FACULTY OF SCIENCE
Chemistry department
Biochemistry division

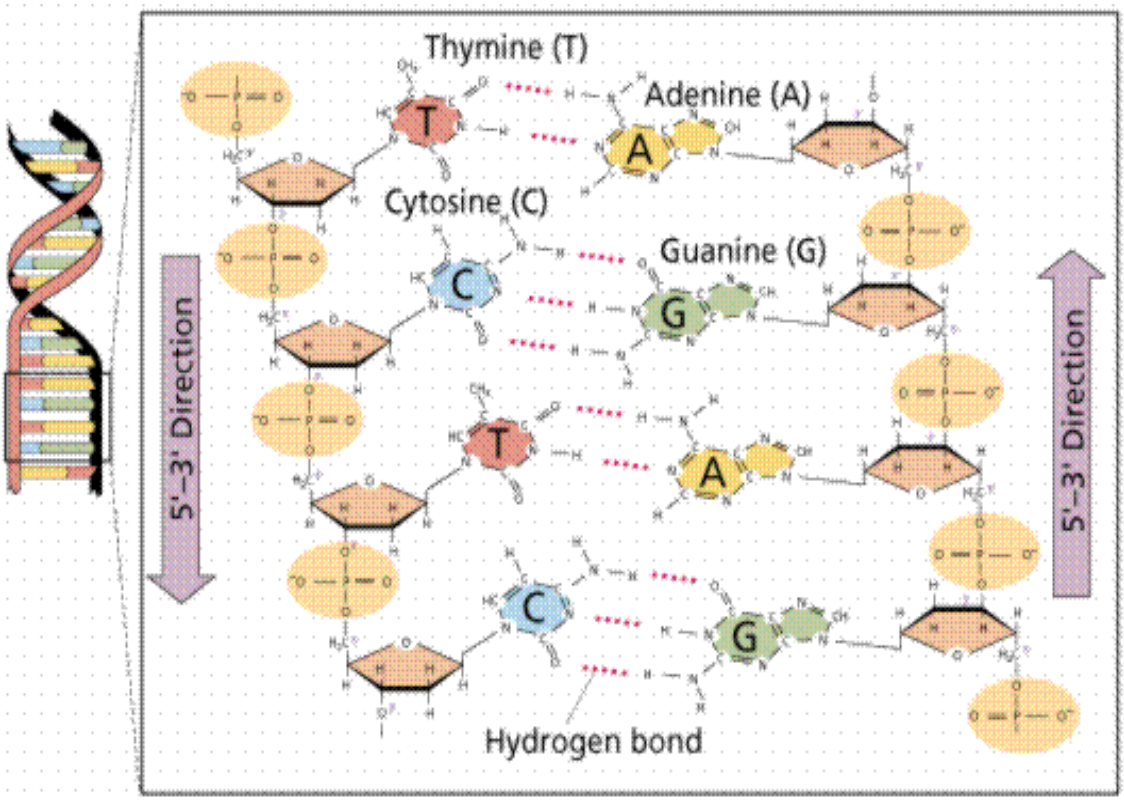
DNA Technology

Postgraduate students

BY

Dr. Walaa Hozayen

2017



Contents

	Page No.
DNA STRUCTURE	**3**
Central Dogma of Molecular Biology	**14**
DNA Replication	**15**
RNA TRANSCRIPTION	**26**
Translation Process	**40**
Tools and molecular techniques	**52**
Applications of PCR	**60**
Mutations	**64**
References	**69**

DNA STRUCTURE

The Structure and Function of DNA as the Molecule of Inheritance:

The DNA's function is to serve as the reservoir of all the genetic information of the organism. This molecule contains genes, and one definition is that they are stretches (pieces) of the DNA molecule that code for specific functions. Crudely speaking, each gene will code for a specific protein that will carry out a specific function in the complex web of the cell's biochemical metabolism. Everything that the organism has to do or make is specified in this DNA macromolecule: when a cell will develop, which type of cell will develop, when to live or die is all specified in the genetic material (DNA).

In addition, the DNA molecule functions as the instrument to pass on information from one generation to the other. When the cell has to duplicate, the DNA helix splits apart and each strand serves as a template to make a complementary strand of itself. This way all the information stored in the parent's DNA is passed on to its offspring.

In humans the genes are contained in 46 chromosomes - you inherit one pair (23 chromosomes) from your father and one pair from your mother.

Experiments Identified the Genetic Material

The nucleic acid DNA (deoxyribonucleic acid) is one of the most familiar molecules, the subject matter of movies and headlines (figure 1). Fictional scientists reconstruct dinosaurs from DNA preserved in an ancient mosquito's gut. Criminal trials hinge on DNA evidence; the idea of cloning animals raises questions about the role of DNA in determining

who we are; and DNA-based discoveries are yielding new diagnostic tests, treatments, and vaccines.

More important than DNA's role in society today is its role in life itself. Of all the characteristics that distinguish the living from the nonliving, the one most important to the continuance of life is the self-replicating cell. At the molecular level, reproduction depends on DNA, a biochemical that has dual abilities. First, it directs the activities of the cell by controlling protein synthesis. Second, it manufactures an exact replica of itself, copying those instructions for the next generation of cells.

The recognition of DNA's vital role in life was a long time in coming. By the early 1900s, researchers had recognized the connection between inheritance and protein. For example, the English physician, Archibald Garrod, noted that people with inherited "inborn errors of metabolism" lacked certain enzymes. Other researchers added supporting evidence: they linked abnormal or missing enzymes to unusual eye color in fruit flies and nutritional deficiencies in bread mold. But how were enzyme deficiencies and inheritance linked? Experiments in bacteria would answer the question.

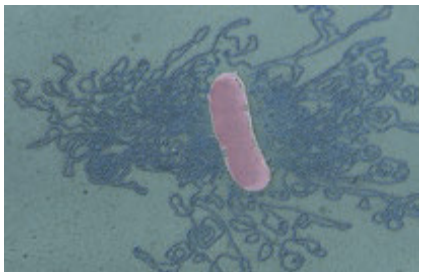


FIGURE 1 DNA—The Molecule in the Media.

(**A**) DNA bursts forth from this treated bacterial cell, illustrating just how much DNA is tightly wound into a single cell. (**B**) *Jurassic Park* was a 1993 blockbuster movie in which fictional scientists recreated dinosaurs. The dinosaur DNA came from blood found in ancient mosquitoes entombed in amber.

A. Griffith Discovered that Bacteria Can Transfer Genetic Information:

In 1928, English microbiologist Frederick Griffith contributed the first step in identifying DNA as the genetic material. Griffith studied mice with pneumonia caused by a bacterium, *Streptococcus pneumoniae.* He identified two types of bacteria: type R and type S (figure 2). Type R bacteria form rough-shaped colonies, and when injected into mice, they do not cause pneumonia.

Type S bacteria form smooth colonies because they are encased in a polysaccharide capsule. When injected into mice, type S bacteria cause pneumonia. Therefore, the smooth polysaccharide
coat seemed to be necessary for infection.

When Griffith heated type S bacteria ("heat-killing" them) and injected them into mice, they no longer caused pneumonia.

However, when he injected mice with a mixture of type R bacteria plus heat-killed type S bacteria, neither of which was able to cause pneumonia alone, the mice died of pneumonia. Their bodies contained live type S bacteria encased in polysaccharide.

How had the previously harmless bacteria acquired the ability to cause disease? In the 1940s, U.S. physicians Oswald Avery, Colin MacLeod, and Maclyn McCarty offered an explanation.

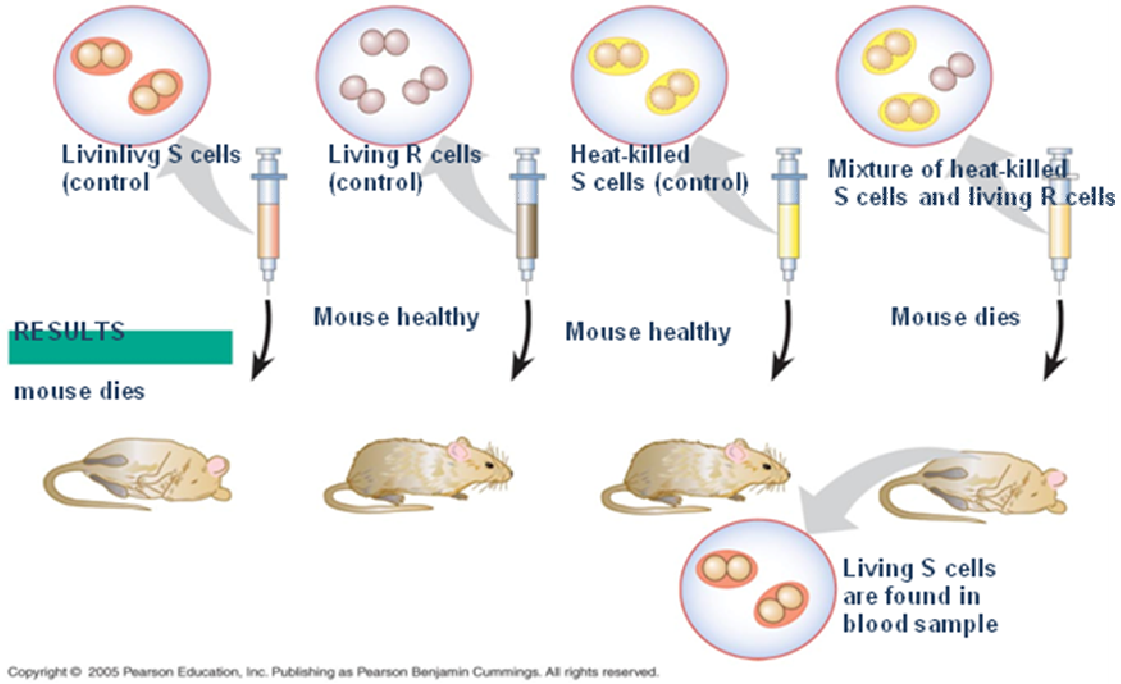
B. Avery, MacLeod, and McCarty Showed that Genetic Information Is DNA

Avery, MacLeod, and McCarty hypothesized that something in the heat-killed type S bacteria entered and "transformed" the normally harmless type R strain into a killer. Was this "transforming principle" a protein? Treating the solution from the type S strain with a protein destroying enzyme (a protease) failed to keep the type R strain from being

transformed into a killer (figure 7.3). Therefore, a protein was not responsible for transmitting the killing trait. Treating the solution from the heat-killed S bacteria with a DNA-destroying enzyme (DNase) first, however, prevented the killing ability.

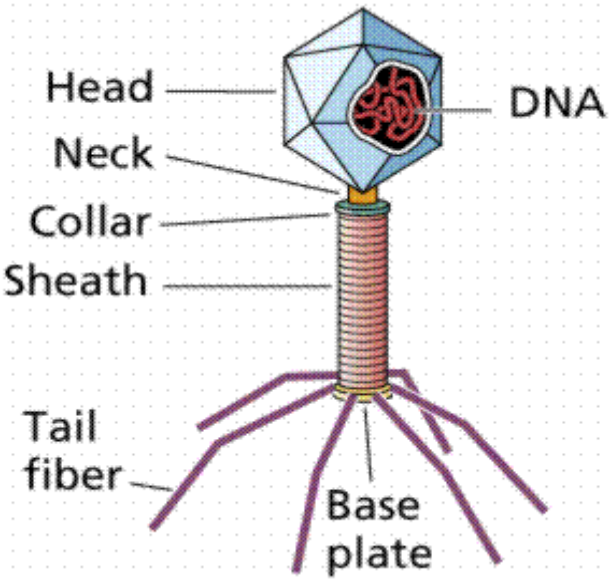
Avery, MacLeod, and McCarty confirmed that DNA transformed the bacteria by isolating DNA from heat-killed type S bacteria and injecting it along with type R bacteria into mice. The mice died, and their bodies contained active type S bacteria. The conclusion: Type S DNA altered the type R bacteria, enabling them to manufacture the smooth coat necessary to cause infection.

At first, biologists hesitated to accept DNA as the biochemical of heredity. They knew more about proteins than about nucleic acids. They also thought that protein, wit its 20 building blocks, was able to encode many more traits than DNA, which includes just four types of building blocks. In 1950, however, U.S. microbiologists Alfred Hershey and Martha Chase conclusively showed that DNA—not protein—is the genetic material.



C. Hershey and Chase Confirmed the Genetic Role of DNA

Hershey and Chase used a very simple system; they infected the bacterium *Escherichia coli* with a bacteriophage called T4. A bacteriophage is a virus that infects only bacteria; most consist of only a protein coat and a nucleic acid core (DNA in this case). We now know that when the virus infects the bacterial cell, it injects its DNA, and the protein coat remains loosely attached to the bacterium (figure 3).



The viral DNA uses the bacterial cell's energy and raw materials to manufacture more of itself. New virus particles then burst from the cell. Hershey and Chase wanted to know which part of the virus controls its replication: the DNA or the protein coat.

The researchers "labeled" two batches of viruses, one with radioactive sulfur that marked protein, and the other with radioactive phosphorus that marked DNA. They used each type of labeled virus to infect a separate batch of bacteria and allowed several minutes for the virus particles to bind to the bacteria and inject their DNA into them (figure 4). Then they agitated each mixture in a blender, which removed the unattached viruses and empty protein coats from the surfaces of the bacteria. They poured the mixtures into test tubes and spun them at high speed. This settled the infected bacteria at the bottom of each test tube because they were heavier than the liberated viral protein coats.

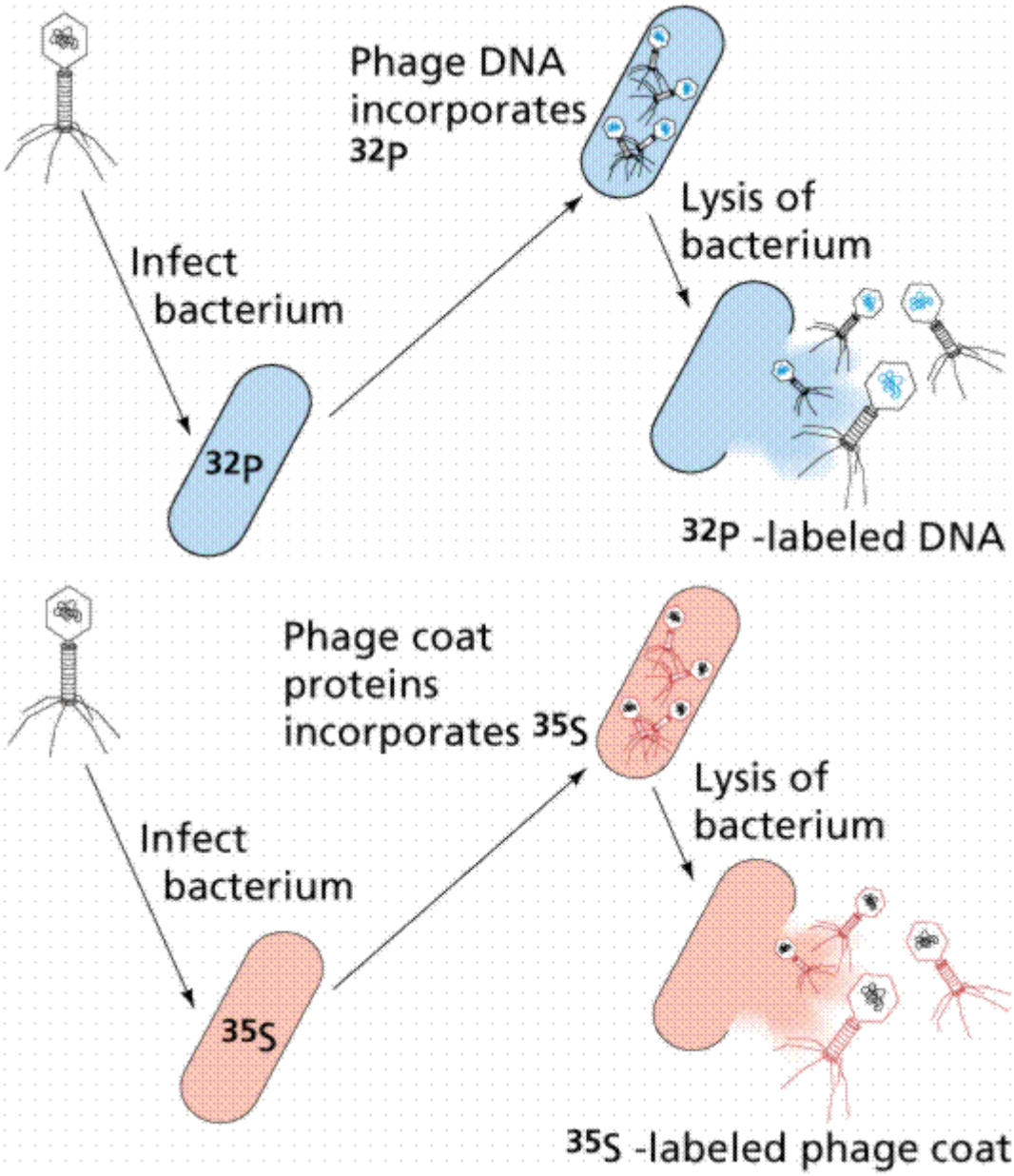


FIGURE 4

Diagrams illlustrating the Hershey and Chase experiment that supported DNA as the hereditary material while it also showed protein was NOT the hereditary material.

Summary A series of experiments revealed DNA, and not protein, to be the genetic material. Further experiments provided clues to its structure.

DNA Is a Double Helix That Encodes "Recipes" for Proteins

The early twentieth century also saw corresponding advances in the study of the structure of DNA. By 1929, biochemists had discovered the distinction between ribonucleic acid (RNA) and DNA, the two types of nucleic acid. Later, they determined that nucleotides, the building blocks of nucleic acids, included sugars, nitrogen-containing groups, and phosphorus-containing components. Another important clue was the observation that DNA and RNA nucleotides always contain the same sugars and phosphates, but they may contain any one of four different nitrogen-containing bases.

Biochemists and Physicists Discovered DNA's Structure

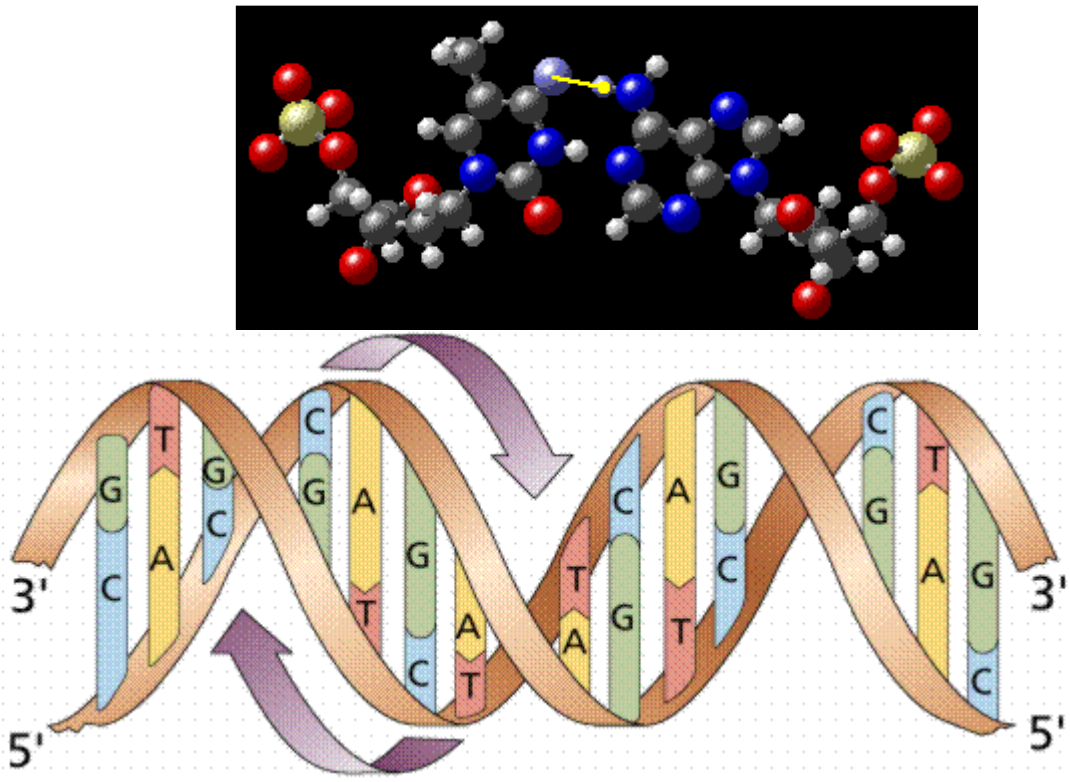
In the early 1950s, two lines of evidence together revealed DNA's chemical structure. Austrian-American biochemist Erwin Chargaff showed that DNA contains equal amounts of the bases adenine (A) and thymine (T) and equal amounts of the bases guanine (G) and cytosine (C). English physicist Maurice Wilkins and chemist Rosalind Franklin bombarded DNA with X-rays, using a technique called X-ray diffraction to determine the three-dimensional shape of the molecule.

The X-ray diffraction pattern revealed a regularly repeating structure of building blocks (figure 5).

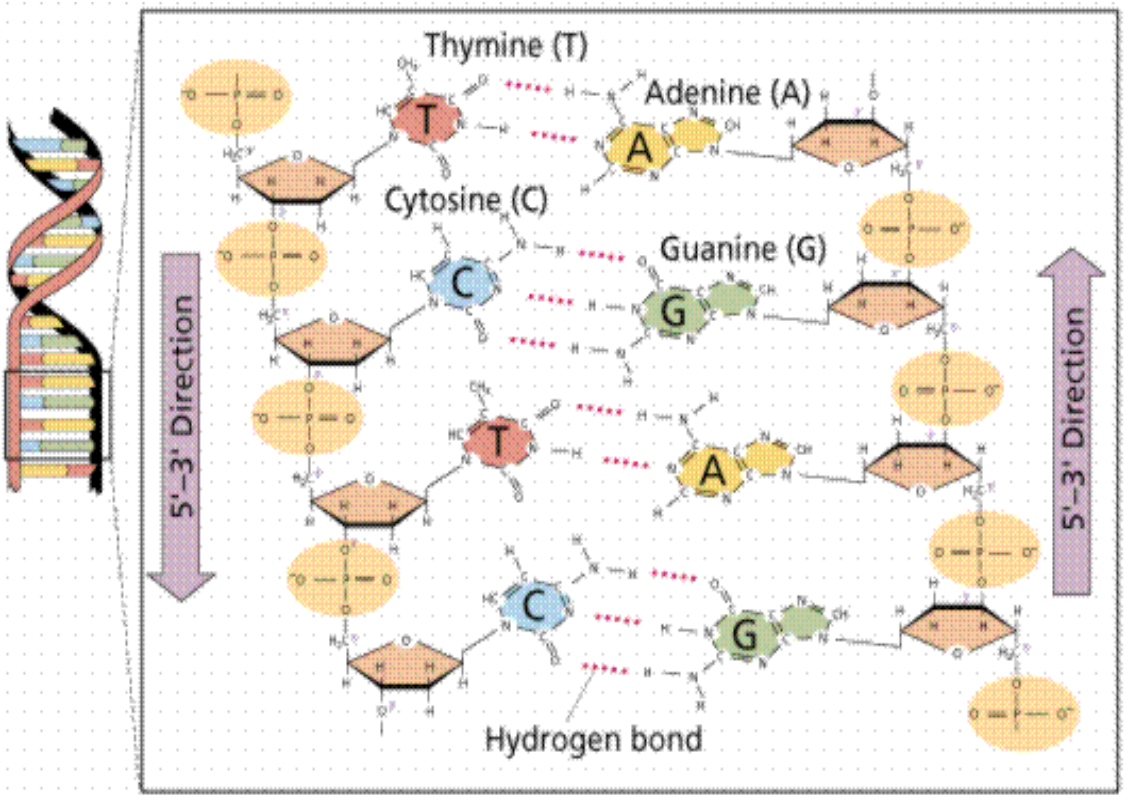
Watson and Crick's Model Fits the Data

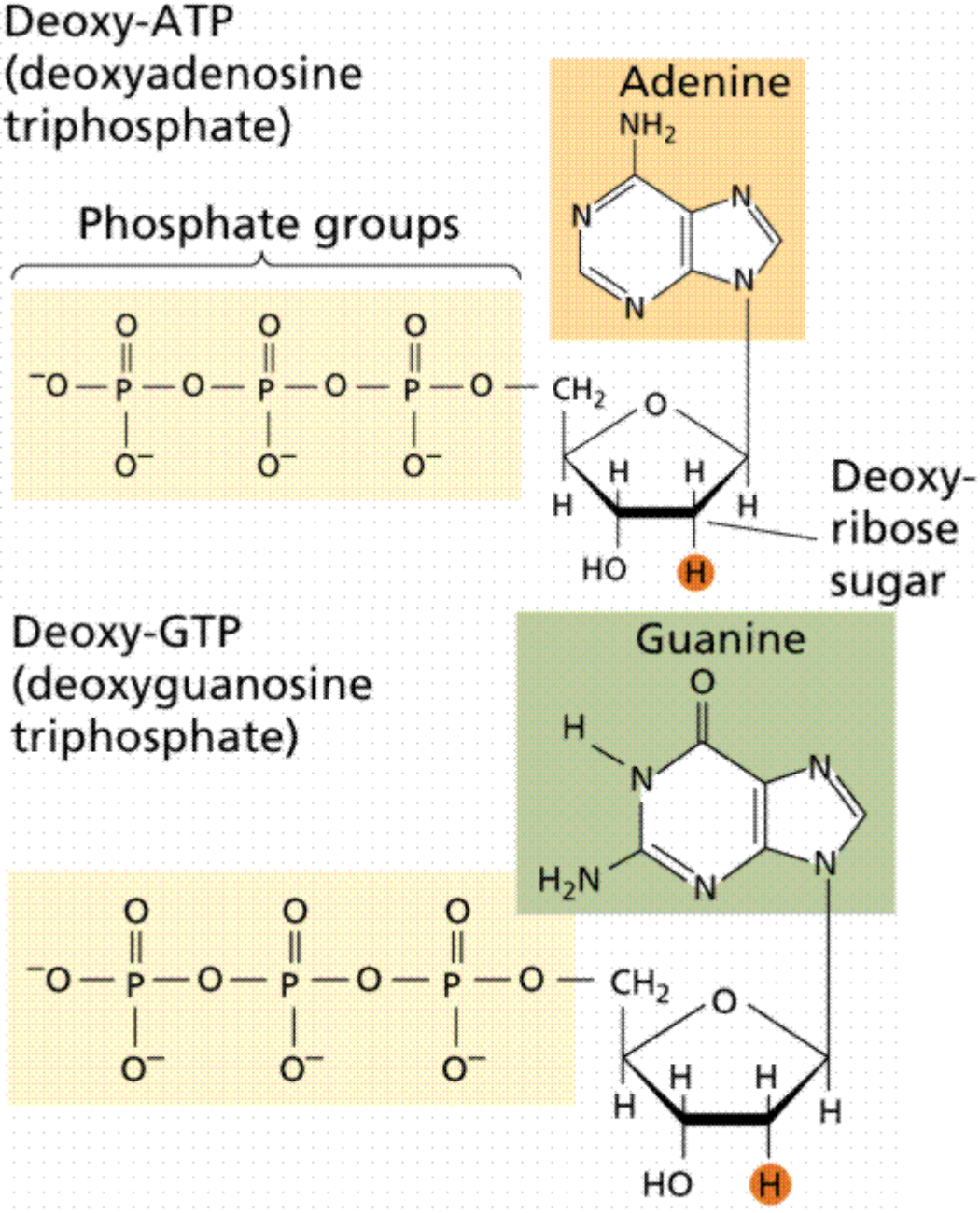
In 1953, U.S. biochemist James Watson and English physicist Francis Crick, working at the Cavendish laboratory in Cambridge in the United Kingdom, used these clues to build a ball-and-stick model of the DNA molecule. The now familiar double helix included equal amounts of G and C and of A and T, and it had the sleek symmetry revealed in the X-ray diffraction pattern. The DNA double helix resembles a twisted ladder.

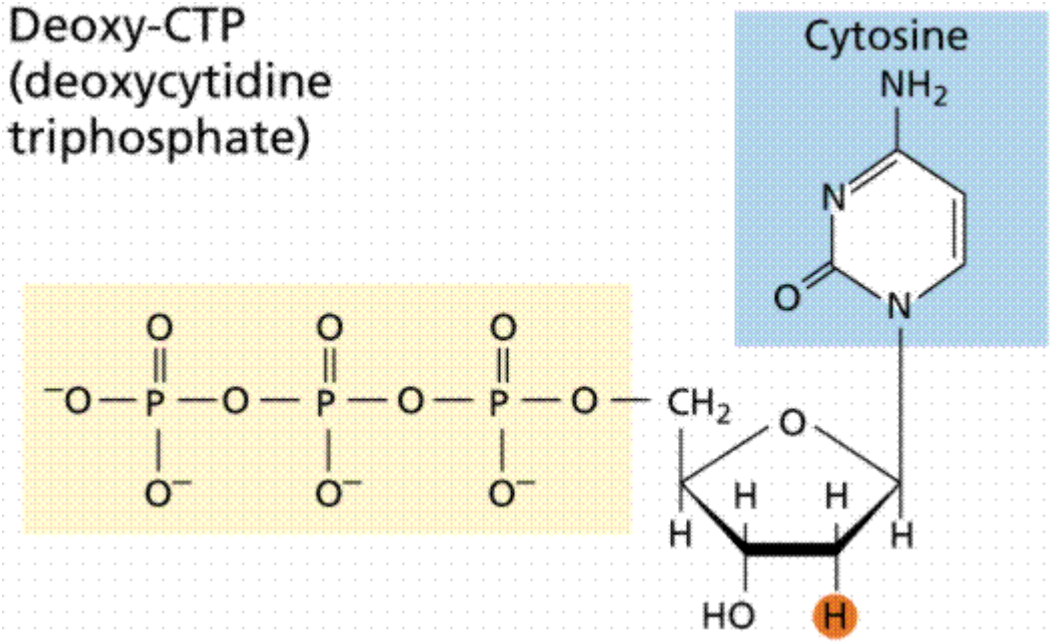
FIGURE 5

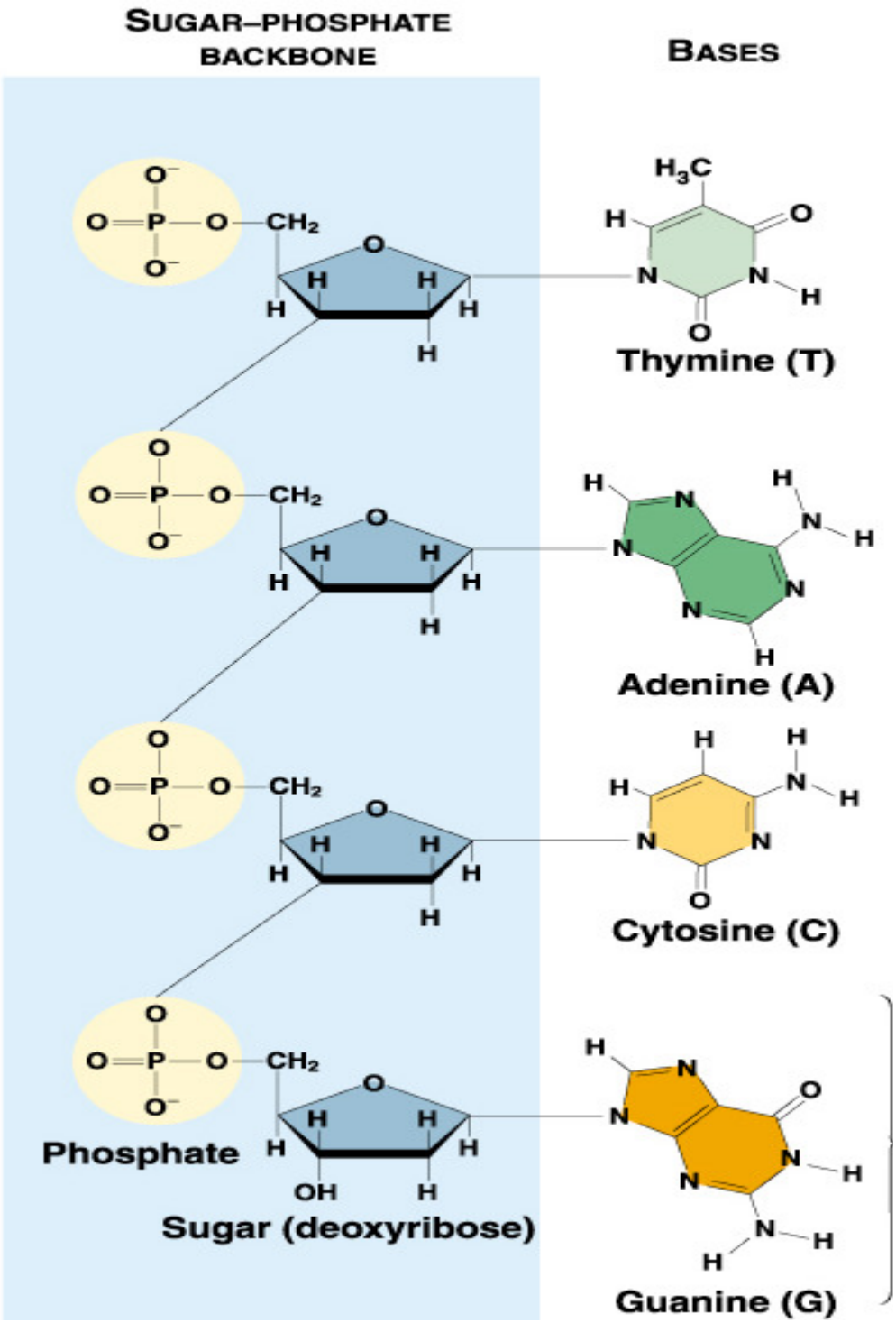


It contained four nitrogenous bases: cytosine, thymine, adenine, and guanine; deoxyribose sugar; and a phosphate group. He concluded that the basic unit (nucleotide) was composed of a base attached to a sugar and that the phosphate also attached to the sugar. He (unfortunately) also erroneously concluded that the proportions of bases were equal and that there was a tetranucleotide that was the repeating structure of the molecule. The nucleotide, however, remains as the fundemantal unit (monomer) of the nucleic acid polymer. There are four nucleotides: those with cytosine (C), those with guanine (G), those with adenine (A), and those with thymine (T).


Thymine (T)
Adenine (A)
Cytosine (C)
Guanine (G)
5'–3' Direction
5'–3' Direction
Hydrogen bond

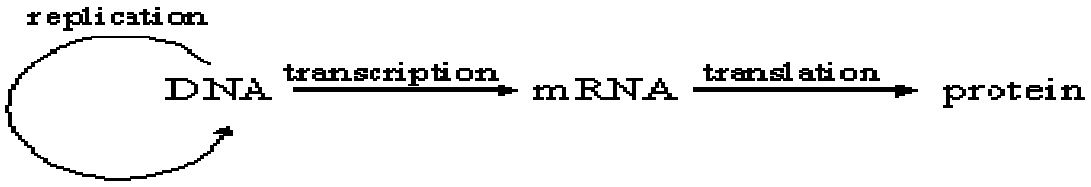

Deoxy-ATP
(deoxyadenosine
triphosphate)
Adenine
Phosphate groups
Deoxy-
ribose
sugar
Deoxy-GTP
(deoxyguanosine
triphosphate)
Guanine


Deoxy-CTP
(deoxycytidine
triphosphate)
Cytosine
NH2
N
O
N
O
P
O
O
O
O
O
O
O
O
O
CH2
O
H
H
H
H
HO
H



Central Dogma of Molecular Biology

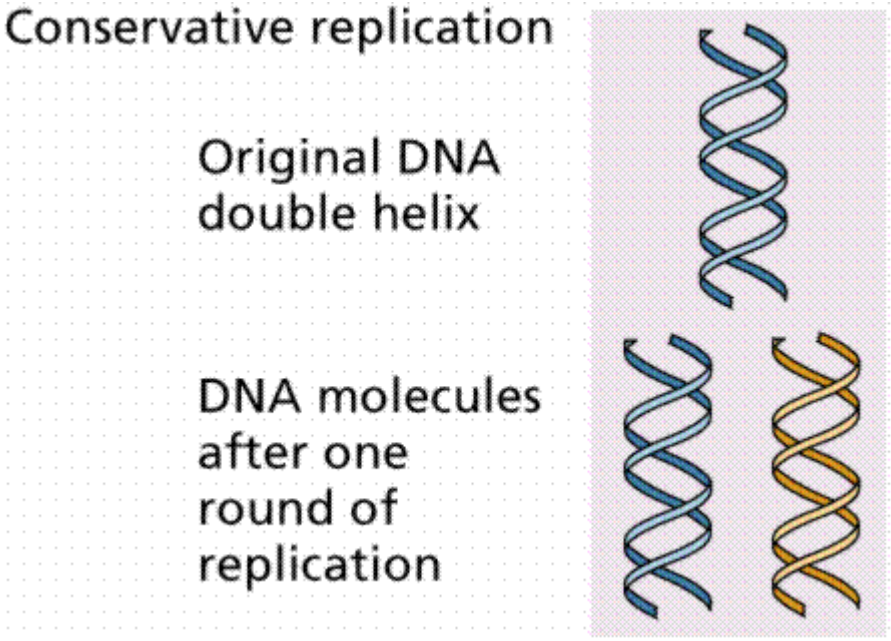
Protein synthesis requires two steps: ***transcription*** and ***translation***.



I - DNA Replication

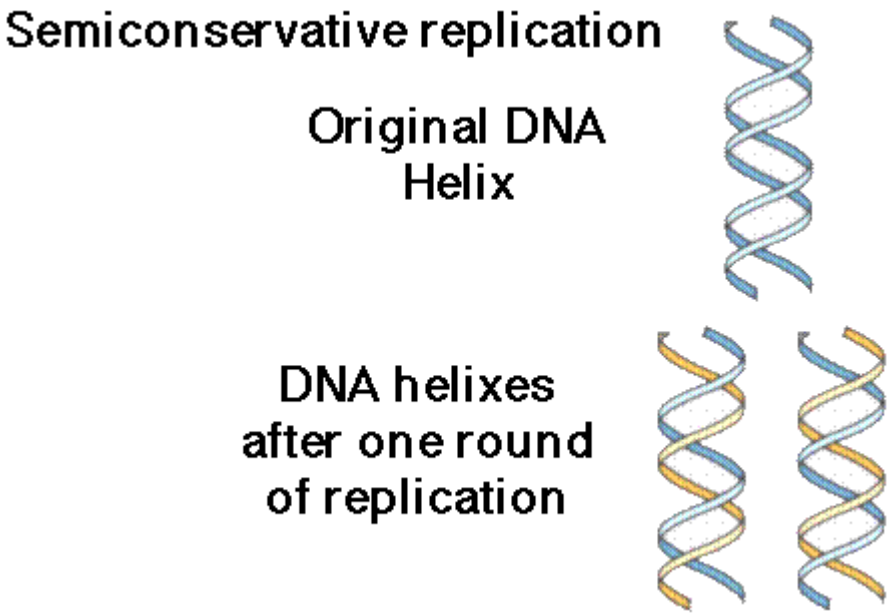
DNA was proven as the hereditary material and Watson et al. had deciphered its structure. What remained was to determine how DNA copied its information and how that was expressed in the phenotype. Matthew Meselson and Franklin W. Stahl designed an experiment to determine the method of DNA replication. Three models of replication were considered likely.

1. **Conservative replication** would produce an entirely new DNA strand during replication.



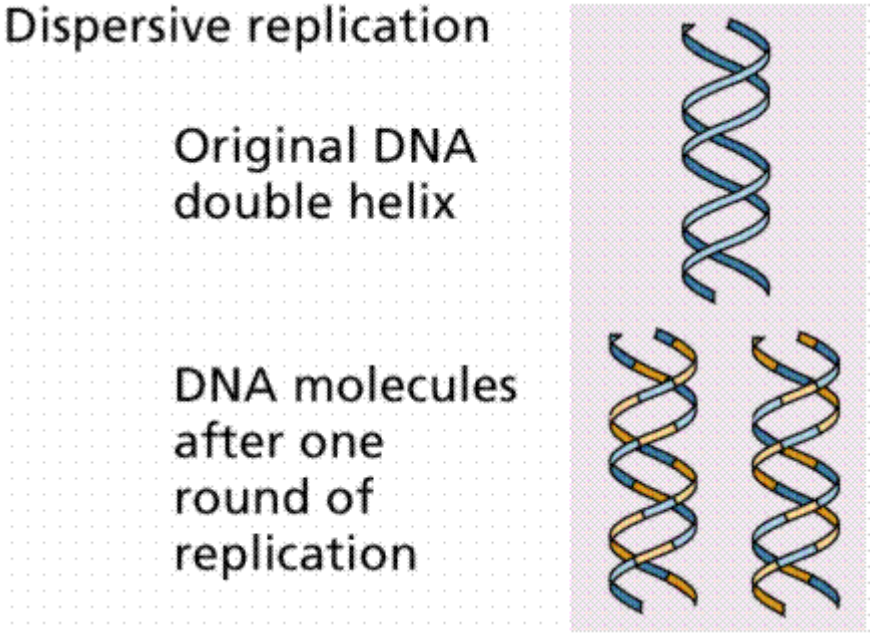
Conservative model of DNA replication.

2. **Semiconservative replication** would produce two DNA molecules, each of which was composed of one-half of the parental DNA along with an entirely new complementary strand. In other words the new DNA would consist of one new and one old strand of DNA. The existing strands would serve as complementary templates for the new strand.



The semiconservative model of DNA structure.

3. **Dispersive replication** involved the breaking of the parental strands during replication, and somehow, a reassembly of molecules that were a mix of old and new fragments on each strand of DNA.



DNA is the genetic molecule of all life. DNA (along with associated proteins) is found in chromosomes. DNA ultimately controls all cell activities.

We also know that the DNA in each of our cells is identical; DNA molecules duplicate* prior to cell division ensuring that the new cells formed are genetically identical to the original cell

A gene locus of DNA stores the information that specifies the sequence of amino acids that form a specific polypeptide. The genes, or more precisely, alleles we inherit from our parents determine the polypeptides we synthesize in our cells, which determine the structure and functioning of our cells and tissues.

DNA was a polymer of nucleotides. Each nucleotide contained:

1-Phosphate (P)

2-The 5-carbon sugar, deoxyribose

3- One of four different nitrogen-containing bases

DNA Replication and Repair:

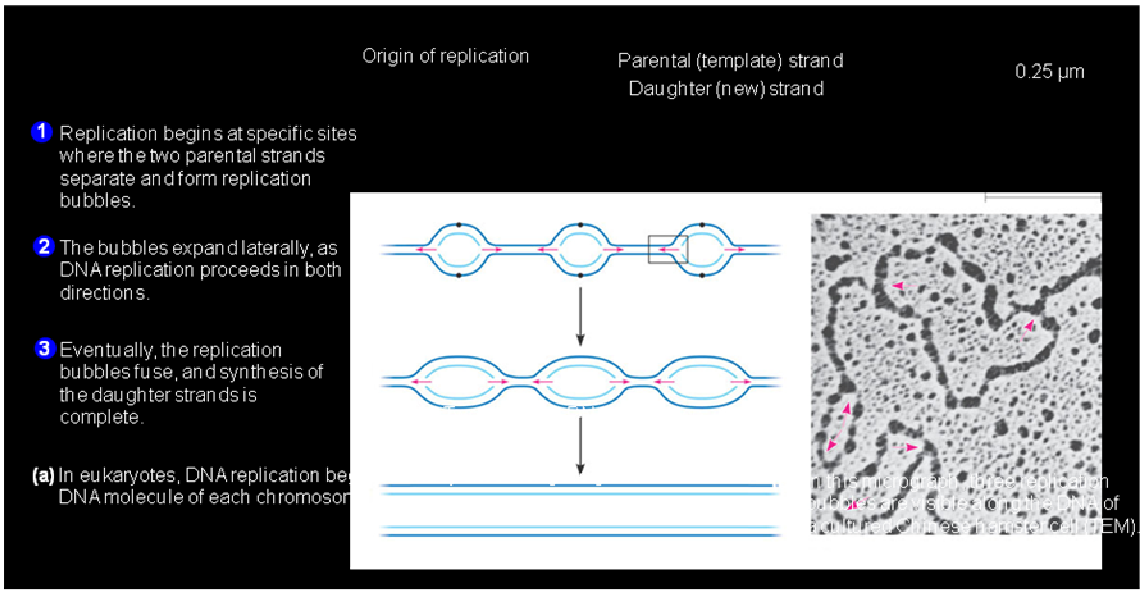
Difference Between the Eukaryotic and Prokaryotic DNA Replication Steps

As mentioned earlier, the eukaryotic DNA replication and the prokaryotic DNA replication steps have a different mechanism. The eukaryotic DNA replication steps are more complex than the prokaryotic DNA replication. Eukaryotic DNA is found in all complex organisms which includes plants and animals, whereas the prokaryotic DNA is present in 'simple' organisms like bacteria and cyanobacteria. The eukaryotic DNA is always present in combination with histone proteins and the prokaryotic DNA is a simple duplex that does not contain histones (that is, basic proteins). In the prokaryotic DNA replication steps, the DNA is replicated during the interval between the cell divisions. The eukaryotic DNA replication steps are highly regulated and the process takes place during the 'S' phase of the cell cycle, that precedes mitosis or meiosis I.

Replication process

Origins

For a cell to divide, it must first replicate its DNA. This process is initiated at particular points in the DNA, known as "origins", which are targeted by proteins that separate the two strands and initiate DNA synthesis. Origins contain DNA sequences recognized by replication initiator proteins (e.g., dnaA in *E. coli*' and the Origin Recognition Complex in yeast). These initiator proteins recruit other proteins to separate the two strands and initiate replication forks.

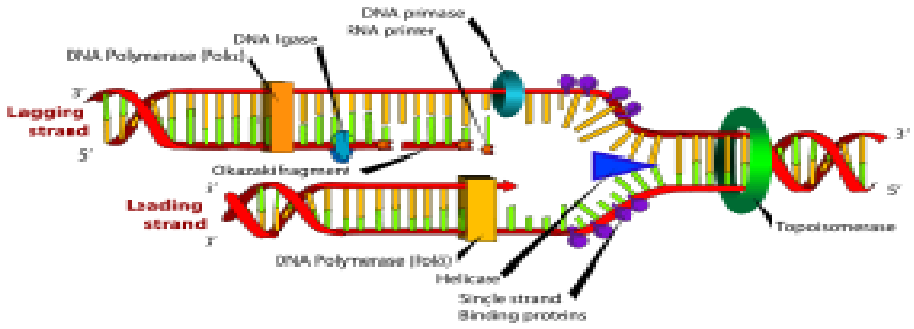


Initiator proteins recruit other proteins to separate the DNA strands at the origin, forming a bubble. Origins tend to be "AT-rich" (rich in adenine and thymine bases) to assist this process, because A-T base pairs have two hydrogen bonds (rather than the three formed in a C-G pair)—in general, strands rich in these nucleotides are easier to separate because a greater number of hydrogen bonds requires more energy to break them. Once strands are separated, RNA primers are created on the template strands. To be more specific, the leading strand receives one RNA primer per active origin of replication while the lagging strand receives several; these several fragments of RNA primers found on the lagging strand of DNA are called Okazaki fragments, named after their discoverer. DNA Polymerase extends the leading strand in one continuous motion and the lagging strand in a discontinuous motion (due to the Okazaki fragments). RNase removes the RNA fragments used to initiate replication by DNA Polymerase, and another DNA Polymerase enters to fill the gaps. When this is complete, a single nick on the leading strand and several nicks on the lagging strand can be found. Ligase

works to fill these nicks in, thus completing the newly replicated DNA molecule.

As DNA synthesis continues, the original DNA strands continue to unwind on each side of the bubble, forming a replication fork with two prongs. In bacteria, which have a single origin of replication on their circular chromosome, this process eventually creates a "theta structure" (resembling the Greek letter theta: θ). In contrast, eukaryotes have longer linear chromosomes and initiate replication at multiple origins within these.

Replication fork



Many enzymes are involved in the DNA replication fork.

The replication fork is a structure that forms within the nucleus during DNA replication. It is created by helicases, which break the hydrogen bonds holding the two DNA strands together. The resulting structure has two branching "prongs", each one made up of a single strand of DNA. These two strands serve as the template for the leading and lagging strands, which will be created as DNA polymerase matches complementary nucleotides to the templates; The templates may be properly referred to as the leading strand template and the lagging strand template.

Leading strand

The leading strand is the template strand of the DNA double helix so that the replication fork moves along it in the 3' to 5' direction. This allows the newly synthesized strand complementary to the original strand to be synthesized 5' to 3' in the same direction as the movement of the replication fork.

On the leading strand, a polymerase "reads" the DNA and adds nucleotides to it continuously. This polymerase is DNA polymerase III (DNA Pol III) in prokaryotes and presumably Pol ε in yeasts. In human cells the leading and lagging strands are synthesized by Pol α and Pol δ within the nucleus and Pol γ in the mitochondria. Pol ε can substitute for Pol δ in special circumstances.

Lagging strand

The lagging strand is the strand of the template DNA double helix that is oriented so that the replication fork moves along it in a 5' to 3' manner. Because of its orientation, opposite to the working orientation of DNA polymerase III, which moves on a template in a 3' to 5' manner, replication of the lagging strand is more complicated than that of the leading strand.

On the lagging strand, primase "reads" the DNA and adds RNA to it in short, separated segments. In eukaryotes, primase is intrinsic to Pol α. DNA polymerase III or Pol δ lengthens the primed segments, forming Okazaki fragments. Primer removal in eukaryotes is also performed by Pol δ. In prokaryotes, DNA polymerase I "reads" the fragments, removes the RNA using its flap endonuclease domain (RNA primers are removed

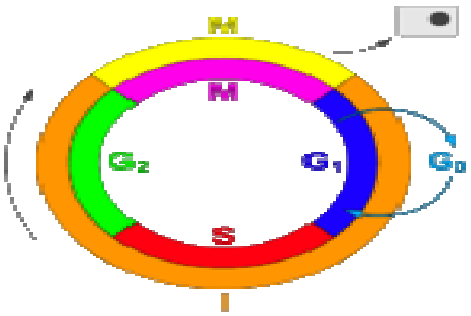
by 5'-3' exonuclease activity of polymerase I [weaver, 2005], and replaces the RNA nucleotides with DNA nucleotides (this is necessary because RNA and DNA use slightly different kinds of nucleotides). DNA ligase joins the fragments together.

As helicase unwinds DNA at the replication fork, the DNA ahead is forced to rotate. This process results in a build-up of twists in the DNA ahead. This build-up would form a resistance that would eventually halt the progress of the replication fork. DNA Gyrase is an enzyme that temporarily breaks the strands of DNA, relieving the tension caused by unwinding the two strands of the DNA helix; DNA Gyrase achieves this by adding negative supercoils to the DNA helix.

Bare single-stranded DNA tends to fold back on itself and form secondary structures; these structures can interfere with the movement of DNA polymerase. To prevent this, single-strand binding proteins bind to the DNA until a second strand is synthesized, preventing secondary structure formation.

Clamp proteins form a sliding clamp around DNA, helping the DNA polymerase maintain contact with its template, thereby assisting with processivity. The inner face of the clamp enables DNA to be threaded through it. Once the polymerase reaches the end of the template or detects double-stranded DNA, the sliding clamp undergoes a conformational change that releases the DNA polymerase. Clamp-loading proteins are used to initially load the clamp, recognizing the junction between template and RNA primers.

Regulation



The cell cycle of eukaryotic cells.

Eukaryotes

Within eukaryotes, DNA replication is controlled within the context of the cell cycle. As the cell grows and divides, it progresses through stages in the cell cycle; DNA replication occurs during the S phase (synthesis phase). The progress of the eukaryotic cell through the cycle is controlled by cell cycle checkpoints. Progression through checkpoints is controlled through complex interactions between various proteins, including cyclins and cyclin-dependent kinases.

The G1/S checkpoint (or restriction checkpoint) regulates whether eukaryotic cells enter the process of DNA replication and subsequent division. Cells that do not proceed through this checkpoint are remain in the G0 stage and do not replicate their DNA.

Replication of chloroplast and mitochondrial genomes occurs independent of the cell cycle, through the process of D-loop replication.

Bacteria

Most bacteria do not go through a well-defined cell cycle but instead continuously copy their DNA; during rapid growth, this can result in the concurrent occurrences of multiple rounds of replication. In *E. coli*, the best-characterized bacteria, DNA replication is regulated through several mechanisms, including: the hemimethylation and sequestering of the origin sequence, the ratio of ATP to ADP, and the levels of protein DnaA. All these control the process of initiator proteins binding to the origin sequences.

Because *E. coli* methylates GATC DNA sequences, DNA synthesis results in hemimethylated sequences. This hemimethylated DNA is recognized by the protein SeqA, which binds and sequesters the origin sequence; in addition, dnaA (required for initiation of replication) binds less well to hemimethylated DNA. As a result, newly replicated origins are prevented from immediately initiating another round of DNA replication.

ATP builds up when the cell is in a rich medium, triggering DNA replication once the cell has reached a specific size. ATP competes with ADP to bind to DnaA, and the DnaA-ATP complex is able to initiate replication. A certain number of DnaA proteins are also required for DNA replication — each time the origin is copied, the number of binding sites for DnaA doubles, requiring the synthesis of more DnaA to enable another initiation of replication.

Termination

Eukaryotes initiate DNA replication at multiple points in the chromosome, so replication forks meet and terminate at many points in the chromosome; these are not known to be regulated in any particular way.

Because eukaryotes have linear chromosomes, DNA replication is unable to reach the very end of the chromosomes, but ends at the telomere region of repetitive DNA close to the end. This shortens the telomere of the daughter DNA strand. This is a normal process in somatic cells. As a result, cells can only divide a certain number of times before the DNA loss prevents further division. (This is known as the Hayflick limit.) Within the germ cell line, which passes DNA to the next generation, telomerase extends the repetitive sequences of the telomere region to prevent degradation. Telomerase can become mistakenly active in somatic cells, sometimes leading to cancer formation.

Because bacteria have circular chromosomes, termination of replication occurs when the two replication forks meet each other on the opposite end of the parental chromosome. *E coli* regulate this process through the use of termination sequences that, when bound by the Tus protein, enable only one direction of replication fork to pass through. As a result, the replication forks are constrained to always meet within the termination region of the chromosome.

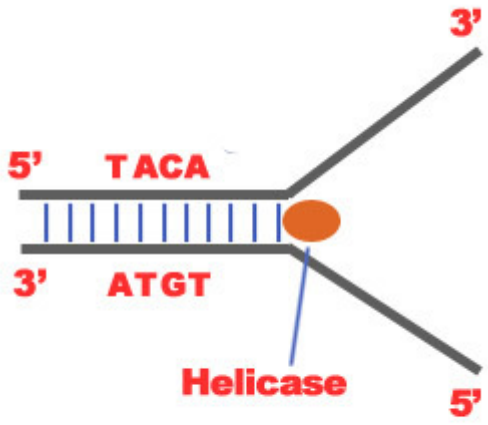
Table 16.1 Bacterial DNA replication proteins and their functions

Protein	Function for Leading and Lagging Strands	
Helicase	Unwinds parental double helix at replication forks	
Single-strand binding protein	Binds to and stabilizes single-stranded DNA until it can be used as a template	
Topoisomerase	Corrects "overwinding" ahead of replication forks by breaking, swiveling, and rejoining DNA strands	
	Function for Leading Strand	**Function for Lagging Strand**
Primase	Synthesizes a single RNA primer at the 5′ end of the leading strand	Synthesizes an RNA primer at the 5′ end of each Okazaki fragment
DNA pol III	Continuously synthesizes the leading strand, adding on to the primer	Elongates each Okazaki fragment, adding on to its primer
DNA pol I	Removes primer from the 5′ end of leading strand and replaces it with DNA, adding on to the adjacent 3′ end	Removes the primer from the 5′ end of each fragment and replaces it with DNA, adding on to the 3′ end of the adjacent fragment
DNA Ligase	Joins the 3′ end of the DNA that replaces the primer to the rest of the leading strand	Joins the Okazaki fragments

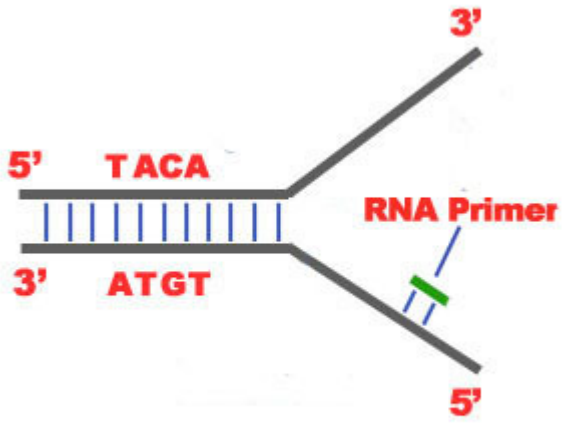
Steps of DNA Replication

The next we have to do is to shed light into the mystery of the steps of DNA Replicationof the Eykaryotes.

1)The first major step for the DNA Replication to take place is the breaking of hydrogen bonds between bases of the two antiparallel strands. The unwounding of the two strands is the starting point. The splitting happens in places of the chains which are rich in A-T. That is because there are only two bonds between Adenine and Thymine (there are three hydrogen bonds between Cytosine and Guanine). Helicase is the enzyme that splits the two strands. The initiation point where the splitting starts is called "origin of replication".The structure that is created is known as "Replication Fork".

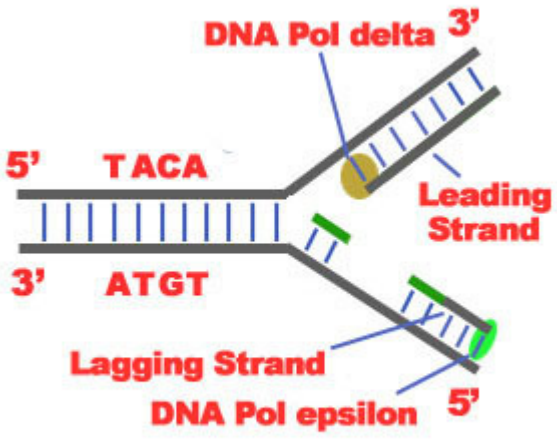


2) One of the most important steps of DNA Replication is the binding of RNA Primase in the the initiation point of the 3'-5' parent chain. RNA Primase can attract RNA nucleotides which bind to the DNA nucleotides of the 3'-5' strand due to the hydrogen bonds between the bases. RNA nucleotides are the primers (starters) for the binding of DNA nucleotides.



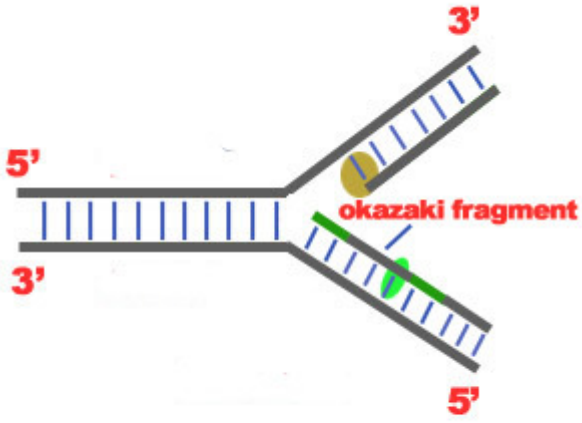
3) The **elongation** process is different for the 5'-3' and 3'-5' template. a)**5'-3' Template**: The 3'-5' proceeding daughter strand -that uses a **5'-3' template**- is called **leading strand** because **DNA Polymerase** can "read" the template and continuously adds nucleotides (complementary

to the nucleotides of the template, for example Adenine opposite to Thymine etc).



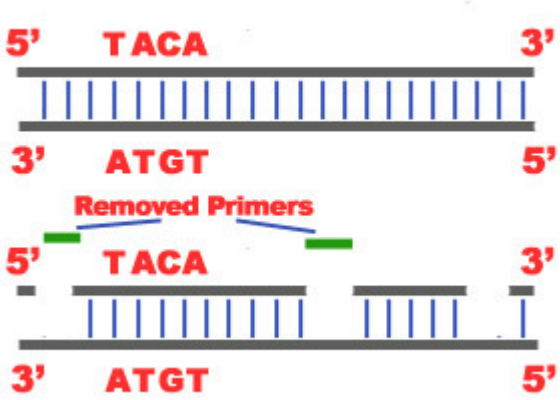
b) **3'-5'Template**: The **3'-5' template** cannot be "read" by DNA Polymerase. The replication of this template is complicated and the new strand is called **lagging strand**. In the lagging strand the RNA Primase adds more RNA Primers. **DNA polymerase** reads the template and lengthens the bursts. The gap between two RNA primers is called "**Okazaki Fragments**".

The RNA Primers are necessary for DNA Polymerase to bind Nucleotides to the 3' end of them. The daughter strand is elongated with the binding of more DNA nucleotides.



4) In the lagging strand the **DNA Pol I** -**exonuclease**- reads the fragments and removes the RNA Primers. The gaps are closed with the action of DNA Polymerase (adds complementary nucleotides to the gaps) and DNA Ligase (adds phosphate in the remaining gaps of the phosphate - sugar backbone).

Each new double helix is consisted of one old and one new chain. This is what we call **semiconservative replication**.

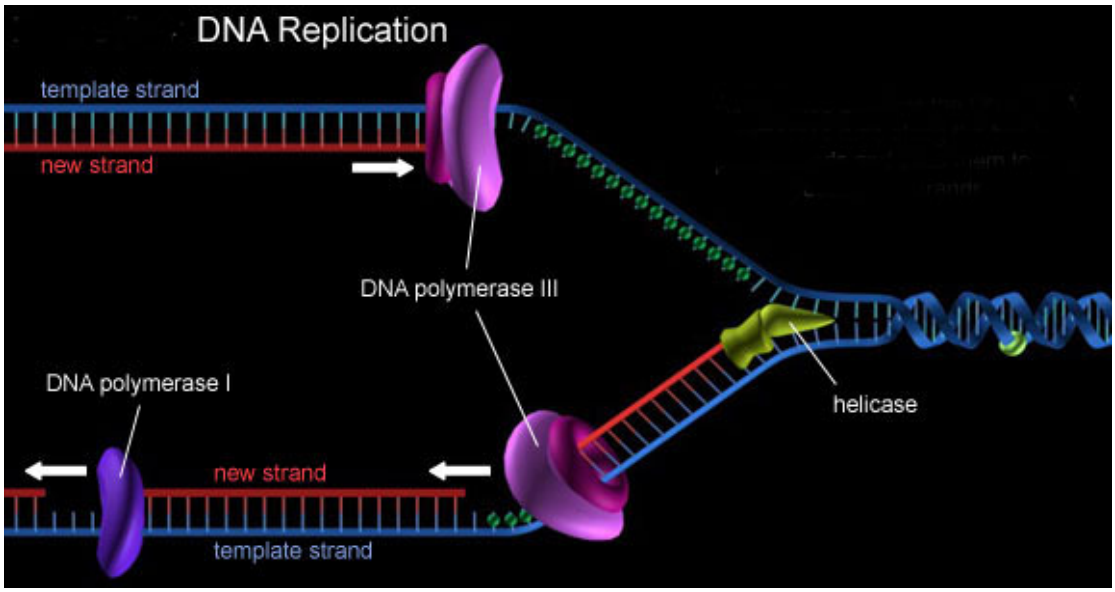


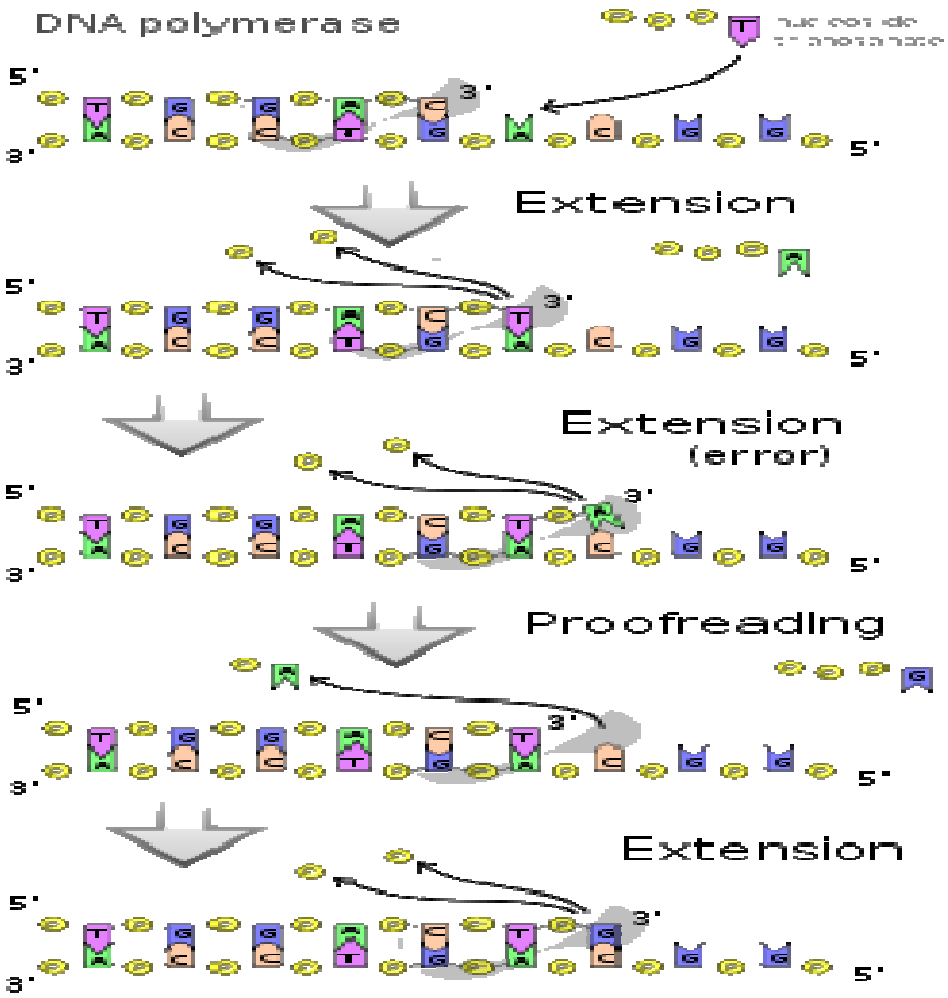
5) The last **step of DNA Replication** is the **Termination**. This process

happens when the DNA Polymerase reaches to an end of the strands. We can easily understand that in the last section of the lagging strand, when the RNA primer is removed, it is not possible for the DNA Polymerase to seal the gap (because there is no primer). So, the end of the parental strand where the last primer binds isn't replicated. These ends of linear (chromosomal) DNA consists of noncoding DNA that contains repeat sequences and are called **telomeres**. As a result, a part of the telomere is removed in every cycle of DNA Replication.

6) The DNA Replication is not completed before a **mechanism of repair** fixes possible errors caused during the replication. Enzymes like **nucleases** remove the wrong nucleotides and the DNA Polymerase fills the gaps.

Similar processes also happen during the **steps of DNA Replication** of prokaryotes though there are some differences.





II- RNA TRANSCRIPTION

Introduction:

Definition:

Transcription is the first step in gene expression. RNA strands are constructed by cellular enzymes that read the genetic code of a single DNA strand, and synthesize a complimentary RNA strand. Usually, particularly in eukaryotic cells, such as humans, plants and animals, transcription is initiated by a complex of enzymes including RNA polymerase, which bind to the DNA strand at a specific location known as the promoter or enhancer region. Initiation of transcription is often a complex process that requires the presence of a number of proteins called transcription factors. These proteins bind to the DNA strand and interact with one another, either turning transcription on or off, depending on the combination and types of transcription factors present.
Transcription is part of the process by which proteins are made. The next step in the process is translation of the RNA code to an amino acid chain, resulting in a protein.

Transcription is the mechanism by which a template strand of DNA is utilized by specific RNA polymerases to generate one of the three different classifications of RNA. These 3 RNA classes are:

1. **Messenger RNAs (mRNAs):** This class of RNAs are the genetic coding templates used by the translational machinery to determine the order of amino acids incorporated into an elongating polypeptide in the process of translation.

2. **Transfer RNAs (tRNAs):** This class of small RNAs form covalent attachments to individual amino acids and recognize the

encoded sequences of the mRNAs to allow correct insertion of amino acids into the elongating polypeptide chain.

3. Ribosomal RNAs (rRNAs): This class of RNAs are assembled, together with numerous ribosomal proteins, to form the ribosomes. Ribosomes engage the mRNAs and form a catalytic domain into which the tRNAs enter with their attached amino acids. The proteins of the ribosomes catalyze all of the functions of polypeptide synthesis.

All RNA polymerases are dependent upon a DNA template in order to synthesize RNA. The resultant RNA is, therefore, complimentary to the template strand of the DNA duplex and identical to the non-template strand. The non-template strand is called the coding strand because its' sequences are identical to those of the mRNA. However, in RNA, U is substituted for T.

Classes of RNA Polymerases

In prokaryotic cells, all 3 RNA classes are synthesized by a single polymerase. In eukaryotic cells there are 3 distinct classes of RNA polymerase, RNA polymerase (pol) I, II and III. Each polymerase is responsible for the synthesis of a different class of RNA. The capacity of the various polymerases to synthesize different RNAs was shown with the toxin α-amanitin. At low concentrations of α-amanitin synthesis of mRNAs are affected but not rRNAs nor tRNAs. At high concentrations, both mRNAs and tRNAs are affected. These observations have allowed the identification of which polymerase synthesizes which class of RNAs. RNA pol I is responsible for rRNA synthesis (excluding the 5*S* rRNA).There are 4 major rRNAs in eukaryotic cells designated by their

sedimentation size. The 28*S*, 5*S* 5.8*S* RNAs are associated with the large ribosomal subunit and the 18*S* rRNA is associated with the small ribosomal subunit. RNA pol II synthesizes the mRNAs and some of the small nuclear RNAs (snRNAs) involved in RNA splicing. RNA pol III synthesizes the tRNAs, the 5*S* rRNA and some snRNAs.

Mechanism of RNA Polymerases

Synthesis of RNA exhibits several features that are synonymous with DNA replication. RNA synthesis requires accurate and efficient initiation, elongation proceeds in the 5'—>3' direction (i.e. the polymerase moves along the template strand of DNA in the 3'—>5' direction), and RNA synthesis requires distinct and accurate termination. Transcription exhibits several features that are distinct from replication.

1. Transcription initiates, both in prokaryotes and eukaryotes, from many more sites than replication.

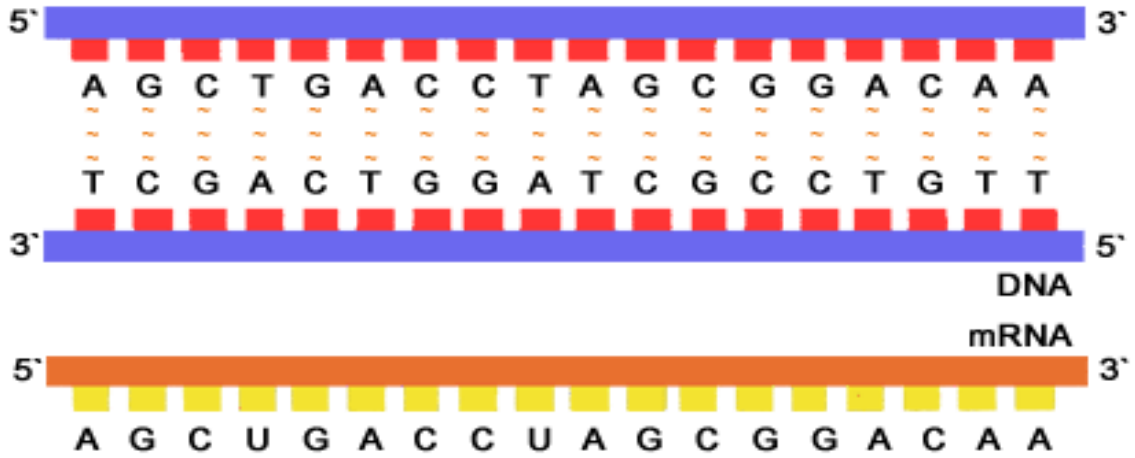
2. There are many more molecules of RNA polymerase per cell than DNA polymerase.

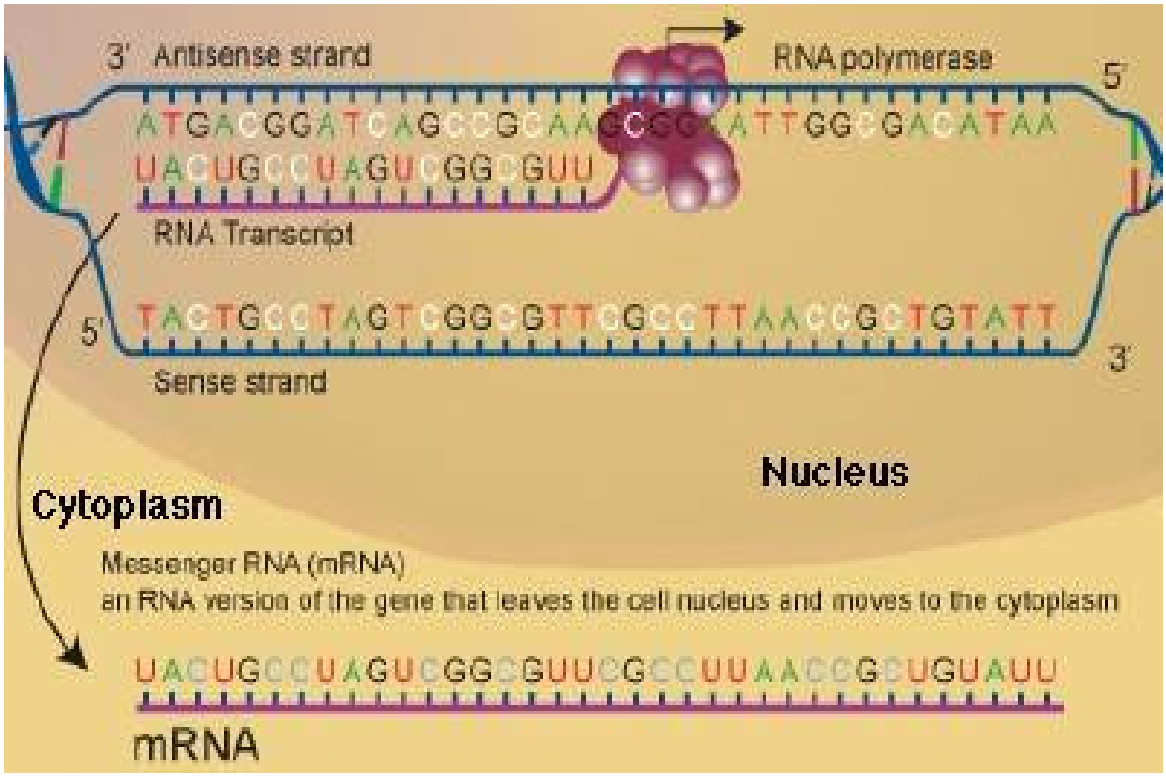
3. RNA polymerase proceeds at a rate much slower than DNA polymerase (approximately 50–100 bases/sec for RNA versus near 1000 bases/sec for DNA).

4. Finally the fidelity of RNA polymerization is much lower than DNA. This is allowable since the aberrant RNA molecules can simply be turned over and new correct molecules made.

Steps of DNA Transcription

DNA transcription is a process that involves the transcribing of genetic information from DNA to RNA. The transcribed DNA message is used to produce proteins. DNA is housed within the nucleus of our cells. It controls cellular activity by coding for the production of enzymes and proteins. The information in DNA is not directly converted into proteins, but must first be copied into RNA. This ensures that the information contained within the DNA does not become tainted.





Transcription can be explained easily in 4 or 5 simple steps, each moving like a wave along the DNA.

1. RNA polymerase unwinds/"unzips" the DNA by breaking the hydrogen bonds between complimentary nucleotides.
2. RNA nucleotides are paired with complementary DNA bases.
3. RNA sugar-phosphate backbone forms with assistance from RNA polymerase.
4. Hydrogen bonds of the untwisted RNA+DNA helix break, freeing the newly synthesized RNA strand.
5. If the cell has a nucleus, the RNA is further processed and then moves through the small nuclear pores to the cytoplasm.

DNA consists of four nucleotide bases [adenine (A), guanine (G), cytosine (C) and thymine (T)] that are paired together (A-T and C-G) to give DNA its double helical shape.

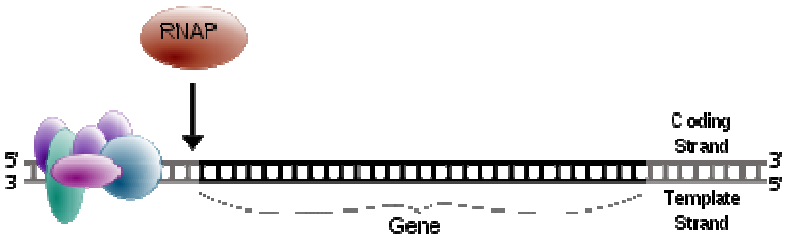
There are three main steps to the process of DNA transcription.

- Initiation

  -RNA Polymerase Binds to DNA
  -DNA is transcribed by an enzyme called RNA polymerase. Specific nucleotide sequences tell RNA polymerase where to begin and where to end. RNA polymerase attaches to the DNA at a specific area called the promoter region.

  -transcription begins by the formation of **preinitiation complex(PIC)** which contains the RNA polymerase and certain proteins called six general transcription factors (GTP)

  -formation of this complex requires the recognition of the TATA box(core of promoter which has a sequence of TATAA).

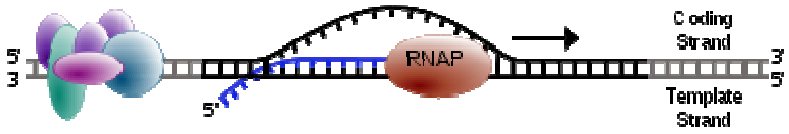


- Elongation

  Transcription factors unwind the DNA strand and allow RNA polymerase to transcribe only a single strand of DNA into a single stranded RNA polymer called messenger RNA (mRNA). The strand that serves as the template is called the antisense strand. The strand that is not transcribed is called the sense strand.

  Like DNA, RNA is composed of nucleotide bases. RNA however,

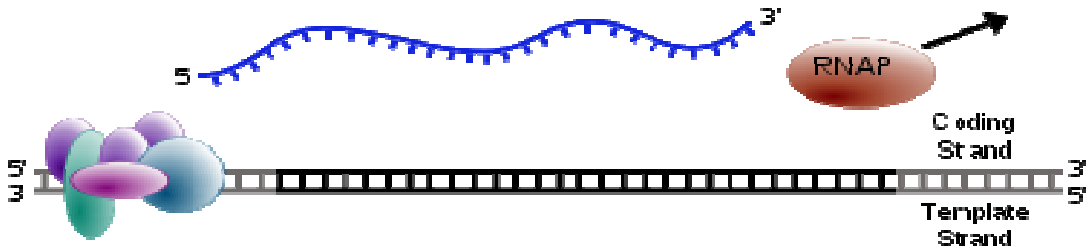
contains the nucleotides adenine, guanine, cytosine and uricil (U). When RNA polymerase transcribes the DNA, guanine pairs with cytosine and adenine pairs with uricil.



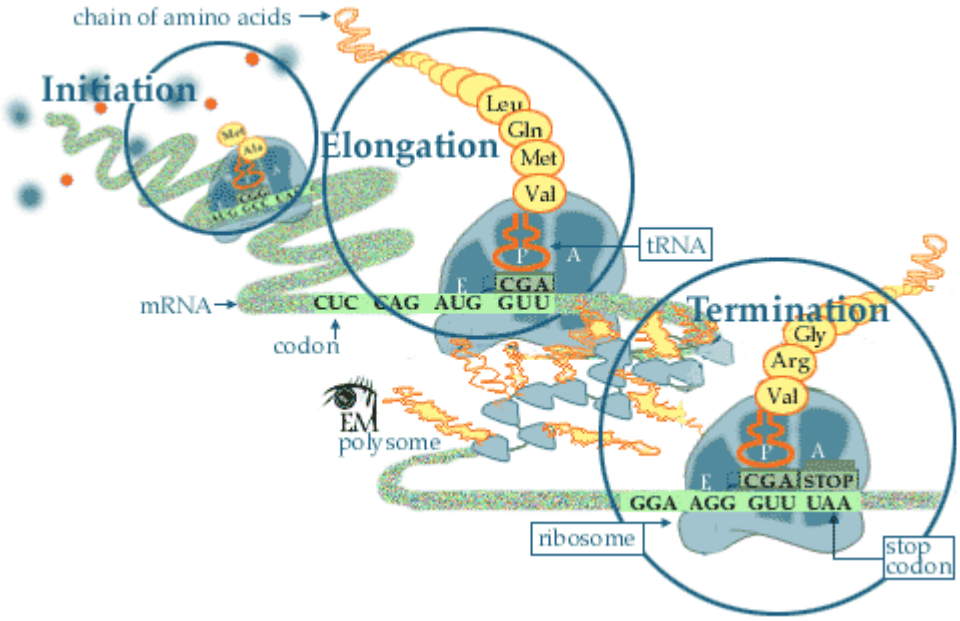
- Termination

  RNA polymerase moves along the DNA until it reaches a terminator sequence. At that point, RNA polymerase releases the mRNA polymer and detaches from the DNA.

Since proteins are constructed in the cytoplasm of the cell by a process called translation, mRNA must cross the nuclear membrane to reach the cytoplasm. Once in the cytoplasm, mRNA along with ribosomes and another RNA molecule called transfer RNA, work together to produce proteins. Proteins can be manufactured in large quantities because a single DNA sequence can be transcribed by many RNA polymerase molecules at once.



III- Translation Process



Translation is the actual synthesis of a protein under the direction of mRNA. During this process the nucleotide sequence of an mRNA (messenger RNA) is translated into the amino acid sequence of a protein. The protein synthesis requires a technical machinery of high complexity. As compared to information transfer between nucleic acid molecules, where direct copying occurs on the basis of base complementarity, the translation process involves a greater number of chemical reactions and the participation of additional nucleic acid and protein components. One of these components, the ribosome, provides the basic machinery for the translation process. The major role of the ribosome is to catalyse coupling of amino acids into protein according to the sequence specified by the mRNA. The amino acids are brought to the ribosome by tRNA (transfer RNA) molecules. The nucleotide sequence of the mRNA is composed of four different nucleotides whereas a protein is built up from 20 amino acids. To allow

the four nucleotides to specify 20 different amino acids, the nucleotide sequence is interpreted in codons, groups of three nucleotides. These codons have their corresponding anticodon in the tRNA. Furthermore each anticodon is linked to one particular amino acid. Thus, each codon specifies one amino acid. This is referred to as the genetic code.

In addition to the main components of the translational system listed above, the translation process also involves a large number of protein factors that facilitate binding of mRNA and tRNA to the ribosome. Protein synthesis consumes a large part of the energy produced in the cell.

Translation may be divided into three distinct steps. The first, initiation, results in the formation of an initiation complex in which the ribosome is bound to the specific initiation (start) site on the mRNA while the initiator tRNA is annealed to the initiator codon and bound to the ribosome. The second stage, elongation, consists of joining amino acids to the growing polypeptide chain according to the sequence specified by the message. Incorporation of each amino acid occurs by the same mechanism. Thus, the same steps are repeated over and over again until the termination codon is reached in the message. The termination codon gives the signal for the third and last stage of protein synthesis, the termination, in which the ready-made protein is released from the ribosome.
Translation involves taking the message that's in the messenger RNA and in a sense decoding the message from the language of nucleic acids to the language of proteins or polypeptides. For translation to happen, the messenger RNA goes to the cytoplasm where it is attached to a cellular structure called a ribosome. Ribosomes are two part molecular assemblies consisting of various proteins plus a special kind of RNA

called ribosomal RNA. Ribosomal RNA is involved in catalyzing some of the chemical reactions of translation.

In addition to the ribosome, another kind of RNA called tRNA carries amino acids to the mRNA when it is attached to a particular part of the ribosome's small subunit, called a binding site. A critical feature of mRNA and how it is translated is the fact that each three nucleotides in the mRNA is called a codon and it is the codon that is translated. Thus the sequence of codons corresponds to the sequence of amino acids in the polypeptide. You will see that the tRNA molecules have a set of three nucleotide bases at one end that are complementary to a corresponding codon. The bases on the tRNA are called the anti codon. This is critical because the anti codons make the connection between the codons and the correct amino acids that go with each codon.

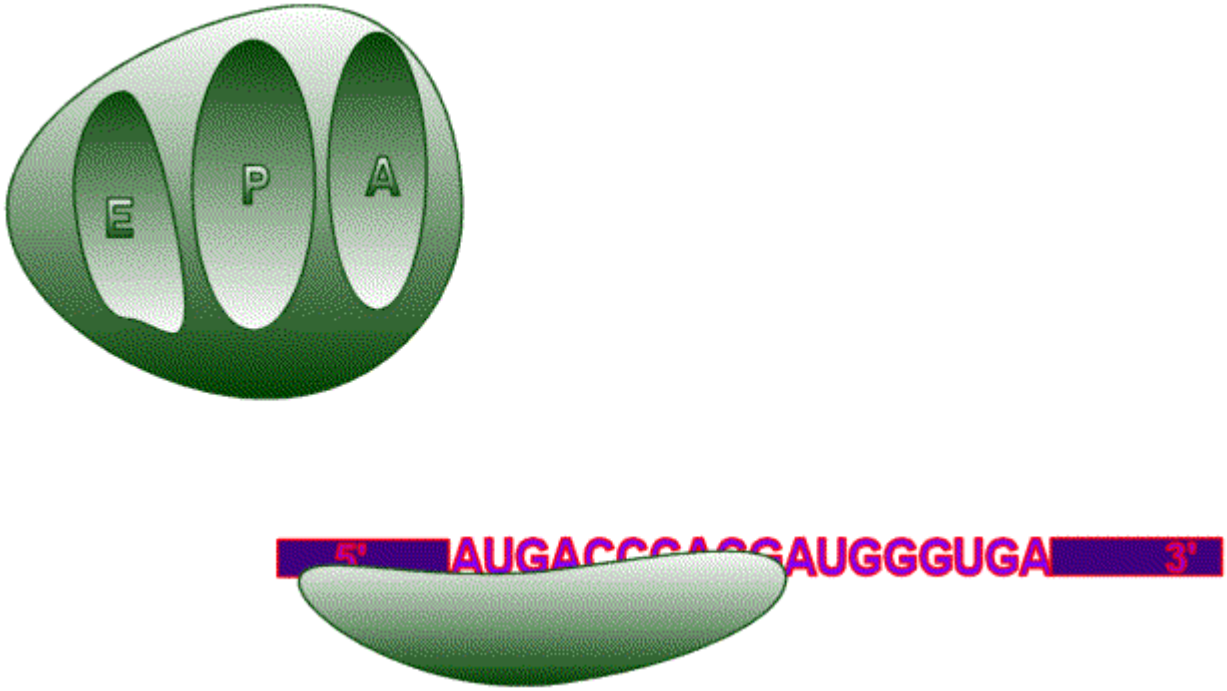
Scientists have cracked the code involved in translation and given a stretch of mRNA can tell what the corresponding sequence of amino acids is. Learning how to do this was one of the two or three big advances in the 20th century and has laid the foundation for many advances in biotechnology as well as more basic biology such as the study of evolution. This genetic code is virtually universal in that a particular codon will usually translate to the same amino acid regardless of of the organism doing the translation.

Sample mRNA transcript showing codon's nucleotide triplets. Read from left to right. This mRNA has 5 codons.

Initiation Step 1.

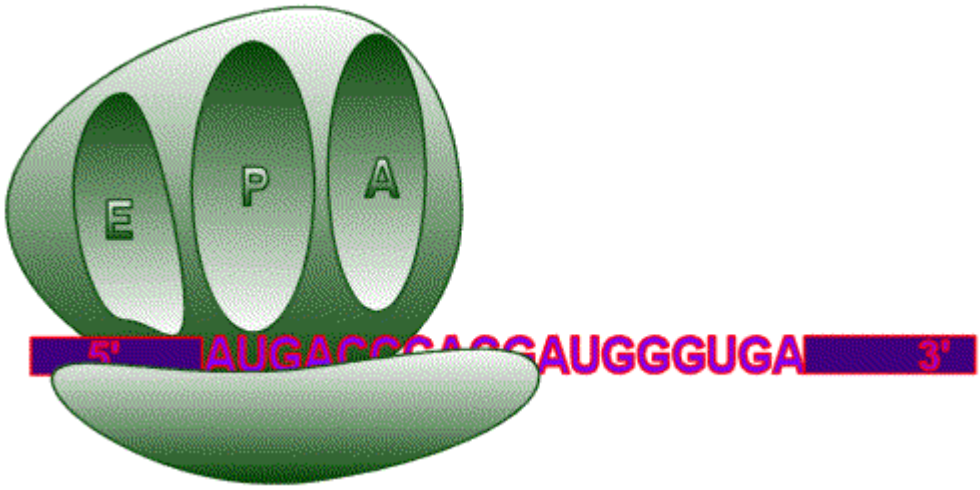
The mRNA joins to the small ribosomal unit at the 5' untranslated region. This binds to a special binding site on the small ribosomal subunit.

The large ribosomal subunit has 3 binding sites, E, P, and A.



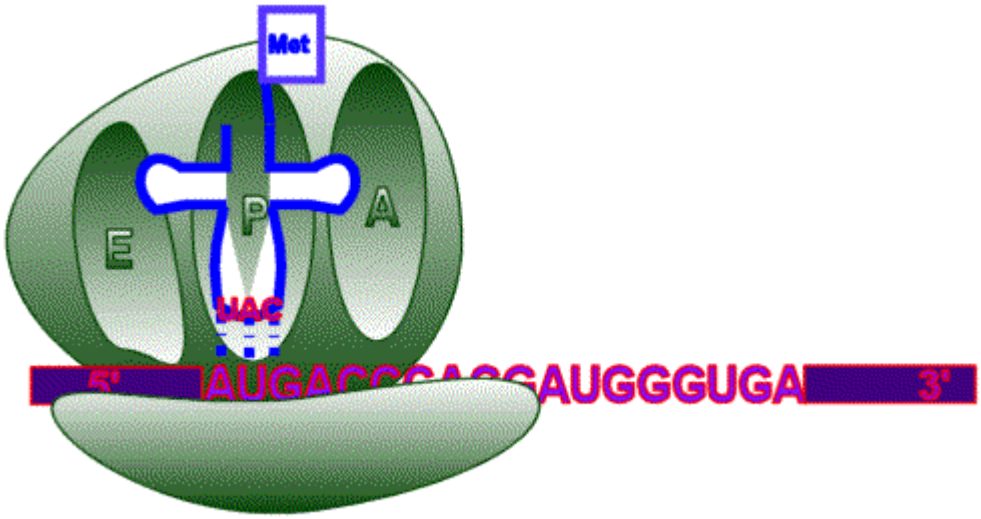
Initiation Step 2.

The large ribosomal subunit attaches to the small subunit such that the first codon is aligned at the P binding site.



Initiation step 3.

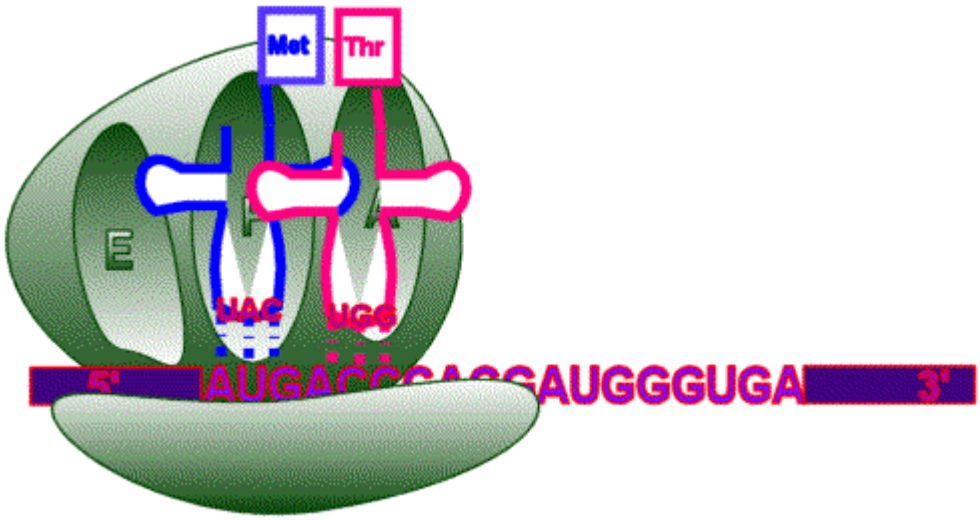
A tRNA carrrying the amino acid methionine attaches to the start codon (AUG) on the messaenger RNA. This inititates elongation.



Elongation step 1.

Attachment of first amino acid carrying tRNA to A binding site.

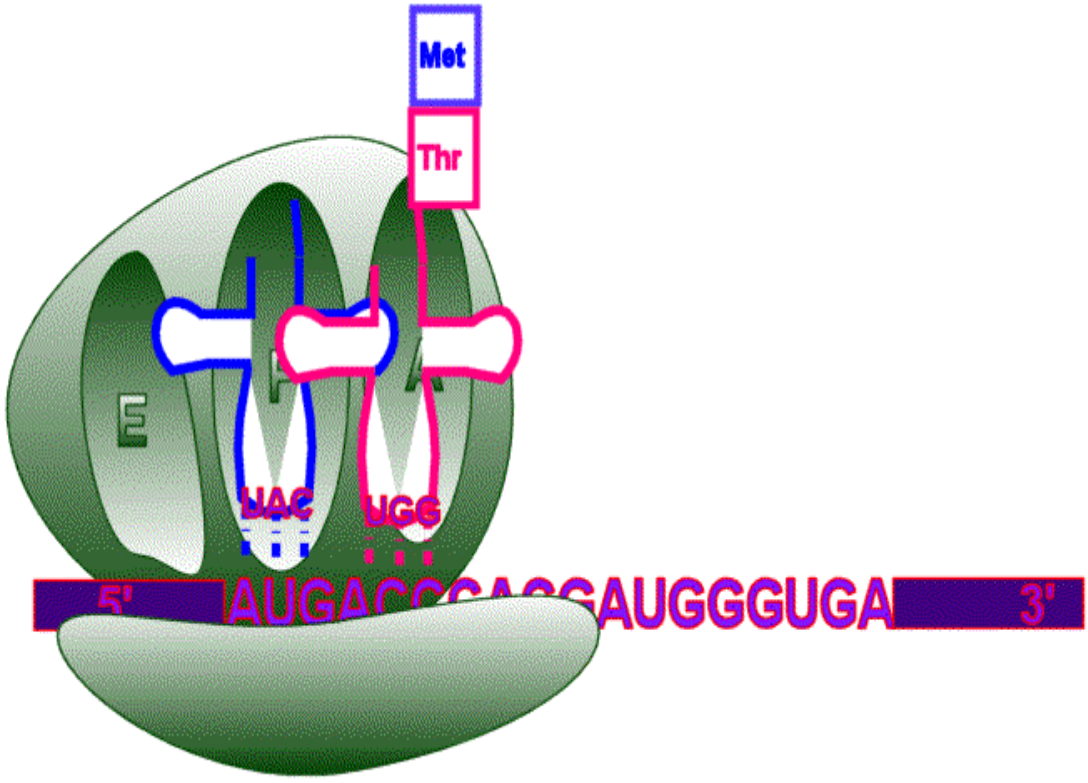
A tRNA and its amino acid attaches to the A binding site.



Elongation Step 2.

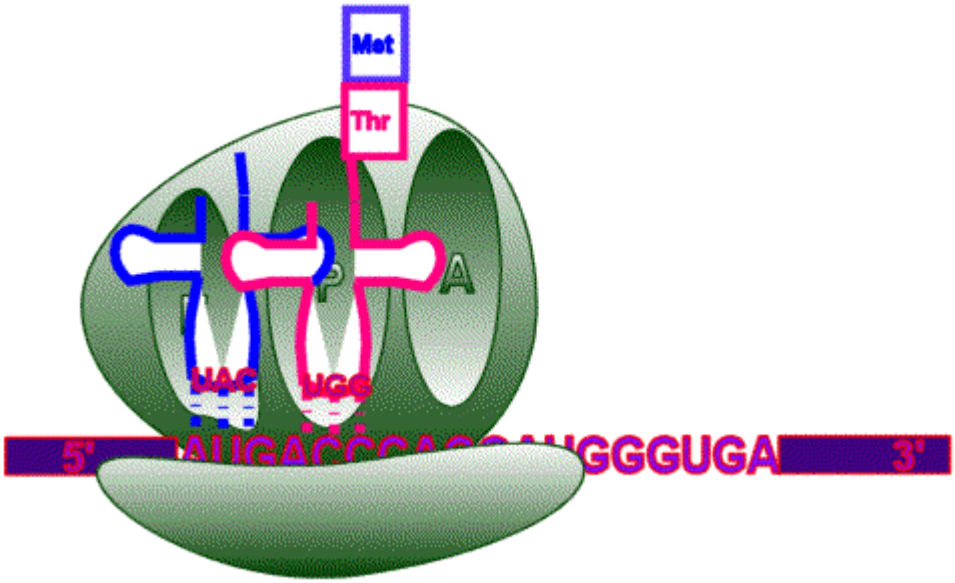
Peptide bond formation between the met and the amino acid carried at the A bindiung site. Our polypeptide chain is now:

Met –Thr



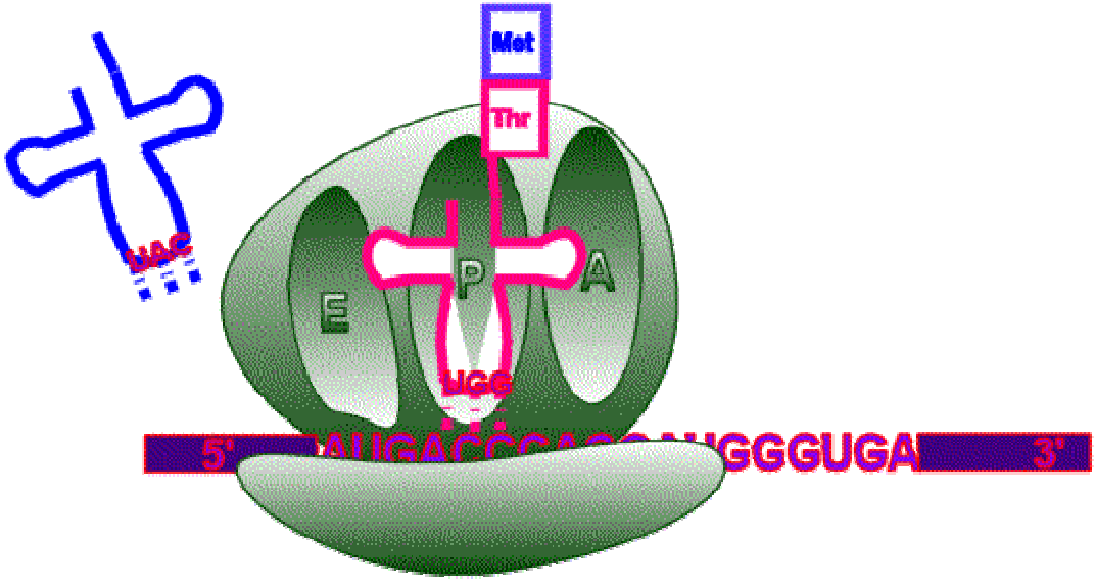
Elongation Step 3.

Ribosome moves in the 3' direction down the messenger RNA by three bases or one codon shifting the tRNA and polypeptide chain to the P Binding site. The A binding site is open and a vacant tRNA is in the E binding site.



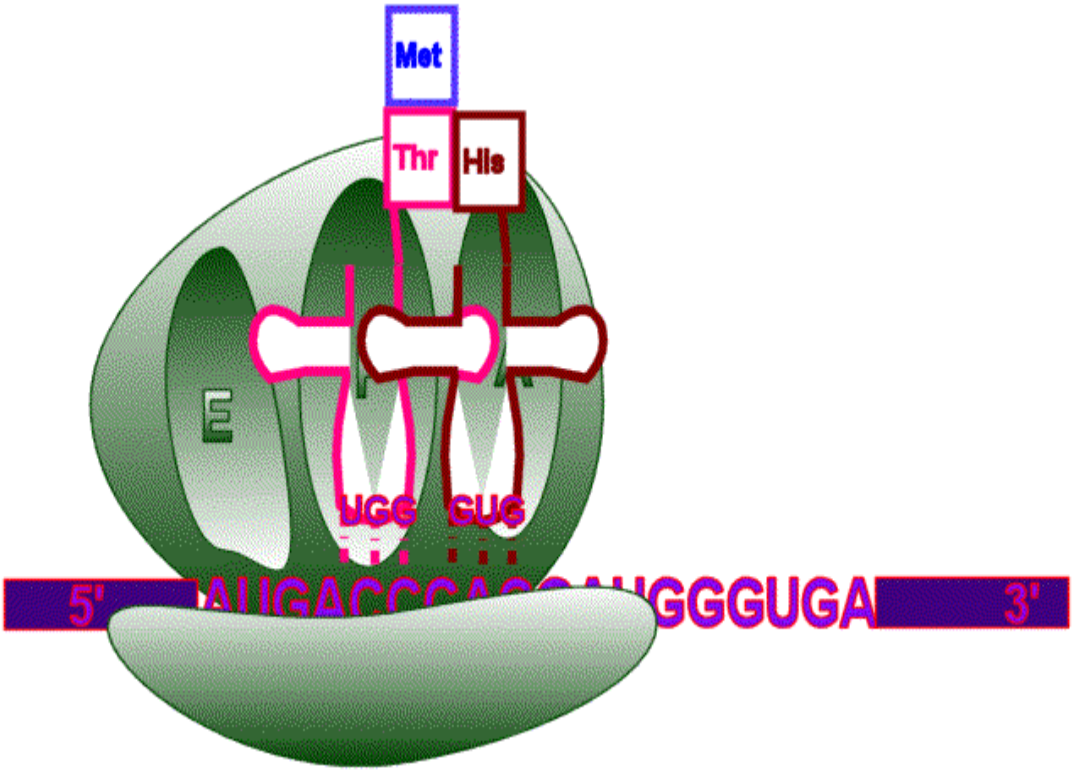
Elongation Step 4.

tRNA ejected from the E binding site.



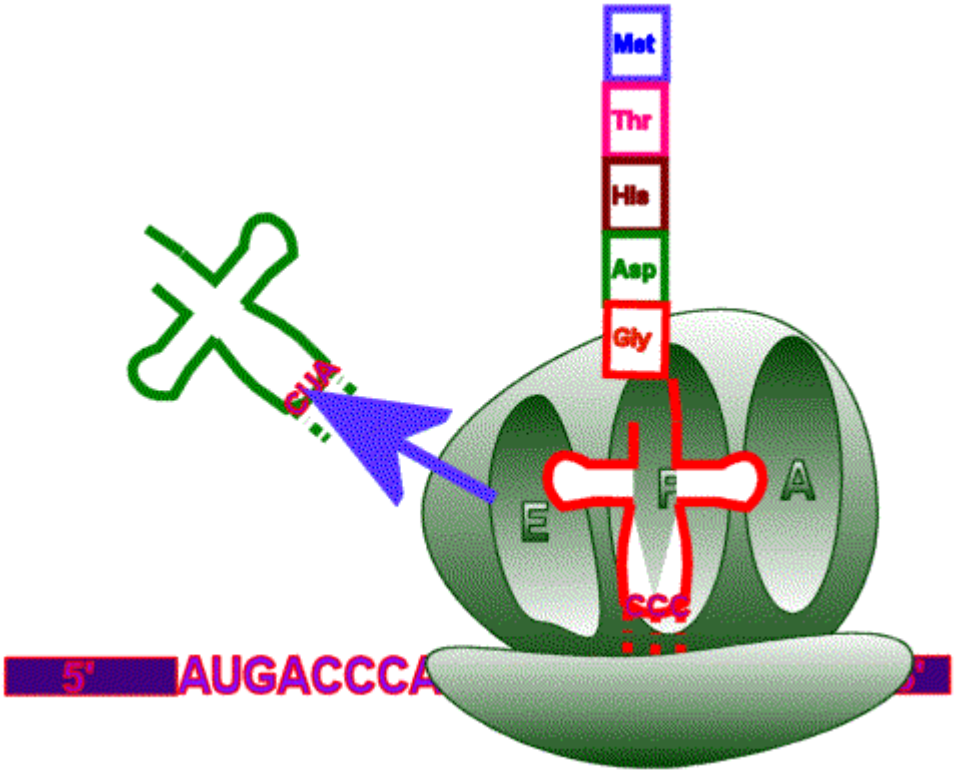
Repeat Elongation Step 1 - 4

until stop codon encountered.



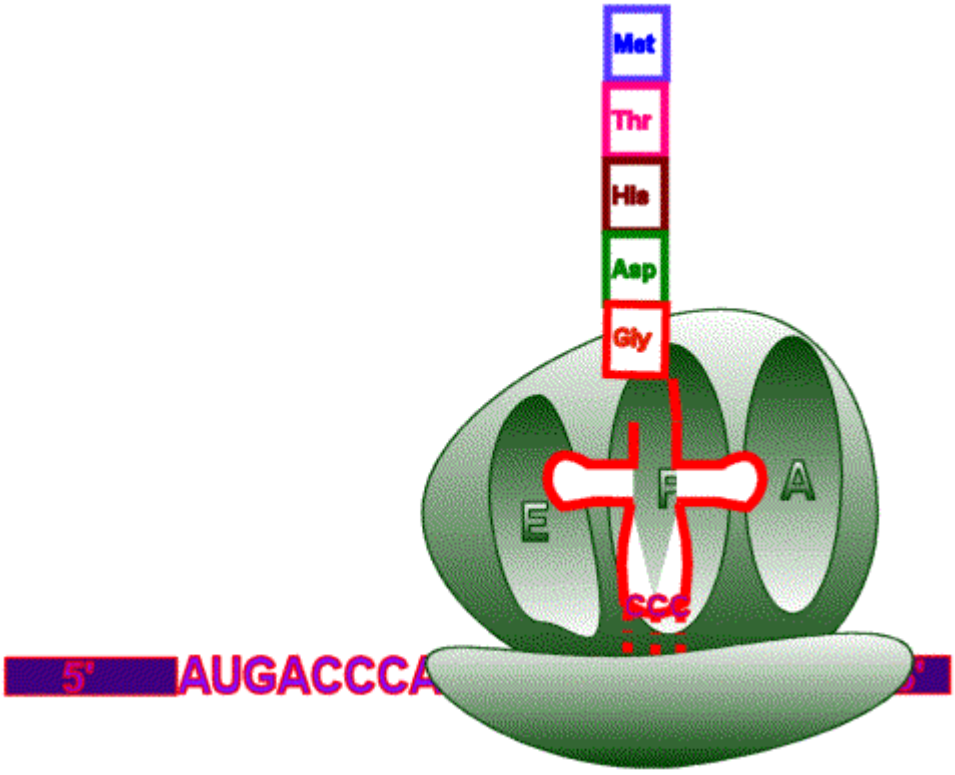
Elongation ends with:

- Old tRNA ejected from the E Binding site



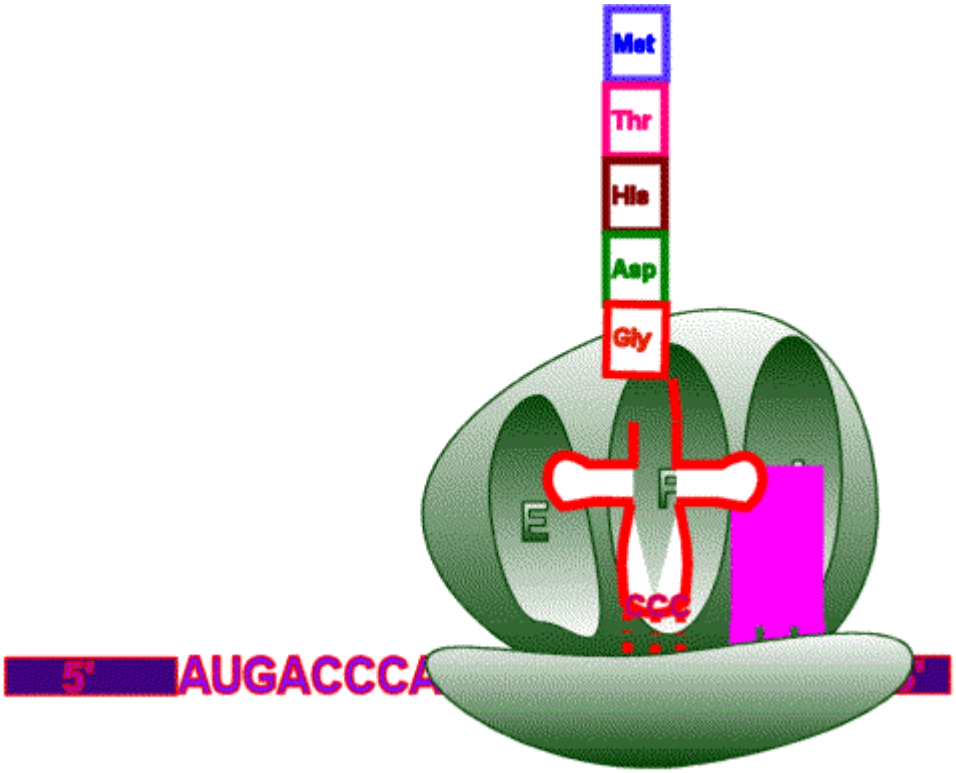
Termination Step 1.

The polypeptide chain is at the P site. The stop codon at the A site.



Termination Step 2.

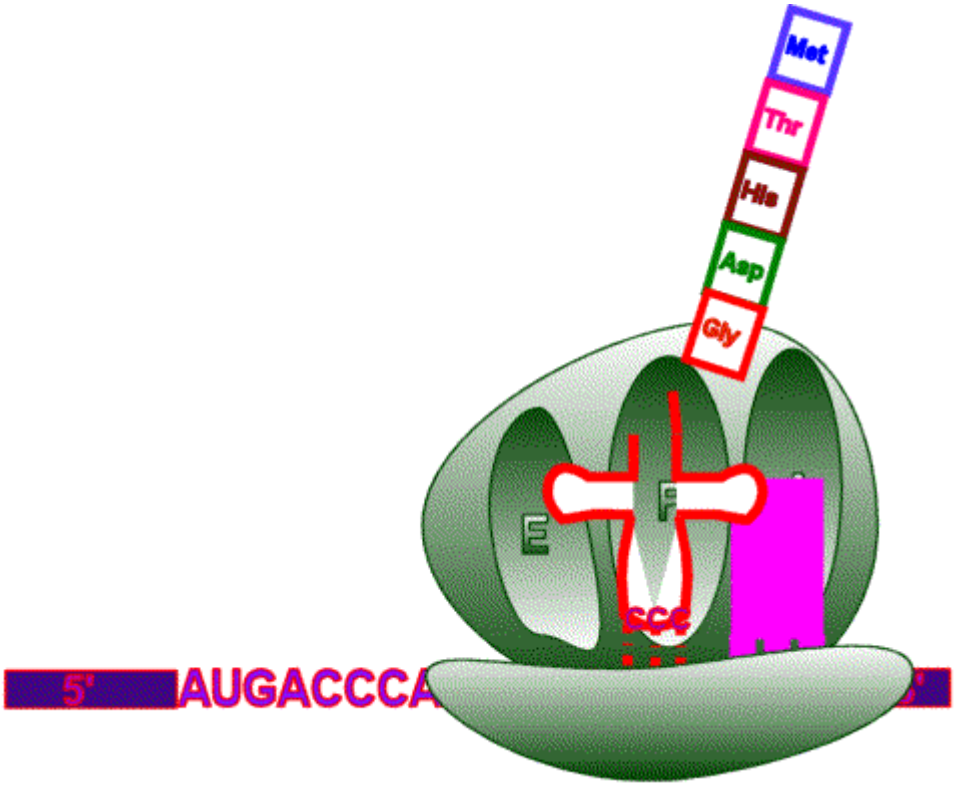
A Release factor protein binds to the stop codon at the A binding site.



Termination Step 3.

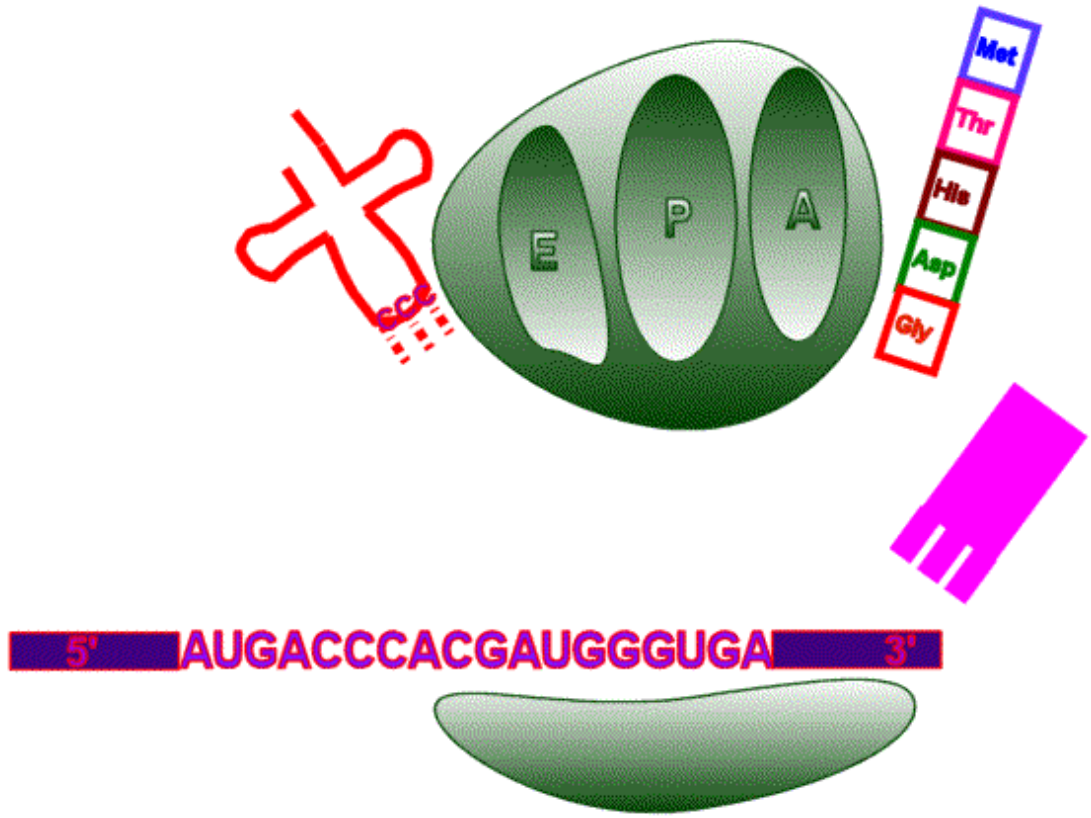
Release factor protein initiates separation of polypeptide chain:

Met-Thr-His-Asp-Gly



Termination Step 4.

Separation of translation machinary. Polypeptide chain may go to cytoplasm for further processing.



Genetic Code

The table below can be used to determine what amino acid corresponds to any 3-letter codon.

First Base	Second Base				Third Base
	U	C	A	G	
U	UUU phenylalanine	UCU serine	UAU tyrosine	UGU cysteine	U
	UUC phenylalanine	UCC serine	UAC tyrosine	UGC cysteine	C
	UUA leucine	UCA serine	UAA stop	UGA stop	A
	UUG leucine	UCG serine	UAG stop	UGG tryptophan	G
C	CUU leucine	CCU proline	CAU histidine	CGU arginine	U

	CUC leucine	CCC proline	CAC histidine	CGC arginine	**C**
	CUA leucine	CCA proline	CAA glutamine	CGA arginine	**A**
	CUG leucine	CCG proline	CAG glutamine	CGG arginine	**G**
A	AUU isoleucine	ACU threonine	AAU asparagine	AGU serine	**U**
	AUC isoleucine	ACC threonine	AAC asparagine	AGC serine	**C**
	AUA isoleucine	ACA threonine	AAA lysine	AGA arginine	**A**
	AUG (start) methionine	ACG threonine	AAG lysine	AGG arginine	**G**
G	GUU valine	GCU alanine	GAU aspartate	GGU glycine	**U**
	GUC valine	GCC alanine	GAC aspartate	GGC glycine	**C**
	GUA valine	GCA alanine	GAA glutamate	GGA glycine	**A**
	GUG valine	GCG alanine	GAG glutamate	GGG glycine	**G**

.

Tools and molecular techniques

1- Southern blot

A Southern blot is a method routinely used in molecular biology for detection of a specific DNA sequence in DNA samples. Southern blotting combines transfer of electrophoresis-separated DNA fragments to a filter membrane and subsequent fragment detection by probe hybridization. The method is named after its inventor, the British biologist Edwin Southern. Other blotting methods (i.e., Western blot, Northern blot, Eastern blot, Southwestern blot) that employ similar principles, but using RNA or protein, have later been named in reference to Edwin Southern's name. As the technique was eponymously named, Southern blot is capitalized as is conventional for proper nouns. The names for other blotting methods may follow this convention, by analogy.

Method

1. Restriction endonucleases are used to cut high-molecular-weight DNA strands into smaller fragments.
2. The DNA fragments are then electrophoresed on an agarose gel to separate them by size.
3. If some of the DNA fragments are larger than 15 kb, then prior to blotting, the gel may be treated with an acid, such as dilute HCl, which depurinates the DNA fragments, breaking the DNA into smaller pieces, thus allowing more efficient transfer from the gel to membrane.
4. If alkaline transfer methods are used, the DNA gel is placed into an alkaline solution (typically containing sodium hydroxide) to denature the double-stranded DNA. The denaturation in an alkaline

environment may improve binding of the negatively charged DNA to a positively charged membrane, separating it into single DNA strands for later hybridization to the probe (see below), and destroys any residual RNA that may still be present in the DNA. The choice of alkaline over neutral transfer methods, however, is often empirical and may result in equivalent results.

5. A sheet of nitrocellulose (or, alternatively, nylon) membrane is placed on top of (or below, depending on the direction of the transfer) the gel. Pressure is applied evenly to the gel (either using suction, or by placing a stack of paper towels and a weight on top of the membrane and gel), to ensure good and even contact between gel and membrane. If transferring by suction 20X SSC buffer is used to ensure a seal and prevent drying of the gel. Buffer transfer by capillary action from a region of high water potential to a region of low water potential (usually filter paper and paper tissues) is then used to move the DNA from the gel on to the membrane; ion exchange interactions bind the DNA to the membrane due to the negative charge of the DNA and positive charge of the membrane.
6. The membrane is then baked in a vacuum or regular oven at 80 ℃ for 2 hours (standard conditions; nitrocellulose or nylon membrane) or exposed to ultraviolet radiation (nylon membrane) to permanently attach the transferred DNA to the membrane.
7. The membrane is then exposed to a hybridization probe—a single DNA fragment with a specific sequence whose presence in the target DNA is to be determined. The probe DNA is labelled so that it can be detected, usually by incorporating radioactivity or tagging the molecule with a fluorescent or chromogenic dye. In some

cases, the hybridization probe may be made from RNA, rather than DNA. To ensure the specificity of the binding of the probe to the sample DNA, most common hybridization methods use salmon or herring sperm DNA for blocking of the membrane surface and target DNA, deionized formamide, and detergents such as SDS to reduce non-specific binding of the probe.

8. After hybridization, excess probe is washed from the membrane (typically using SSC buffer), and the pattern of hybridization is visualized on X-ray film by autoradiography in the case of a radioactive or fluorescent probe, or by development of color on the membrane if a chromogenic detection method is used.

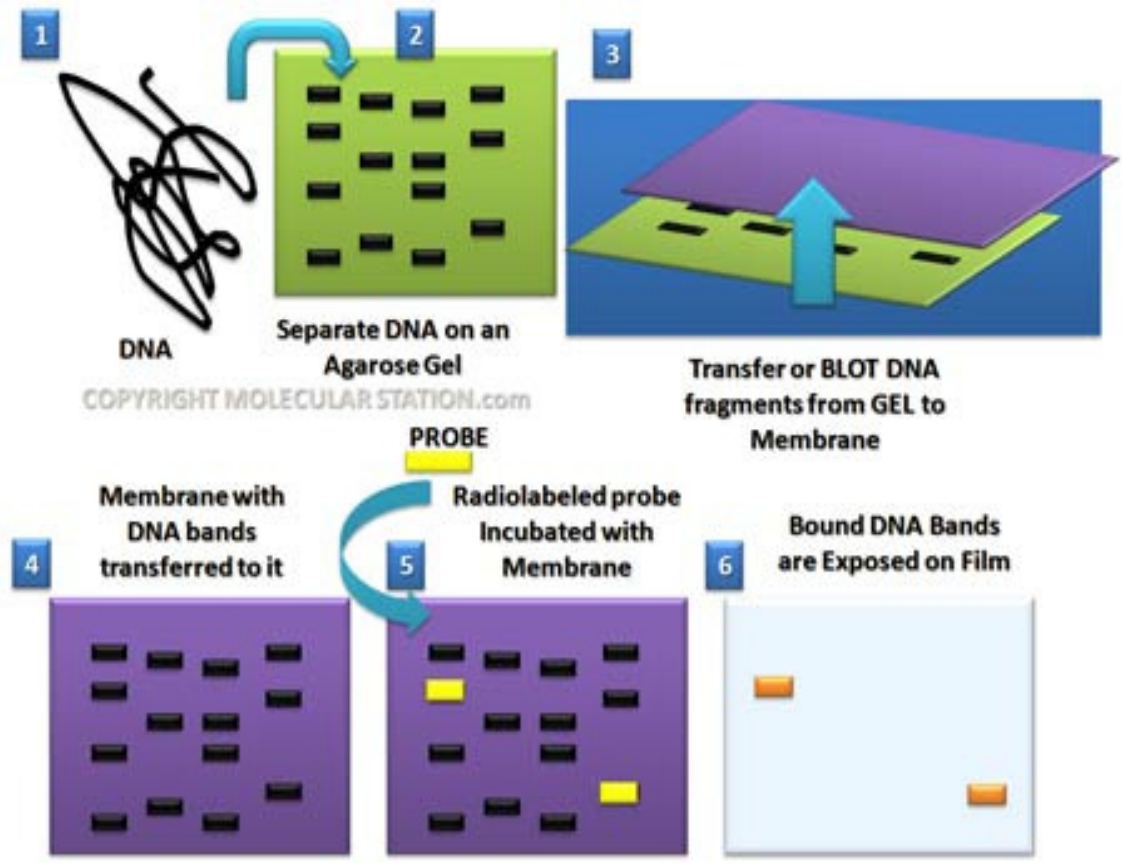
Result

Hybridization of the probe to a specific DNA fragment on the filter membrane indicates that this fragment contains DNA sequence that is complementary to the probe.

The transfer step of the DNA from the electrophoresis gel to a membrane permits easy binding of the labeled hybridization probe to the size-fractionated DNA. It also allows for the fixation of the target-probe hybrids, required for analysis by autoradiography or other detection methods.

Southern blots performed with restriction enzyme-digested genomic DNA may be used to determine the number of sequences (e.g., gene copies) in a genome. A probe that hybridizes only to a single DNA segment that has not been cut by the restriction enzyme will produce a single band on a Southern blot, whereas multiple bands will likely be observed when the probe hybridizes to several highly similar sequences (e.g., those that

may be the result of sequence duplication). Modification of the hybridization conditions (for example, increasing the hybridization temperature or decreasing salt concentration) may be used to increase specificity and decrease hybridization of the probe to sequences that are less than 100% similar.



2-The polymerase chain reaction

The polymerase chain reaction (PCR) is a powerful technique that has rapidly become one of the most widely used techniques in molecular biology because it is quick, inexpensive and simple. The technique amplifies specific DNA fragments from minute quantities of source DNA material, even when that source DNA is of relatively poor quality.

Principle of the PCR

The purpose of a PCR (Polymerase Chain Reaction) is to make a huge number of copies of a gene. This is necessary to have enough starting template for sequencing.

1. **The cycling reactions :**

   There are three major steps in a PCR, which are repeated for 30 or 40 cycles. This is done on an automated cycler, which can heat and cool the tubes with the reaction mixture in a very short time.

   The DNA polymerase, known as 'Taq polymerase', is named after the hot-spring bacterium Thermus aquaticus from which it was originally isolated. The enzyme can withstand the high temperatures needed for DNA-strand separation, and can be left in the reaction tube.

   The cycle of heating and cooling is repeated over and over, stimulating the primers to bind to the original sequences and to newly synthesised sequences. The enzyme will again extend primer sequences. This cycling of temperatures results in copying and then copying of copies, and so on, leading to an exponential increase in the number of copies of specific sequences. Because the amount of DNA placed in the tube at the beginning is very small, almost all the DNA at the end of the reaction cycles is copied sequences.

   The reaction products are separated by gel electrophoresis. Depending on the quantity produced and the size of the amplified fragment, the reaction products can be visualised

directly by staining with ethidium bromide or a silver-staining protocol, or by means of radioisotopes and autoradiography.

PCR procedures steps

- Denaturation: DNA fragments are heated at high temperatures, which reduce the DNA double helix to single strands. These strands become accessible to primers.
- Annealing: The reaction mixture is cooled down. Primers anneal to the complementary regions in the DNA template strands, and double strands are formed again between primers and complementary sequences.
- Extension: The DNA polymerase synthesizes a complementary strand. The enzyme reads the opposing strand sequence and extends the primers by adding nucleotides in the order in which they can pair. The whole process is repeated over and over

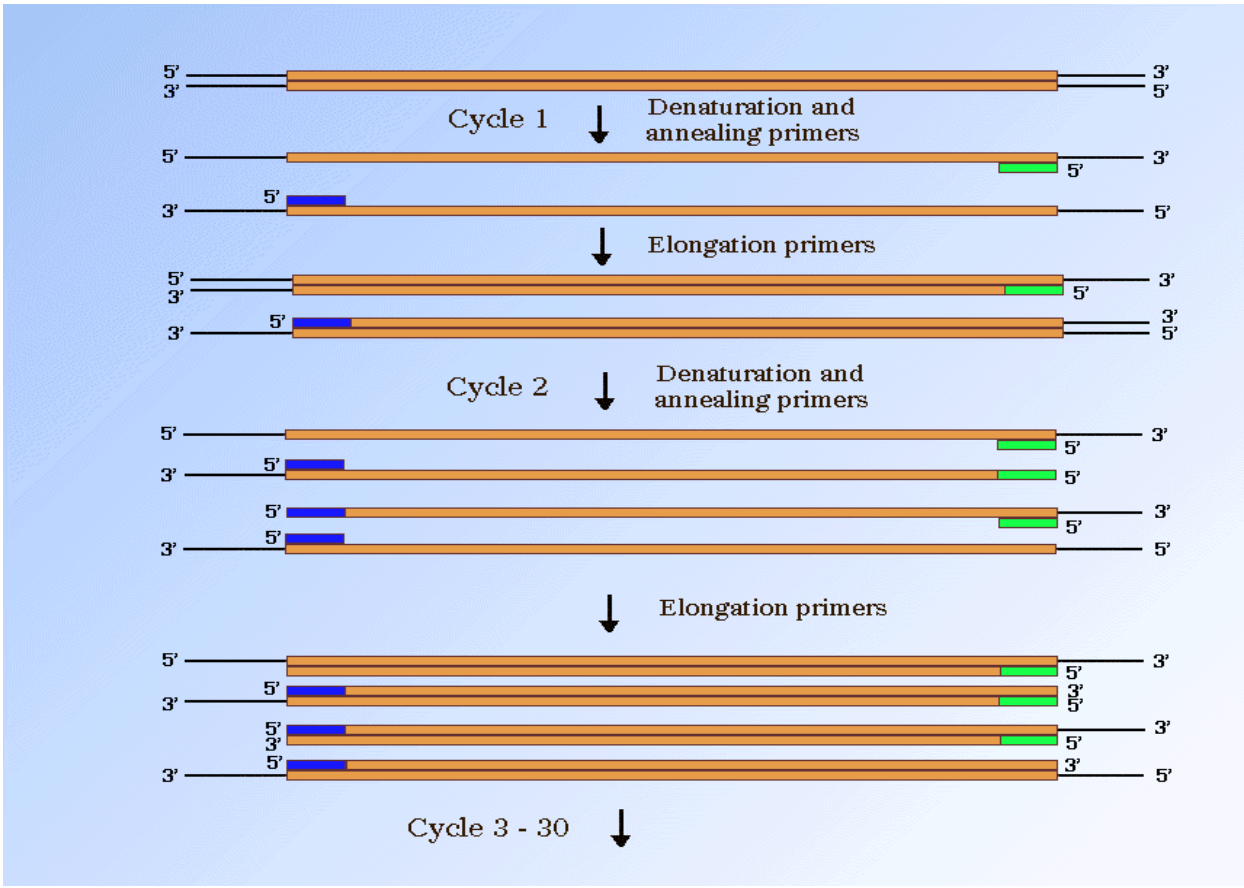


Figure 3 : The different steps in PCR. Animated picture of PCR

Because both strands are copied during PCR, there is an **exponential** increase of the number of copies of the gene. Suppose there is only one copy of the wanted gene before the cycling starts, after one cycle, there will be 2 copies, after two cycles, there will be 4 copies, three cycles will result in 8 copies and so on.

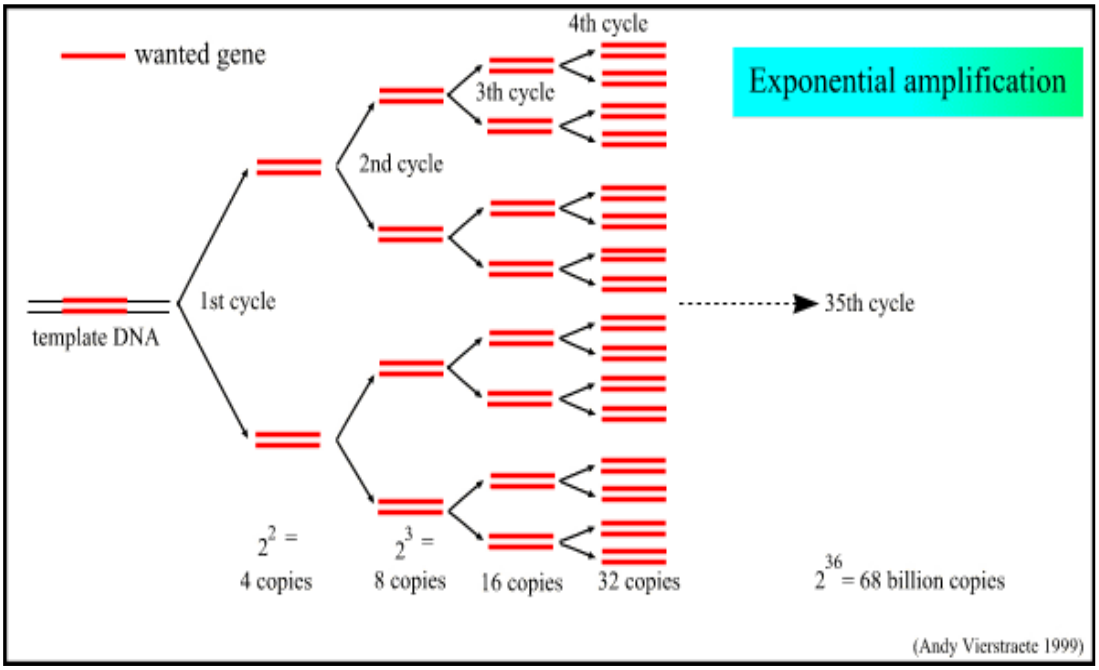


Figure 4 : The exponential amplification of the gene in PCR.

2. **Is there a gene copied during PCR and is it the right size ?**

Before the PCR product is used in further applications, it has to be checked if :

1. There is a product formed.
   Though biochemistry is an exact science, not every PCR is successful. There is for example a possibility that the quality of the DNA is poor, that one of the primers doesn't fit, or that there is too much starting template
2. The product is of the right size
   It is possible that there is a product, for example a band of 500 bases, but the expected gene should be 1800 bases long. In that case, one of the primers probably fits on a part of the gene closer to the other primer. It is also possible that both primers fit on a totally different gene.

3. Only one band is formed.
   As in the description above, it is possible that the primers fit on the desired locations, and also on other locations. In that case, you can have different bands in one lane on a gel.

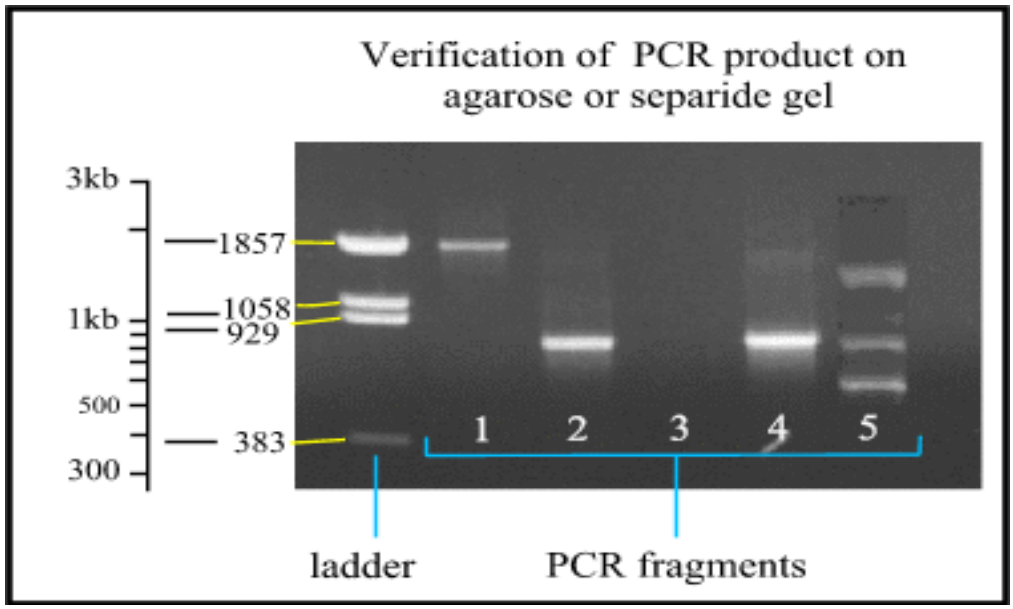


Figure 5 : Verification of the PCR product on gel.

The ladder is a mixture of fragments with known size to compare with the PCR fragments. Notice that the distance between the different fragments of the ladder is logarithmic. Lane 1 : PCR fragment is approximately 1850 bases long. Lane 2 and 4 : the fragments are approximately 800 bases long. Lane 3 : no product is formed, so the PCR failed. Lane 5 : multiple bands are formed because one of the primers fits on different places.

Applications of PCR:

Medical applications

PCR has been applied to a large number of medical procedures:

- The first application of PCR was for genetic testing, where a sample of DNA is analyzed for the presence of genetic disease mutations. Prospective parents can be tested for being genetic carriers, or their children might be tested for actually being affected by a disease. DNA samples for Prenatal testing can be obtained by amniocentesis, chorionic villus sampling, or even by the analysis of rare fetal cells circulating in the mother's bloodstream. PCR analysis is also essential to Preimplantation genetic diagnosis, where individual cells of a developing embryo are tested for mutations.

- PCR can also be used as part of a sensitive test for tissue typing, vital to organ transplantation. As of 2008, there is even a proposal to replace the traditional antibody-based tests for blood type with PCR-based tests.

- Many forms of cancer involve alterations to oncogenes. By using PCR-based tests to study these mutations, therapy regimens can sometimes be individually customized to a patient.

Infectious disease applications

Characterization and detection of infectious disease organisms have been revolutionized by PCR:

- The Human Immunodeficiency Virus (or HIV), responsible for AIDS, is a difficult target to find and eradicate. The earliest tests for infection relied on the presence of antibodies to the virus circulating in the bloodstream. However, antibodies don't appear until many weeks after infection, maternal antibodies mask the infection of a newborn, and therapeutic agents to fight the infection don't affect the

antibodies. PCR tests have been developed that can detect as little as one viral genome among the DNA of over 50,000 host cells. Infections can be detected earlier, donated blood can be screened directly for the virus, newborns can be immediately tested for infection, and the effects of antiviral treatments can be quantified.

- Some disease organisms, such as that for Tuberculosis, are difficult to sample from patients and slow to be grown in the laboratory. PCR-based tests have allowed detection of small numbers of disease organisms (both live or dead), in convenient samples. Detailed genetic analysis can also be used to detect antibiotic resistance, allowing immediate and effective therapy. The effects of therapy can also be immediately evaluated.

- The spread of a disease organism through populations of domestic or wild animals can be monitored by PCR testing. In many cases, the appearance of new virulent sub-types can be detected and monitored. The sub-types of an organism that were responsible for earlier epidemics can also be determined by PCR analysis.

Forensic applications

The development of PCR-based genetic (or DNA) fingerprinting protocols has seen widespread application in forensics:

- In its most discriminating form, Genetic fingerprinting can uniquely discriminate any one person from the entire population of the world. Minute samples of DNA can be isolated from a crime scene, and compared to that from suspects, or from a DNA database of earlier evidence or convicts. Simpler versions of these tests are often used to

rapidly rule out suspects during a criminal investigation. Evidence from decades-old crimes can be tested, confirming or exonerating the people originally convicted.

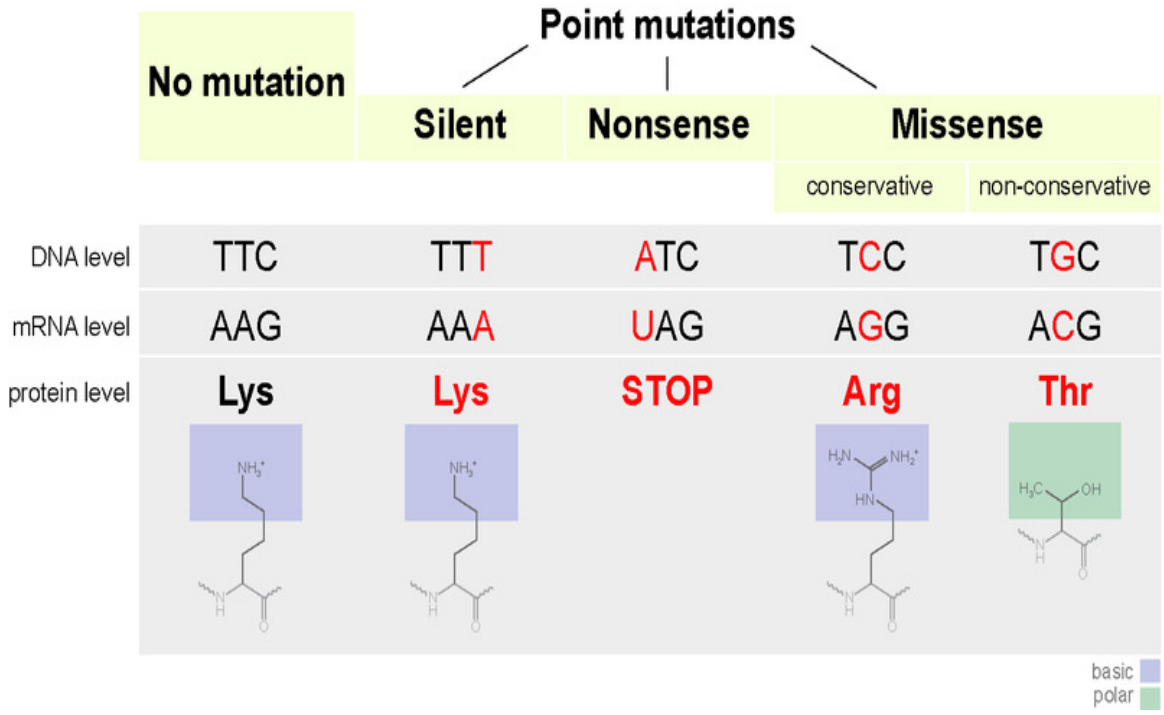
- Less discriminating forms of DNA fingerprinting can help in Parental testing, where an individual is matched with their close relatives. DNA from unidentified human remains can be tested, and compared with that from possible parents, siblings, or children. Similar testing can be used to confirm the biological parents of an adopted (or kidnapped) child. The actual biological father of a newborn can also be confirmed (or ruled out).

Mutations

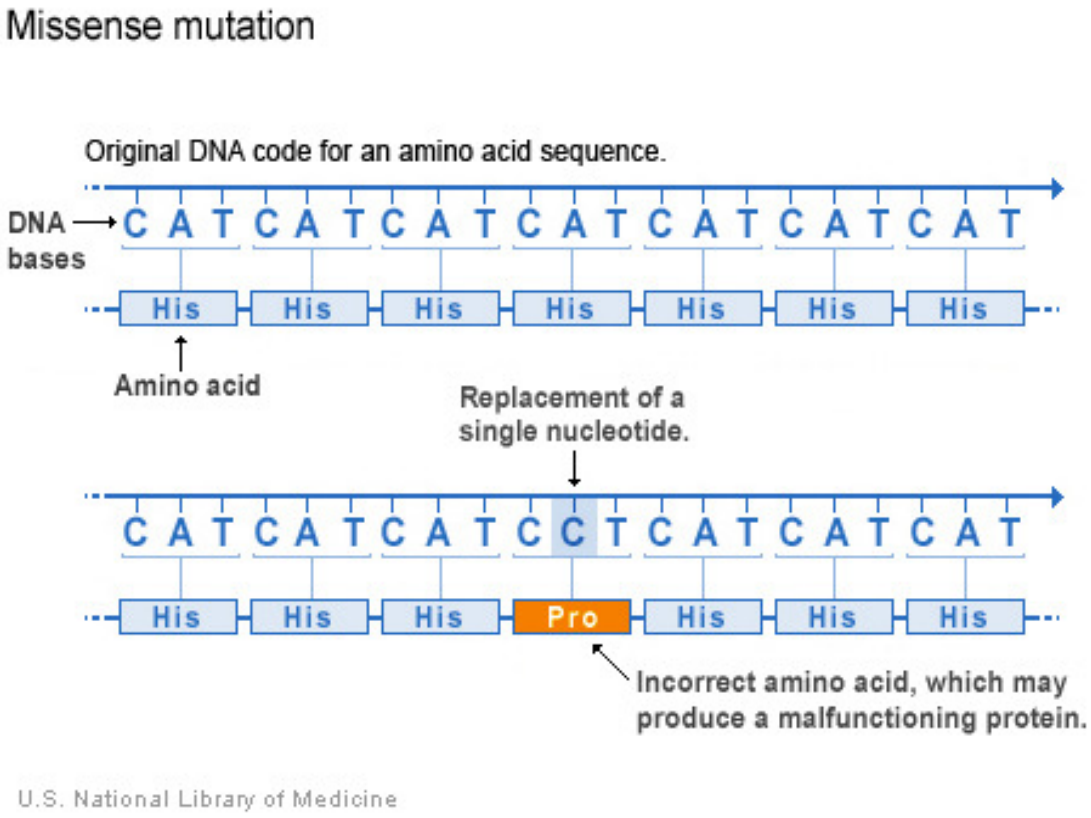
A mutation involving a change in a single base pair, often called a point mutation, or a deletion of a few base pairs generally affects the function of a single gene. Changes in a single base pair may produce one of two types of mutation:

- **Point Mutations:** Includes the deletion, insertion, or substitution of ONE nucleotide in a gene (If one purine [A or G] or pyrimidine [C or T] is replaced by the other, the substitution is called a transition. If a purine is replaced by a pyrimidine or vice-versa, the substitution is called a transversion).

- **Gross mutations:** It involves long stretches of DNA``

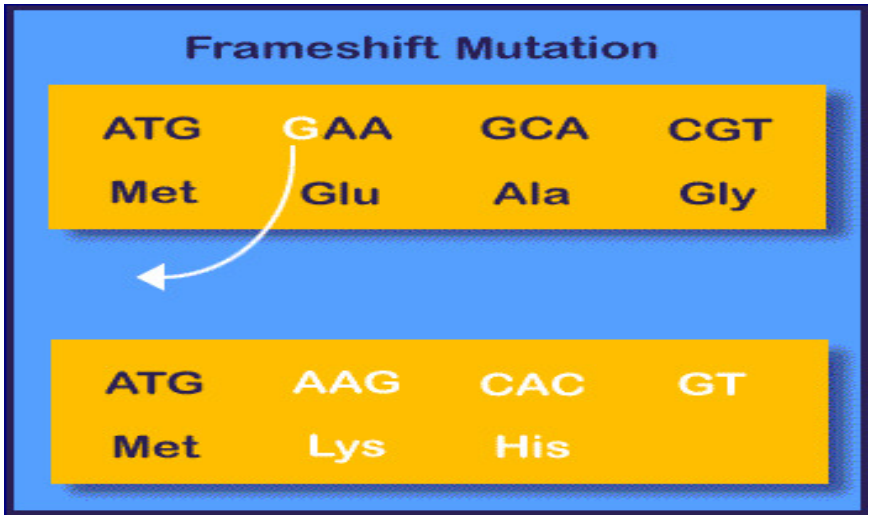


Mis sense mutation, which results in a protein in which one amino acid is substituted for another:

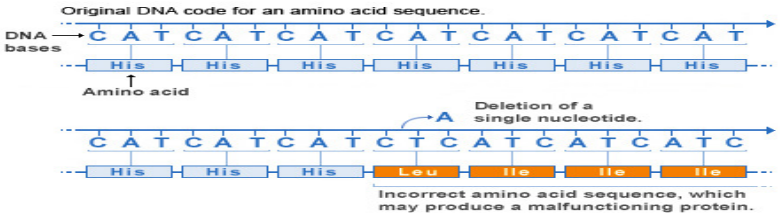


• Nonsense mutation, in which a stop codon replaces an amino acid codon, leading to premature termination of translation

• Frame shift mutation, which causes a change in the reading frame, leading to introduction of unrelated amino acids into the protein, generally followed by a stop codon.

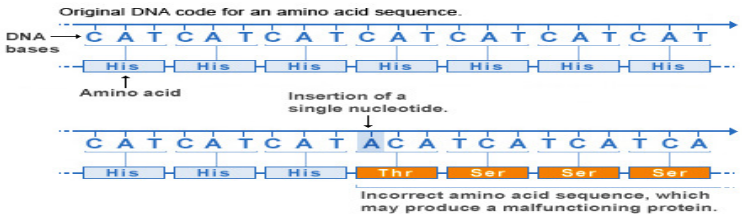


Deletion mutation

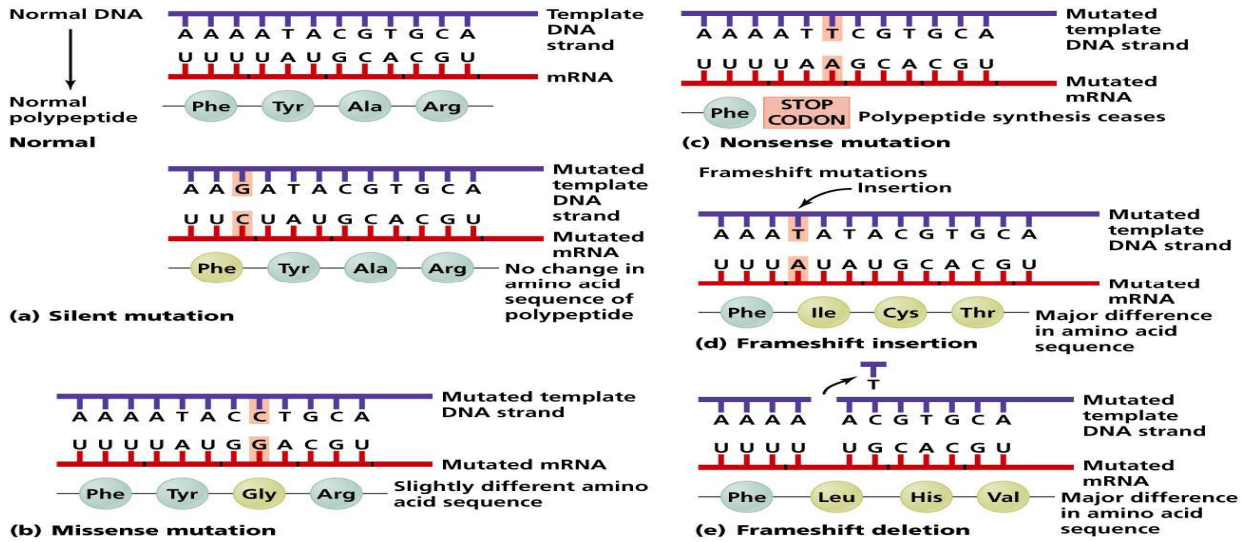


U.S. National Library of Medicine

Insertion mutation



U.S. National Library of Medicine



Gross Mutations:

Cause alteration to DNA involving long stretches of sequences

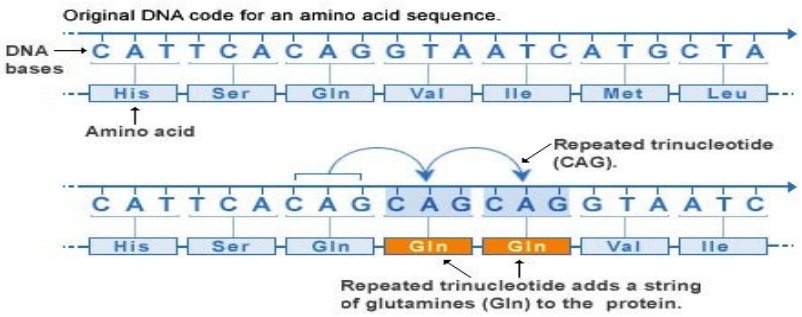
1- deletion

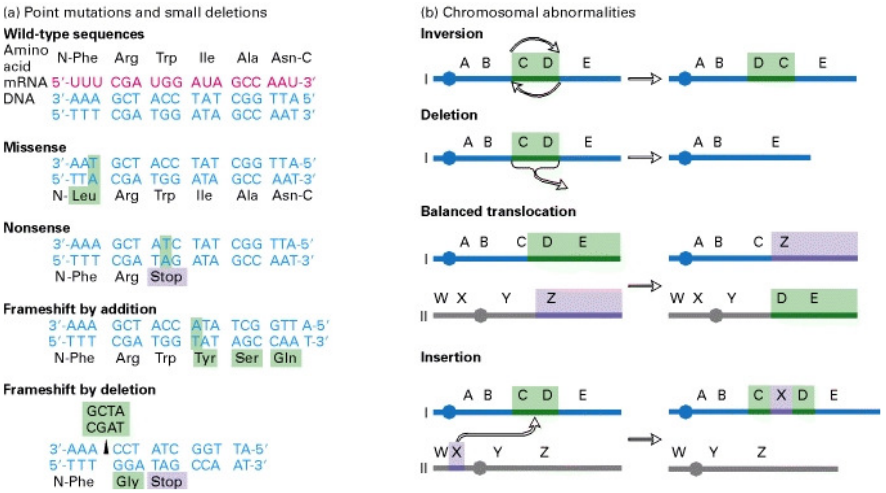
2- insertion

3-rearangement

4- repetition

Repeat expansion mutation




(a) Point mutations and small deletions
Wild-type sequences
Amino acid N-Phe Arg Trp Ile Ala Asn-C
mRNA 5′-UUU CGA UGG AUA GCC AAU-3′
DNA 3′-AAA GCT ACC TAT CGG TTA 5′
5′-TTT CGA TGG ATA GCC AAT 3′
Missense
3′-AAT GCT ACC TAT CGG TTA-5′
5′-TTA CGA TGG ATA GCC AAT-3′
N-Leu Arg Trp Ile Ala Asn-C
Nonsense
3′-AAA GCT ATC TAT CGG TTA-5′
5′-TTT CGA TAG ATA GCC AAT-3′
N-Phe Arg Stop
Frameshift by addition
3′-AAA GCT ACC ATA TCG GTT A-5′
5′-TTT CGA TGG TAT AGC CAA T-3′
N-Phe Arg Trp Tyr Ser Gln
Frameshift by deletion
GCTA
CGAT
3′-AAA CCT ATC GGT TA-5′
5′-TTT GGA TAG CCA AT-3′
N-Phe Gly Stop
(b) Chromosomal abnormalities
Inversion
A B C D E
A B D C E
Deletion
A B C D E
A B E
Balanced translocation
A B C D E
A B C Z
W X Y Z
W X Y D E
Insertion
A B C D E
A B C X D E
W X Y Z
W Y Z

References

1. **Jump up ^** Russell, Peter (2001). *iGenetics*. New York: Benjamin Cummings. ISBN 0-8053-4553-1.
2. **Jump up ^** Mashaghi A, Katan A (2013). "A physicist's view of DNA". *De Physicus* **24e** (3): 59–61. arXiv:1311.2545v1. Bibcode:2013arXiv1311.2545M.
3. **Jump up ^** Saenger, Wolfram (1984). *Principles of Nucleic Acid Structure*. New York: Springer-Verlag. ISBN 0-387-90762-9.
4. ^ Jump up to: **a** **b** Alberts, Bruce; Alexander Johnson, Julian Lewis, Martin Raff, Keith Roberts and Peter Walters (2002). *Molecular Biology of the Cell; Fourth Edition*. New York and London: Garland Science. ISBN 0-8153-3218-1. OCLC 145080076 48122761 57023651 69932405.
5. **Jump up ^** Butler, John M. (2001). *Forensic DNA Typing*. Elsevier. ISBN 978-0-12-147951-0. OCLC 223032110 45406517. pp. 14–15.
6. ^ Jump up to: **a** **b** **c** Watson JD, Crick FH (1953). "A Structure for Deoxyribose Nucleic Acid" (PDF). *Nature* **171** (4356): 737–738. Bibcode:1953Natur.171..737W. doi:10.1038/171737a0. PMID 13054692.
7. **Jump up ^** Mandelkern M, Elias JG, Eden D, Crothers DM (1981). "The dimensions of DNA in solution". *J Mol Biol* **152** (1): 153–61. doi:10.1016/0022-2836(81)90099-1. PMID 7338906.
8. **Jump up ^** Gregory SG, Barlow KF, McLay KE, Kaul R, Swarbreck D, Dunham A, et al. (2006). "The DNA sequence and biological annotation of human chromosome 1". *Nature* **441** (7091): 315–21. Bibcode:2006Natur.441..315G. doi:10.1038/nature04727. PMID 16710414.
9. **Jump up ^** Watson JD, Crick FH (1953). "A Structure for Deoxyribose Nucleic Acid" (PDF). *Nature* **171** (4356): 737–738. Bibcode:1953Natur.171..737W. doi:10.1038/171737a0. PMID 13054692. Retrieved 4 May 2009.
10. ^ Jump up to: **a** **b** **c** Berg J., Tymoczko J. and Stryer L. (2002) *Biochemistry.* W. H. Freeman and Company ISBN 0-7167-4955-6
11. **Jump up ^** Abbreviations and Symbols for Nucleic Acids, Polynucleotides and their Constituents IUPAC-IUB Commission on Biochemical Nomenclature (CBN). Retrieved 3 January 2006.
12. ^ Jump up to: **a** **b** Ghosh A, Bansal M (2003). "A glossary of DNA structures from A to Z". *Acta Crystallogr D* **59** (4): 620–6. doi:10.1107/S0907444903003251. PMID 12657780.
13. **Jump up ^** Created from PDB 1D65
14. **Jump up ^** Yakovchuk P, Protozanova E, Frank-Kamenetskii MD (2006). "Base-stacking and base-pairing contributions into thermal stability of the DNA double helix". *Nucleic Acids Res.* **34** (2): 564–74. doi:10.1093/nar/gkj454. PMC 1360284. PMID 16449200.
15. **Jump up ^** Verma S, Eckstein F (1998). "Modified oligonucleotides: synthesis and strategy for users". *Annu. Rev. Biochem.* **67**: 99–134. doi:10.1146/annurev.biochem.67.1.99. PMID 9759484.
16. **Jump up ^** Kiljunen S, Hakala K, Pinta E, Huttunen S, Pluta P, Gador A, Lönnberg H, Skurnik M (2005). "Yersiniophage phiR1-37 is a tailed bacteriophage having a 270 kb DNA genome with thymidine replaced by

deoxyuridine". *Microbiology* **151** (12): 4093–4102. doi:10.1099/mic.0.28265-0. PMID 16339954.

17. Anthony J. F. Griffiths et al. (1999). "8. The Structure and Replication of DNA". *An Introduction to genetic analysis*. San Francisco: W.H. Freeman. ISBN 0-7167-3520-2.
18. ^ Jump up to: *a* *b* Meselson, M.; Stahl, F.W. (1958). "The Replication of DNA in Escherichia coli". *PNAS* **44**: 671–82. doi:10.1073/pnas.44.7.671. PMC 528642. PMID 16590258.
19. **Jump up ^** Meselson M, Stahl FW. (2007) Demonstration of the semiconservative mode of DNA duplication. In “Phage and the Origins of Molecular Biology” (editors Cairns J, Stent GS, Watson JD). Cold Spring Harbor Laboratory Press ISBN 9780879698003
20. **Jump up ^** McCarthy D, Minner C, Bernstein H, Bernstein C (October 1976). "DNA elongation rates and growing point distributions of wild-type phage T4 and a DNA-delay amber mutant". *J. Mol. Biol.* **106** (4): 963–81. doi:10.1016/0022-2836(76)90346-6. PMID 789903.
21. **Jump up ^** Drake JW, Charlesworth B, Charlesworth D, Crow JF (April 1998). "Rates of spontaneous mutation". *Genetics* **148** (4): 1667–86. PMC 1460098. PMID 9560386.
22. Blow JJ, Dutta A (June 2005). "Preventing re-replication of chromosomal DNA". *Nat. Rev. Mol. Cell Biol.* **6** (6): 476–86. doi:10.1038/nrm1663. PMC 2688777. PMID 15928711.