anti-Shine-Dalgarno sequence Sequence on 16S rRNA that is complementary to the Shine-Dalgarno sequence of mRNA

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

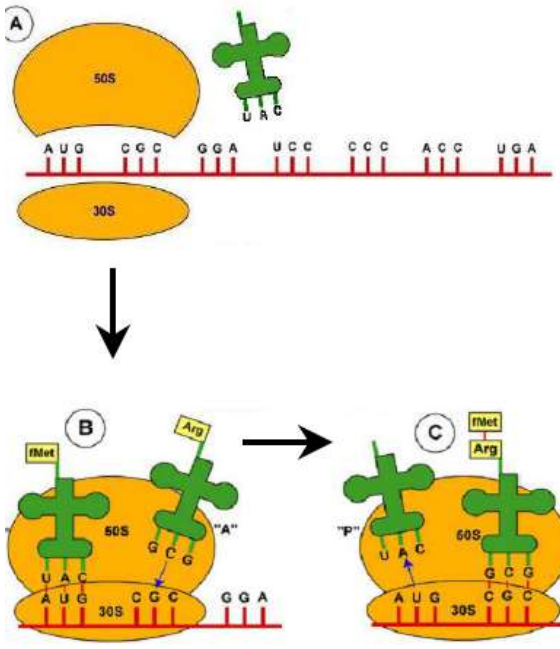
ribosome binding site (RBS) Same as Shine-Dalgarno sequence; sequence close to the front of mRNA that is recognized by the ribosome; only found in prokaryotic cells

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



.

INITIATION



(A) RIBOSOME, tRNA WITH START ANTICODON, AND mRNA COME TOGETHER

(B) START ANTICODON OF tRNA BINDS TO CODON ON mRNA

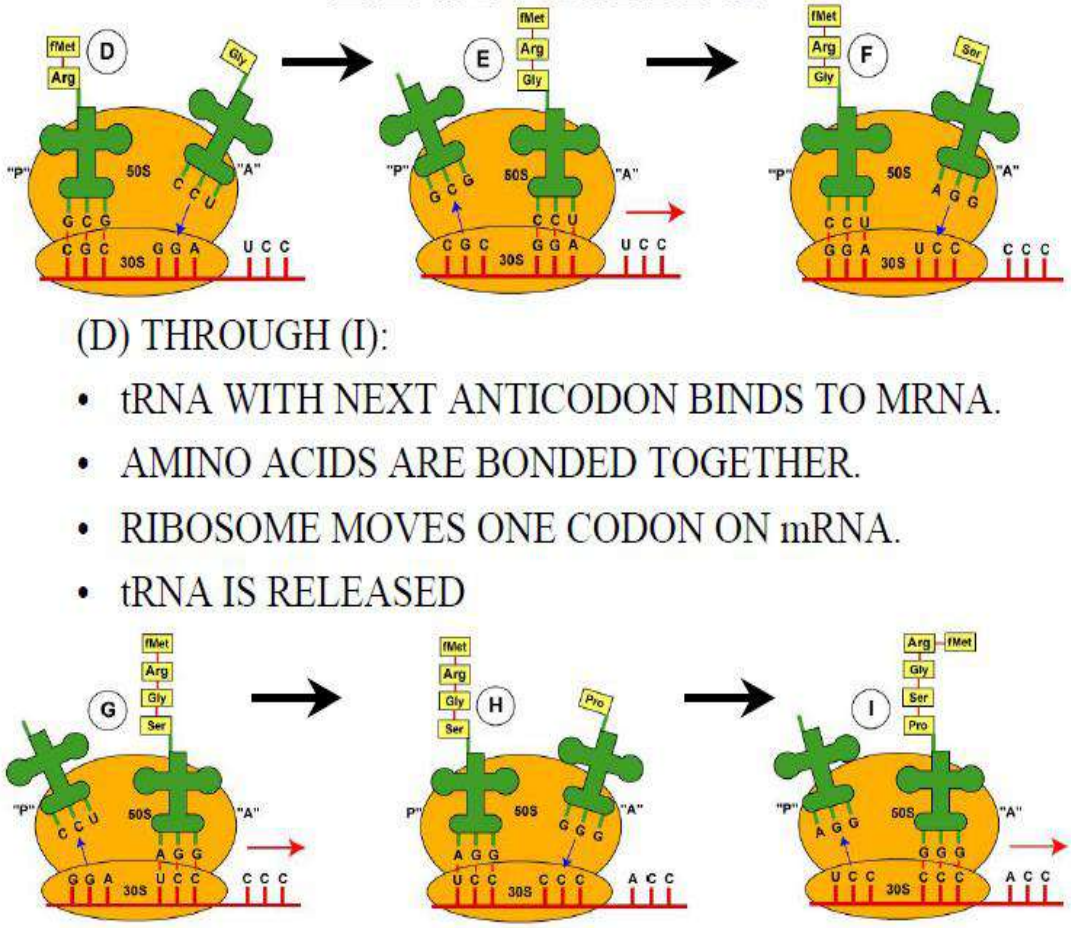
(C) tRNA WITH SECOND ANTICODON BINDS TO SECOND CODON ON mRNA.

AMINO ACIDS ARE BONDED COVALENTLY TOGETHER.

RIBOSOME MOVES ONE CODON ON mRNA.

FIRST tRNA IS RELEASED.

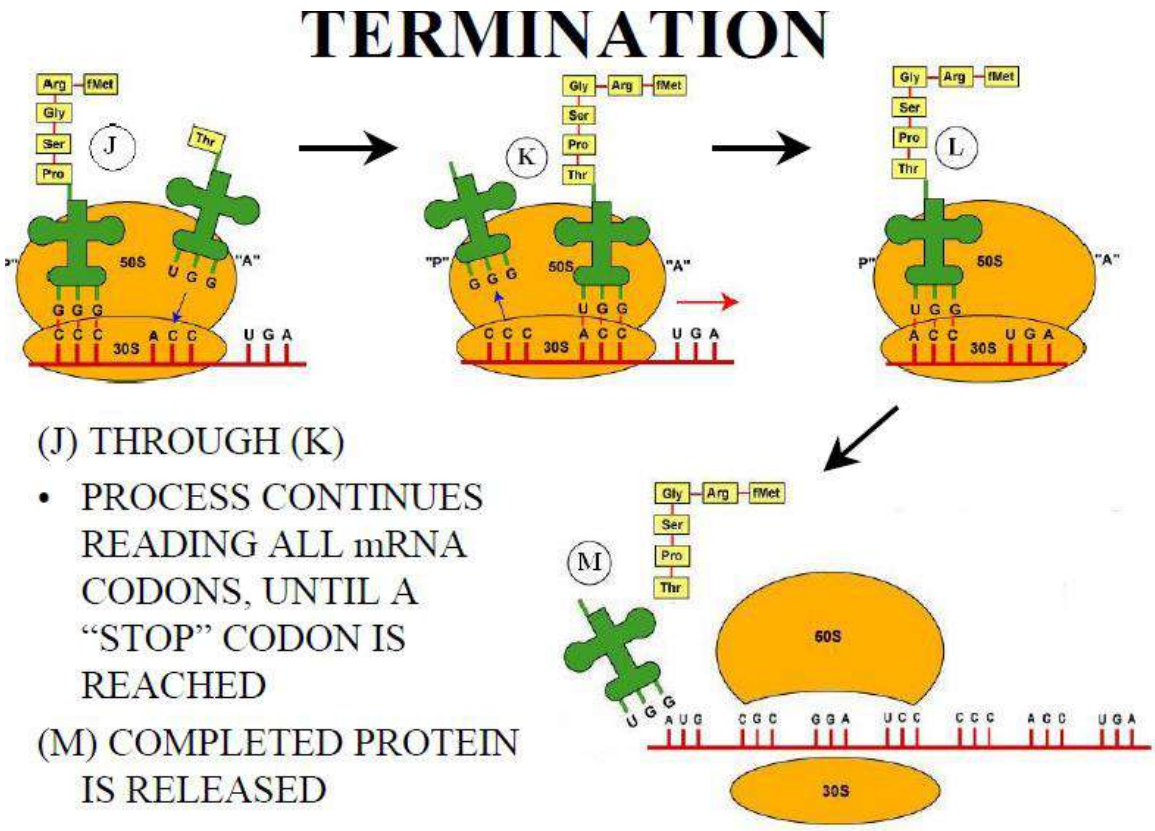
ELONGATION



(D) THROUGH (I):

- tRNA WITH NEXT ANTICODON BINDS TO MRNA.
- AMINO ACIDS ARE BONDED TOGETHER.
- RIBOSOME MOVES ONE CODON ON mRNA.
- tRNA IS RELEASED

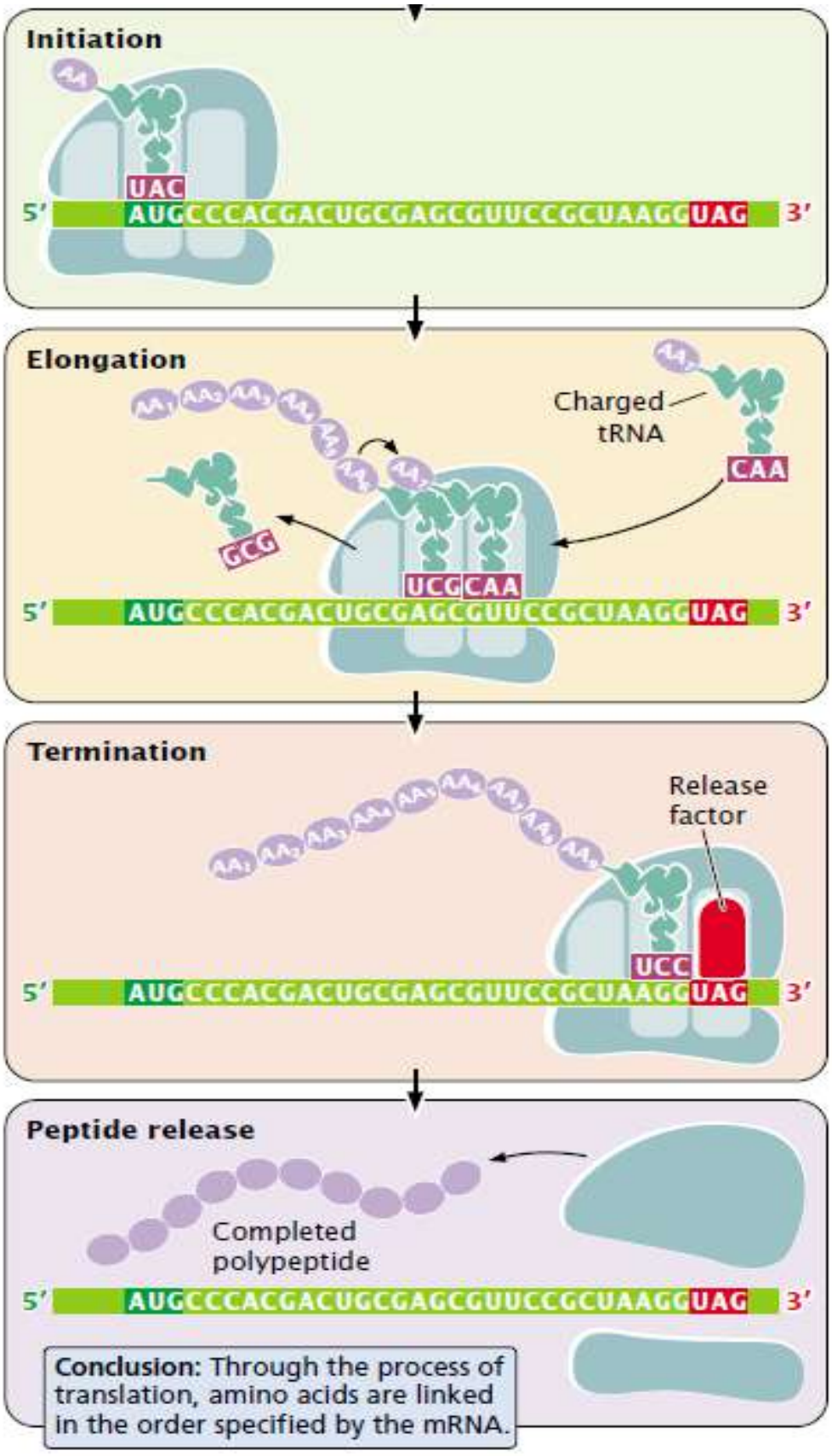
Elongation consists of three steps: (1) a charged tRNA enters the A site, (2) a peptide bond is created between amino acids in the A and P sites, and (3) the ribosome translocates to the next codon. Elongation requires several elongation factors (EF-Tu, EF-Ts, and EF-G) and GTP.



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.

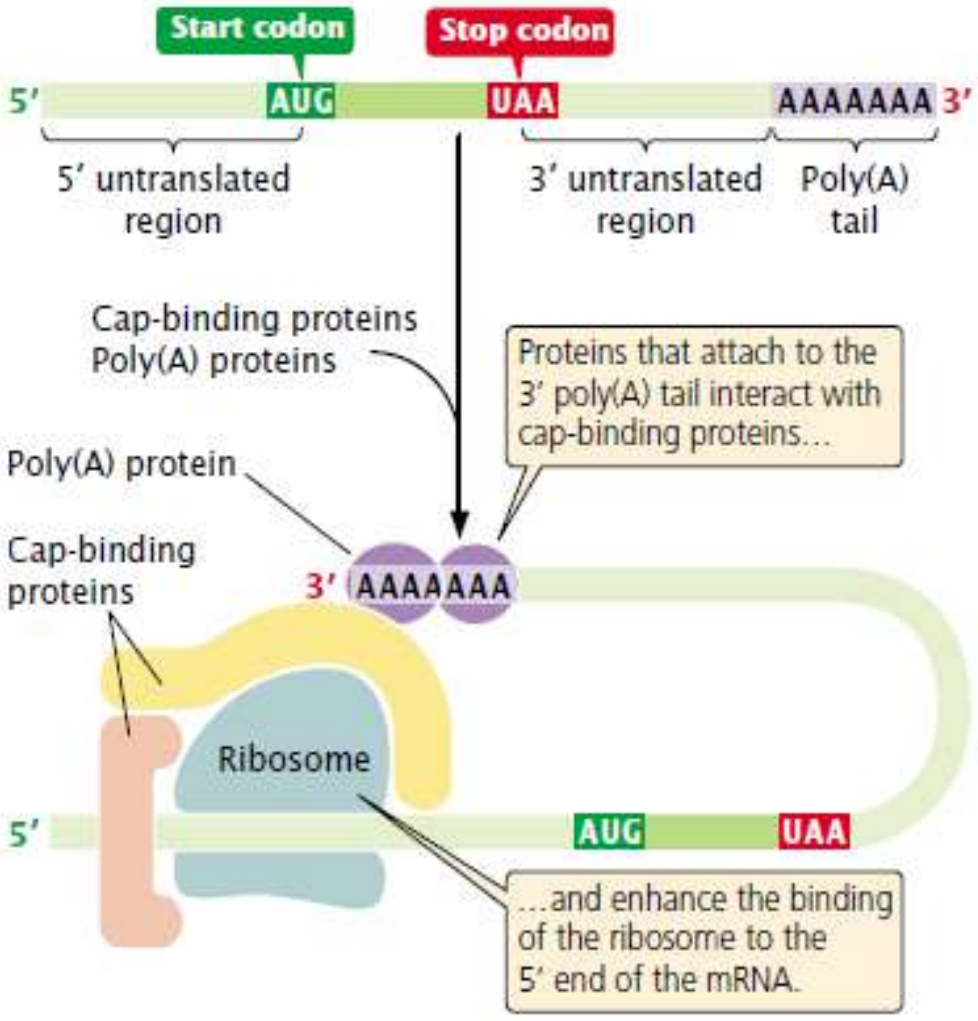
Concepts

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


Initiation
UAC
5′ AUGCCCACGACUGCGAGCGUUCCGCUAAGGUAG 3′
Elongation
Charged tRNA
CAA
GCG
UCG CAA
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.