DNA structure:

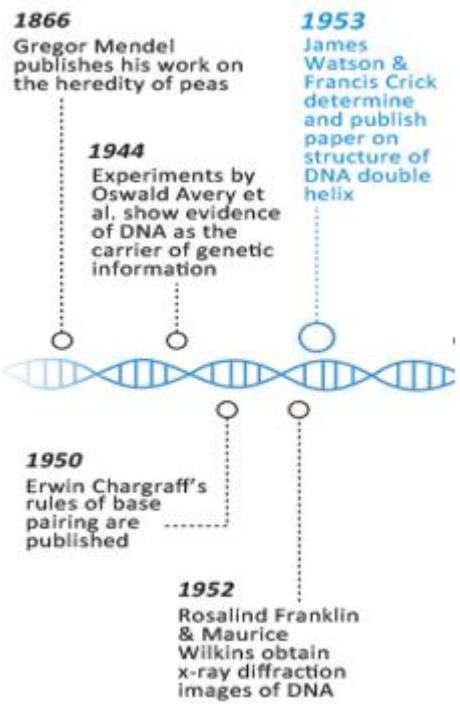


Figure 1- Timeline of events in the discovery of DNA structure. Image courtesy www.genscript.com

The race for DNA structure:

The structure of DNA remained elusive until experimental evidence of many types were considered together in a theoretical framework. Figure 1 indicates the various experiments which acted as steps towards the discovery of the correct DNA structure.

The most crucial evidence was obtained using X-ray crystallography. Some chemical substances, when they are isolated and purified, can be made to form crystals. The positions of atoms in a crystallized substance can be inferred from the diffraction pattern of X- rays passing through the substance.

The events that provided information about this vital molecule are described in the following text.

Chemical composition of DNA:

Biochemists knew that DNA was a polymer of nucleotides. Each nucleotide consists of a molecule of the sugar deoxyribose, a phosphate group, and a nitrogen containing base.

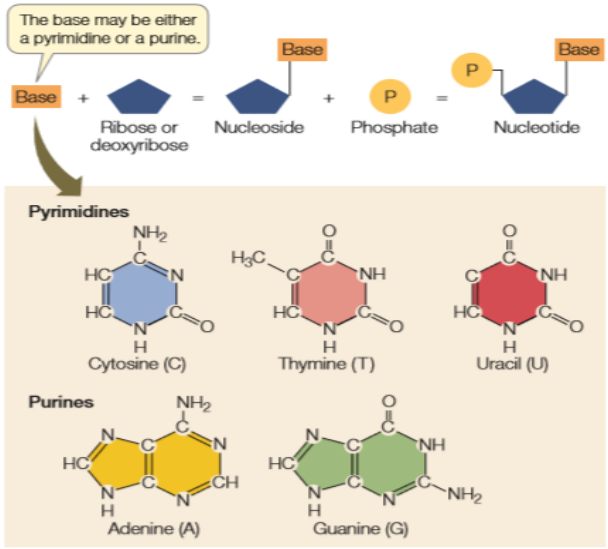


Figure 2- Chemical composition of DNA monomers. The only differences among the four nucleotides of DNA are their nitrogenous bases: the purines adenine (A) and guanine (G), and the pyrimidines cytosine (C) and thymine (T). Image courtesy- Sadava et al, Life: The science of Biology, 9th edition.

In 1950, biochemist Erwin Chargaff reported that DNA from many different species—and from different sources within a single organism—exhibits certain regularities. In almost all DNA, the following rule holds: The amount of adenine equals the amount of thymine (A= T), and the amount of guanine equals the amount of cytosine (G = C).

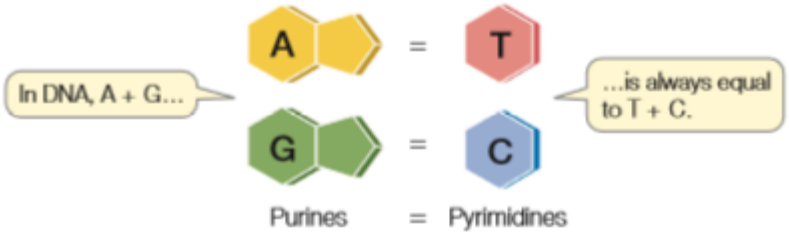


Figure 3- Chargaff's rule- In DNA, total abundance of purines is equal to total abundance of pyrimidines. Image courtesy- Sadava et al, Life: The science of Biology, 9th edition.

In 1952, Rosalind Franklin was able to obtain an X-ray diffraction pattern for certain DNA fibers. This experiment provided the greatest help required by the scientists to deduce the DNA structure.

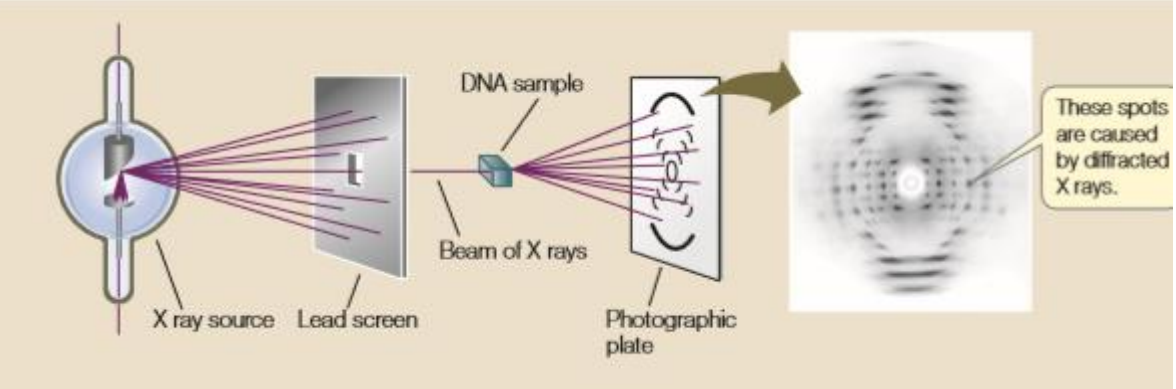


Figure 4- The positions of atoms in a crystallized chemical substance can be inferred by the pattern of diffraction of X rays passed through it. The pattern of DNA is both highly regular and repetitive. Image courtesy- Sadava et al, Life: The science of Biology, 9th edition.

Around the same time, Linus Pauling proposed a triple stranded helical structure for DNA. Linus Pauling had discovered the helical nature of protein folding and deduced that the same folding pattern may be followed by DNA. The structure proposed by Linus Pauling had the phosphate groups of the nucleotides facing inside the helical core. However, such a structure would result in the phosphate group repulsion (due to the negatively charged oxygen groups). Almost unbelievable that the man who had such a command over chemical bonds would get this wrong.

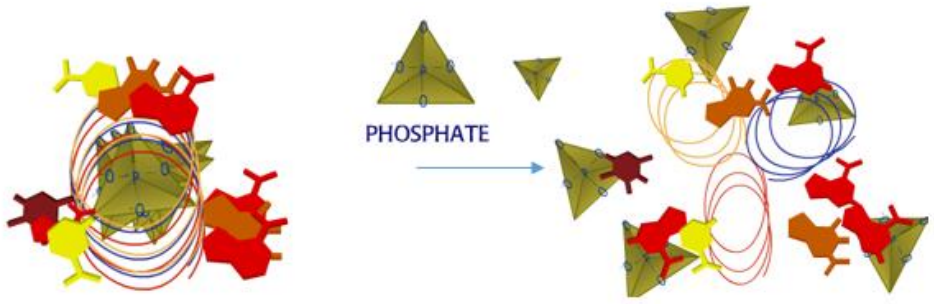


Figure 5- DNA structure as proposed by Linus Pauling (Top view). The negatively charged oxygen atoms repel each other and would cause the strands to disassociate. Image courtesy- http://www.dnai.org/

Double stranded structure discovery by Watson and Crick:

The English physicist Francis Crick and the American geneticist James D. Watson, who were both then at the Cavendish Laboratory of Cambridge University, used model building to solve the structure of DNA.

Watson and Crick attempted to combine all that had been learned so far about DNA structure into a single coherent model. Rosalind Franklin's crystallography results (see Figure 4) convinced Watson and Crick that the DNA molecule must be helical (cylindrically spiral). Density measurements and previous model building results suggested that there are two polynucleotide chains in the molecule. Modeling studies also showed that the strands run in opposite directions, that is, they are antiparallel; that two strands would not fit together in the model if they were parallel.

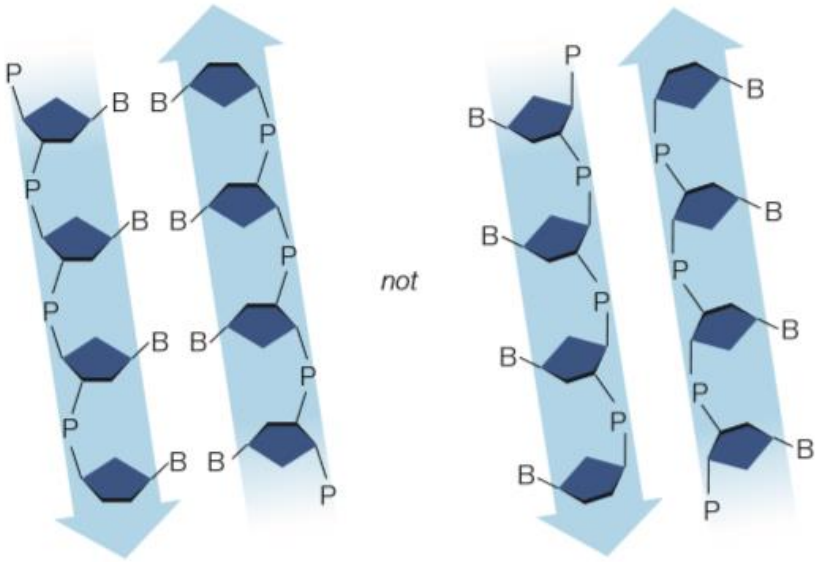


Figure 6- DNA model proposed by Watson and Crick made several assumptions. The nucleotide bases are on the interior of the two strands, with a sugar-phosphate backbone on the outside. Image courtesy- Sadava et al, Life: The science of Biology, 9th edition.

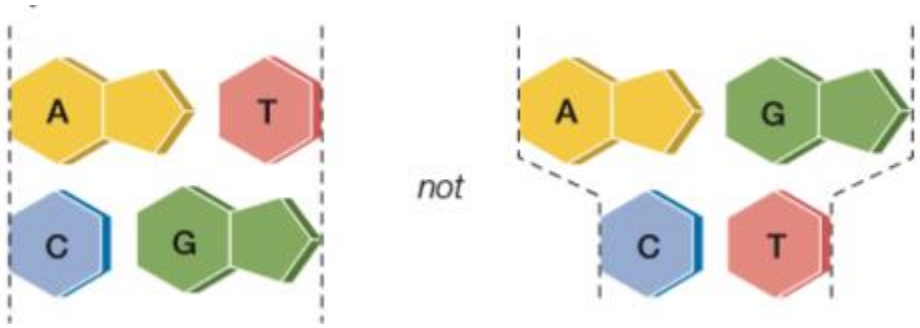


Figure 7-To satisfy Chargaff's rule (purines = pyrimidines), a purine on one strand is always paired with a pyrimidine on the opposite strand. These base pairs (A-T and G-C) have the same width down the double helix, a uniformity shown by x-ray diffraction. Image courtesy- Sadava et al, Life: The science of Biology, 9th edition.

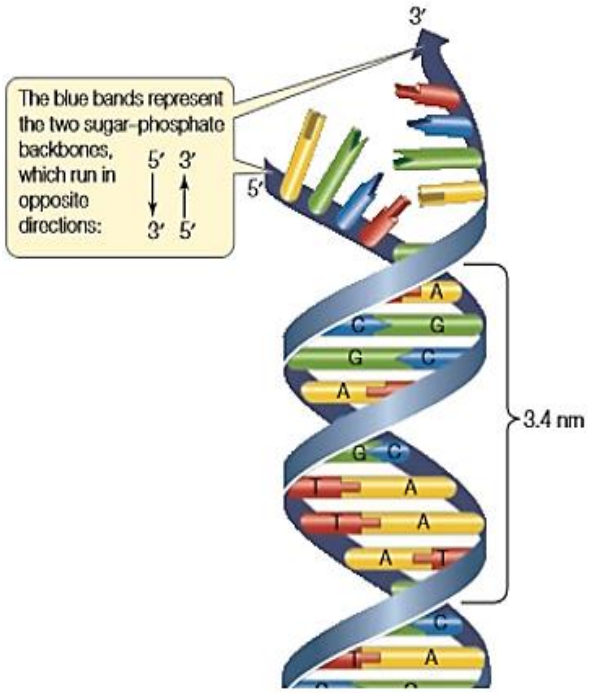


Figure 8- In late February of 1953, Crick and Watson built a model out of tin that established the general structure of DNA. This structure explained all the known chemical properties of DNA, and it opened the door to understanding its biological functions. Image courtesy- Sadava et al, Life: The science of Biology, 9th edition.

Important properties of DNA structure:

1. The double stranded helix has a uniform diameter
2. The two strands rum in opposite direction (antiparallel).
3. The backbone of each strand is made up of sugar phosphate groups linked by phosphodiester bonds
4. The two strands are held together by hydrogen bonding between the nitrogenous bases.

Meselson and Stahl experiment:

This experiment was instrumental in proving that DNA follows a semiconservative model of replication. A cell while undergoing division requires that old cell to produce copies of DNA that can be transferred to the new cells. The new cells receive a copy of the parent cell's DNA.

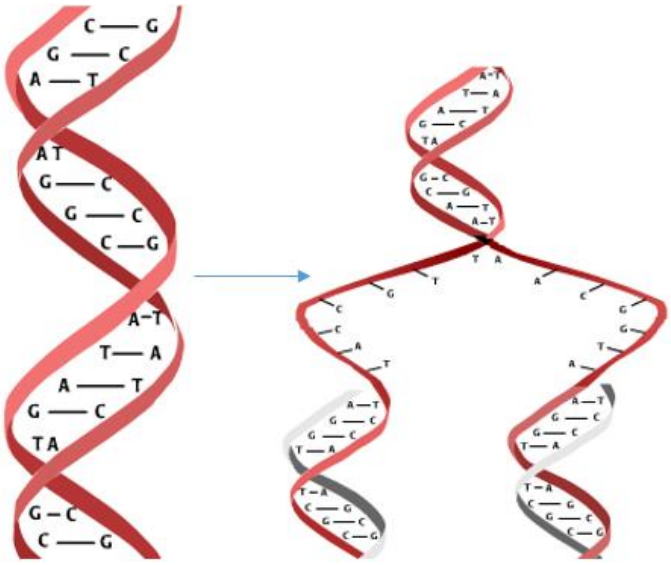


Figure 9- Watson and Crick suggested a semiconservative model of replication, wherein each parental strand acts as a template for synthesizing a new complementary strand. In this model, each daughter molecules consists of one old strand (from parent molecule) and one newly synthesized strand. Image courtesy- www.dnai.org

Meselson and Stahl made clever use of radiolabelled (^{15}N) heavy isotope of nucleotides and a density gradient of Cesium chloride (CsCl) to provide evidence for the semiconservative model of replication.

They collected some of the bacteria after each division and extracted DNA from the samples. To separate the DNA from the cells at different generation on basis of density, they developed a density gradient solution in a test tube using CsCl. They found that the density gradient was different in each bacterial generation:

• At the time of the transfer to the ^{14}N medium, the DNA was uniformly labeled with ^{15}N, and hence formed a single band corresponding with dense DNA.

• After one generation in the ^{14}N medium, when the DNA had been duplicated once, all the DNA was of intermediate density.

• After two generations, there were two equally large DNA bands: one of low density and one of intermediate density.

• In samples from subsequent generations, the proportion of low-density DNA increased steadily.

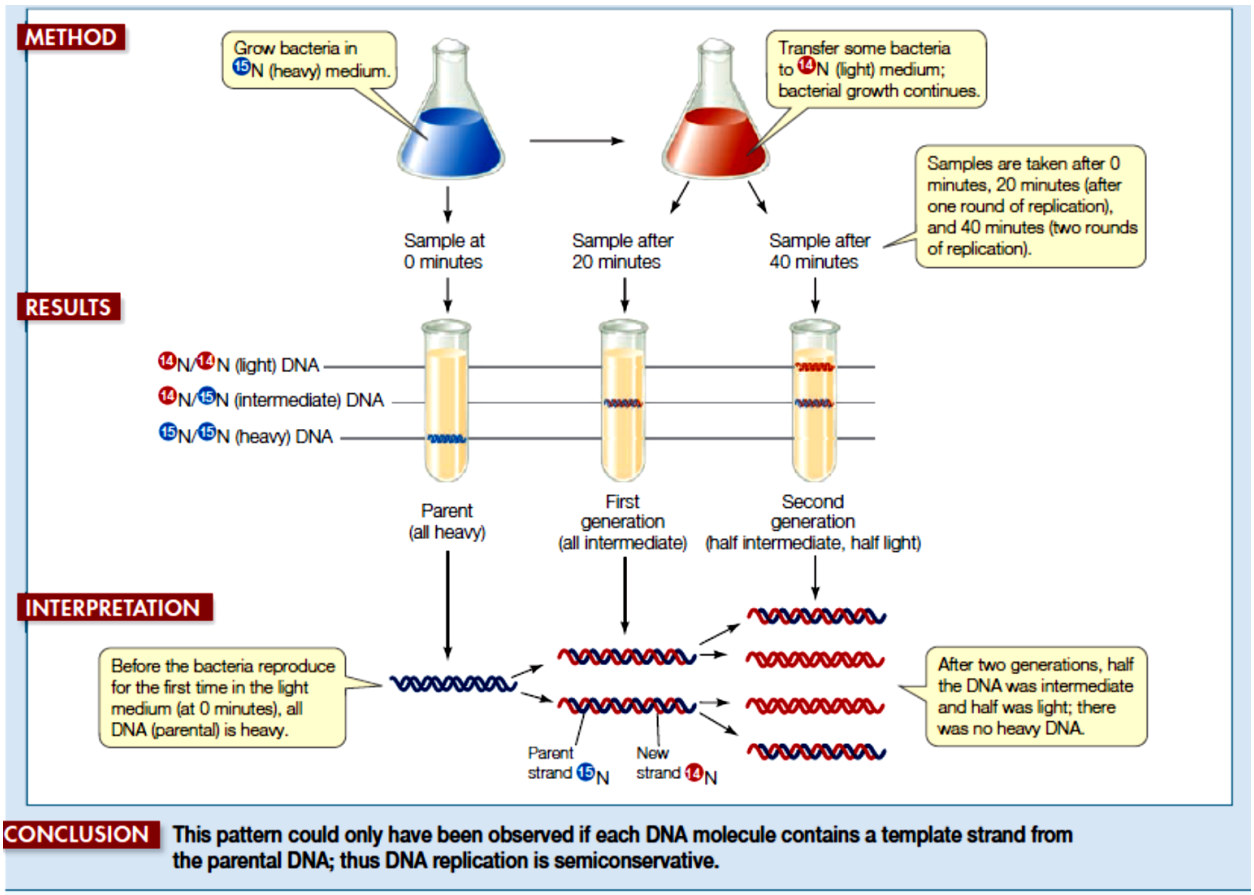


Figure 10- Meselson and Stahl experiment to prove that DNA replicates in a semiconservative manner. The researchers grew another E. coli culture on 15N medium, then transferred it to normal 14N medium and allowed the bacteria to continue growth. Image courtesy- Sadava et al, Life: The science of Biology, 9th edition.

The results of this experiment can be explained only by the semiconservative model of DNA replication. In the first round of DNA replication in the ^{14}N medium, the strands of the double helix—both heavy with ^{15}N—separated. Each strand then acted as the template for a second strand, which contained only ^{14}N and hence was less dense. Each double helix then consisted of one ^{15}N strand and one ^{14}N strand, and was of intermediate density. In the second replication, the ^{14}N-containing strands directed the synthesis of partners with ^{14}N, creating low-density DNA, and the ^{15}N strands formed new ^{14}N partners. The crucial observation demonstrating the semiconservative model was that intermediate-density DNA (^{15}N–^{14}N)

appeared in the first generation and continued to appear in subsequent generations. With the other models, the results would have been quite different:

• If conservative replication had occurred, the first generation would have had both high-density DNA (^{15}N–^{15}N) and low-density DNA (^{14}N–^{14}N), but no intermediate density DNA.

• If dispersive replication had occurred, the density of the new DNA would have been intermediate, but DNA of this density would not continue to appear in subsequent generations.

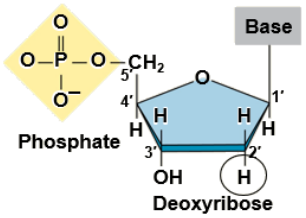


Figure 11- Monomer of DNA (nucleotide monophosphate). The monomers exist within the strand in monophosphate form.

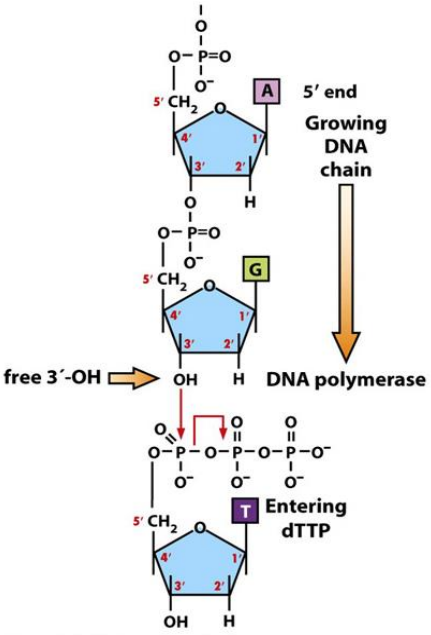


Figure 12- During the formation of a phosphodiester bond, the incoming monomers are in nucleotide triphosphate form. The cleavage of the beta and gamma phosphate provides the energy needed for phophodiester bond formation. This reaction is catalyzed by DNA polymerase.

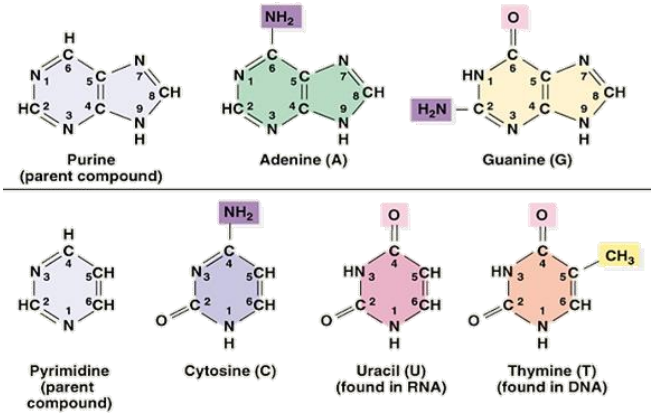


Figure 13- Nitrogenous bases of DNA and RNA. Uracil is present in RNA instead of Thymine.

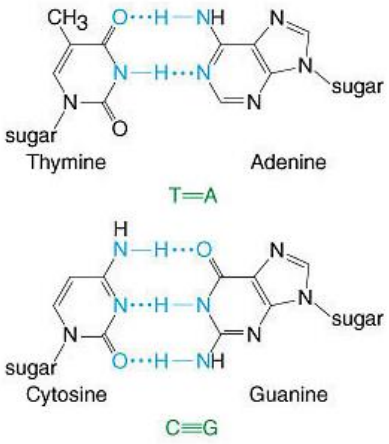


Figure 14- Hydrogen bonding between the nitrogenous bases of DNA.

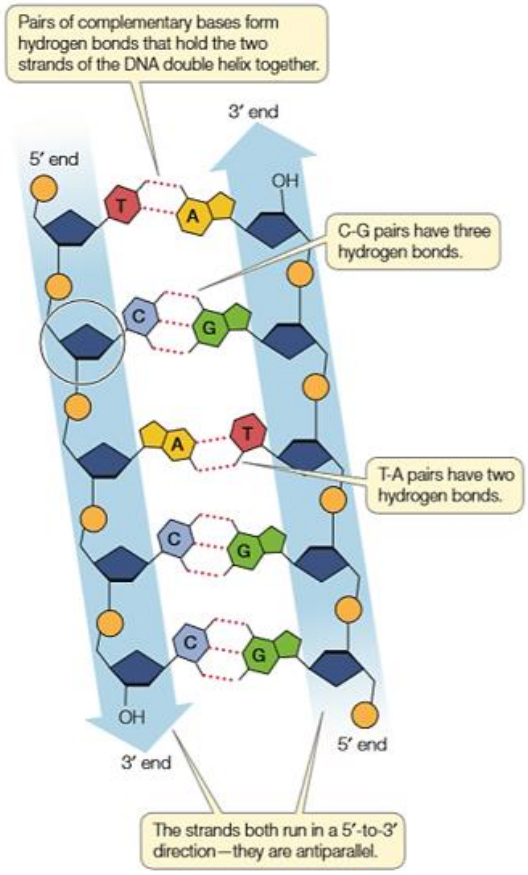


Figure 15- The purines (A and G) pair with the pyrimidines (T and C, respectively) to form base pairs that are equal in size and resemble the rungs on a ladder whose sides are formed by the sugar–phosphate backbones. The deoxyribose sugar (left) is where the 3′ and 5′ carbons are located. The two strands are antiparallel. Image courtesy- Sadava et al, Life: The science of Biology, 9th edition.

This concludes the chase for discovery of DNA structure. The next section describes the process of DNA replication.

Kornberg, Arthur, an American physician and biochemist, was the first to discover how monomers of deoxyribonucleic acid (DNA) duplicate within bacterial cells and also the first to devise a cell free method for replicating DNA. For these achievements he shared the 1959 Nobel Prize in physiology.

Arthur Kornberg isolated a new enzyme from E. coli, "DNA polymerase", which had the property to assemble nucleotides and manufacture DNA. The cell free setup for replicating DNA was developed *in vitro* (in a test tube) by providing a pool of free nucleotides, a DNA primer (he used calf thymus DNA), a source of magnesium ions, and ATP.

Kornberg used DNA polymerase to verify one of the essential elements of the Watson - Crick Model of DNA structure: DNA is always polymerized in the 5´ to 3´ direction (H-CH_2 sugar phosphate bonds to H-O sugar phosphate bond; new nucleotides are added at the 3´ end).

The findings of Kornberg experiment provided conclusive evidence that the double stranded helical model provided by Watson- Crick was accurate.

Kornberg designed the following experiment to determine the growing end for DNA replication process

The reaction used certain key ingredients as follows-

1. Labelled nucleotides (contained the radioactive phosphorus isotope ^{32}P)
2. DNA primer (he used calf thymus DNA),
3. a source of magnesium ions, and ATP

The procedure for the assignment is as follows:

Initiation of the reaction using unlabelled nucleotide as the nucleotide precursor.

Add radioactive (^{32}P) nucleotide for a brief period, and then quickly stop the reaction.

At the completion of the experiment, Kornberg observed that ^{32}P nucleotides attached only to the 3'-OH end of the growing strand.

The results of the experiment clearly indicated that the newly synthesized strand grows at the 3'-OH end.

DNA replication:

Semiconservative DNA replication in the cell involves a number of different enzymes and other proteins. It takes place in two general steps:

• The DNA double helix is unwound to separate the two template strands and make them available for new base pairing.

• As new nucleotides form complementary base pairs with template DNA, they are covalently linked together by phosphodiester bonds, forming a polymer whose base sequence is complementary to the bases in the template strand.

The process of replication occurs within a region called the replication bubble, which contains all the enzymes required for replication and the DNA strands to be replicated. The replication bubble consists of two replication forks, in which the synthesis of DNA proceeds in the two opposite directions.

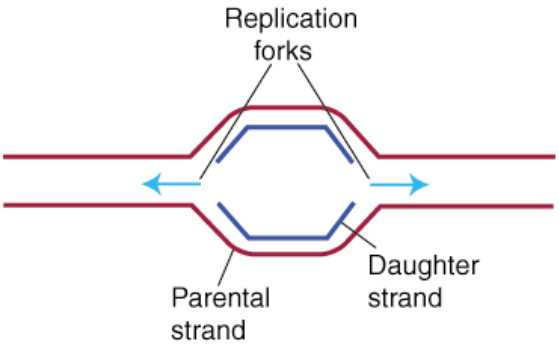


Figure 19- Replication bubble with two replication forks proceeding in opposite directions

Replication is carried out due to the efficient working of various enzymes involved in DNA replication.

DNA is replicated through the interaction of the template strand with a huge protein complex called the replication complex, which contains at least four proteins, including DNA polymerase. All chromosomes have at least one region called the origin of replication (ori), to which the replication complex binds with high specificity.

The first event at the origin of replication is the localized unwinding and separation (denaturation) of the DNA strands. There are several forces that hold the two strands together, including hydrogen bonding and the hydrophobic interactions of the bases. An enzyme called **DNA helicase** uses energy from ATP hydrolysis to unwind and separate the strands. Proteins called **single-strand binding proteins** bind to the unwound strands to keep them from reassociating into a double helix. This process makes each of the two template strands available for complementary base pairing.

Three types of DNA polymerase exist which assist in the process of DNA replication.

- DNA polymerase I
- DNA polymerase II
- DNA polymerase III

The function of these DNA polymerases is revealed as the process of DNA replication is studied further.

The first DNA polymerase to act in the replication complex is **DNA polymerase III**, whose function is to extend the new strand.

A DNA polymerase elongates a polynucleotide strand by covalently linking new nucleotides to a previously existing strand. However, it cannot start this process without a short "starter" strand, called a primer. In DNA replication, the primer is usually a short single strand of RNA (Figure 14). This RNA primer strand is complementary to the DNA template, and is synthesized one nucleotide at a time by an enzyme called a primase. The DNA polymerase III then adds nucleotides to the 3' end of the primer and continues until the replication of that section of DNA has been completed. Thus the basic function of DNA polymerase III is to extend the new strand starting from one end of the RNA primer.

DNA polymerase I (discovered by Arthur Kornberg) degrades the RNA primer, and adds DNA in its place.

This action of DNA polymerase I causes the resulting new strand to have a small gap. This gap is filled in by DNA ligases. When DNA replication is complete, each new strand consists only of DNA.

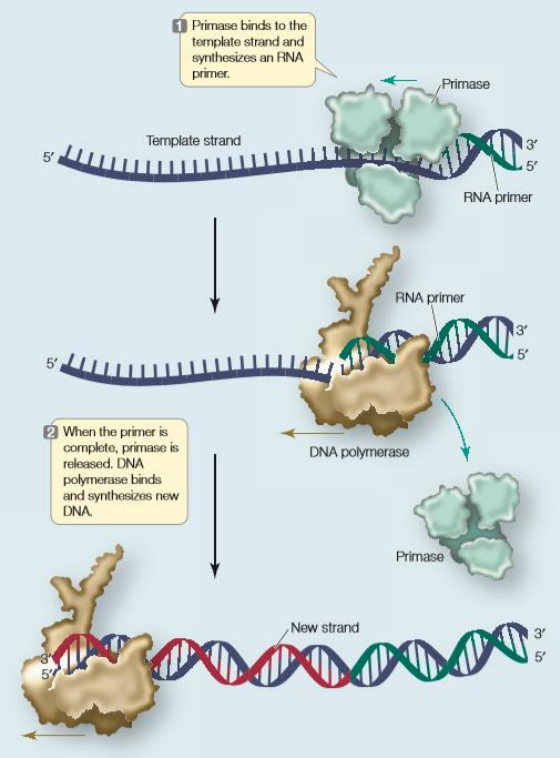


Figure 20- DNA polymerase requires a primer to initiate replication. Primase is the enzyme which provides an RNA primer. DNA polymerase attaches nucleotides to the end of this RNA primer and extends the new strand. Note: the new strand is DNA in nature and not RNA. Image courtesy- Sadava et al, Life: The science of Biology, 9th edition.

The two daughter strands resulting from a double stranded parent grow differently. The DNA double helix is antiparallel in nature, therefore:

• One newly replicating strand (the leading strand) is oriented so that it can grow continuously at its 3'end as the fork opens up.

• The other new strand (the lagging strand) is oriented so that as the fork opens up, its exposed 3'end gets farther and farther away from the fork, and an unreplicated gap is formed. This gap would get bigger and bigger if there were not a special mechanism to overcome this problem.

Synthesis of the lagging strand requires the synthesis of relatively small, discontinuous stretches of sequence. These discontinuous stretches are synthesized just as the leading strand is, by the addition of new nucleotides one at a time to the 3' end of the new strand, but the synthesis of this new strand moves in the direction opposite to that in which the replication fork is moving. These stretches of new DNA are called Okazaki fragments (after their discoverer, the Japanese biochemist Reiji Okazaki). While the leading strand grows continuously "forward," the lagging strand grows in shorter, "backward" stretches with gaps between them. A single primer is needed for synthesis of the leading strand, but each Okazaki fragment requires its own primer to be synthesized by the primase. DNA polymerase III then synthesizes an Okazaki fragment by adding nucleotides to one primer until it reaches the primer of the previous fragment. At this point, DNA polymerase I removes the old primer and replaces it with DNA. Left behind is a tiny nick—the final phosphodiester linkage between the adjacent Okazaki fragments is missing. The enzyme DNA ligase catalyzes the formation of that bond, linking the fragments and making the lagging strand whole.

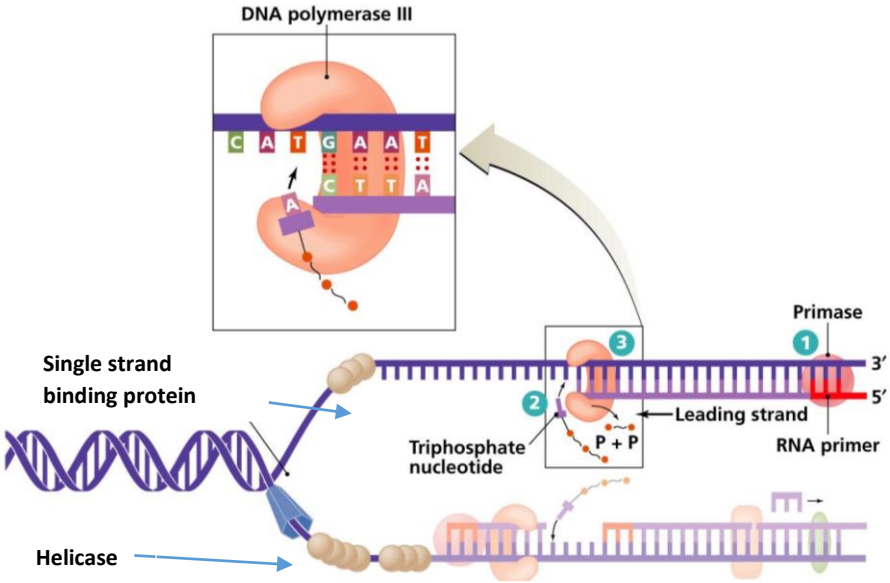


Figure 21- Synthesis of the leading strand. Image courtesy- Biology, by Campbell et al.

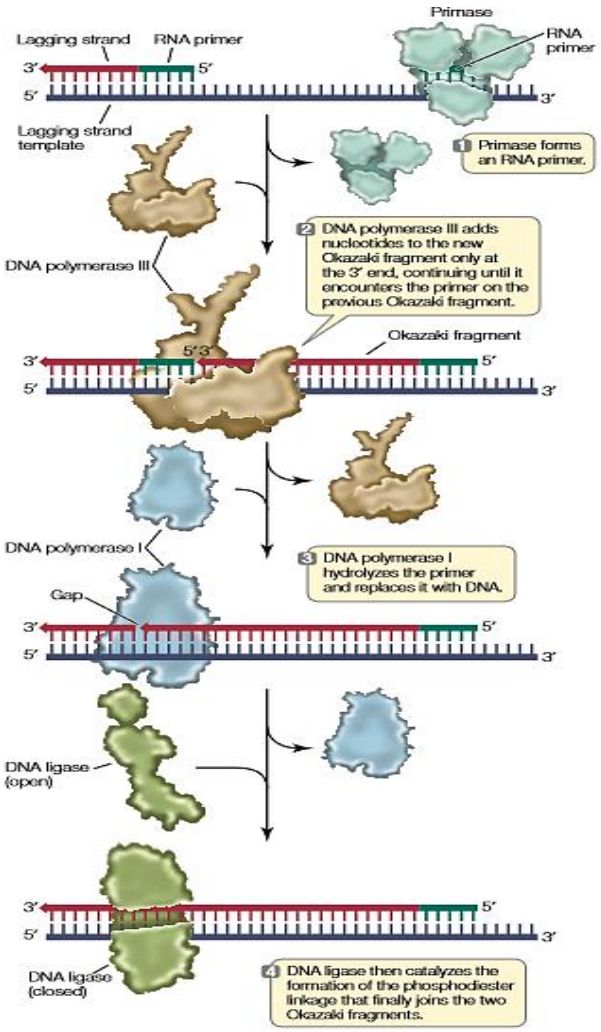


Figure 22- Synthesis of the lagging strand. Image courtesy- Sadava et al, Life: The science of Biology, 9th edition.

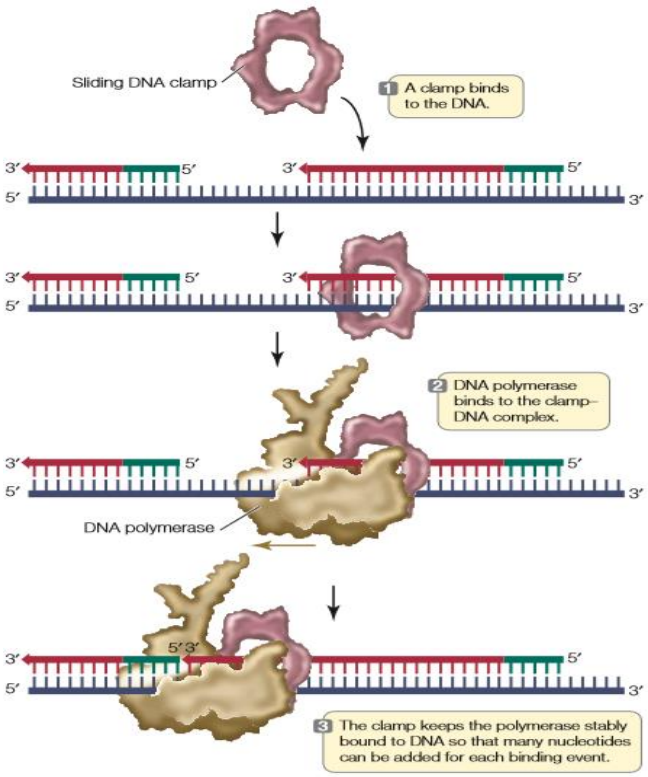


Figure 23- Sliding DNA clamp increases the efficiency of polymerization by keeping the enzyme bound to the substrate, so the enzyme does not have to repeatedly bind to template and substrate. Image courtesy- Sadava et al, Life: The science of Biology, 9th edition.

During DNA replication, the progress of the replication fork generates positive supercoils ahead of the replication machinery and negative supercoils behind it. The DNA can be supercoiled to such an extent that if left unchecked it could impede the progress of the protein machinery involved. This is prevented by **DNA topoisomerase**, which makes single-stranded nicks to relax the helix.

Small circular chromosomes, such as those of bacteria (consisting of 1–4 million base pairs), have a single origin of replication. Two replication forks form at this ori, and as the DNA moves through the replication complex, the replication forks extend around the circle. Two interlocking circular DNA molecules are formed, and they are separated by an enzyme called DNA topoisomerase. DNA polymerases are very fast. In E. coli, replication can be as fast as 1,000 bases per second, and it takes 20–40 minutes to replicate the bacterium's 4.7 million base pairs.

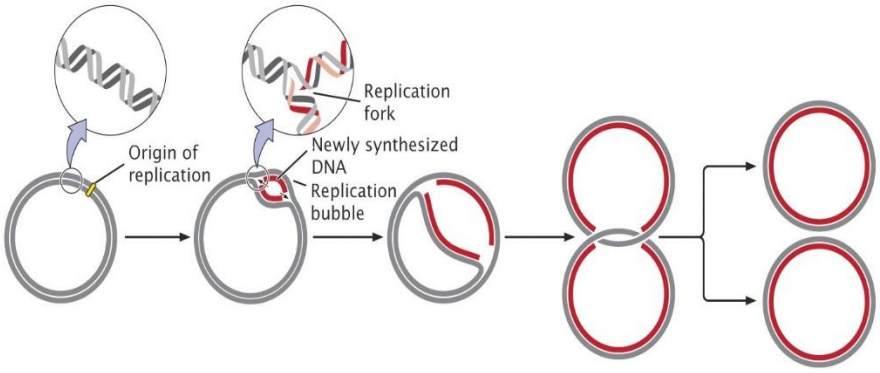
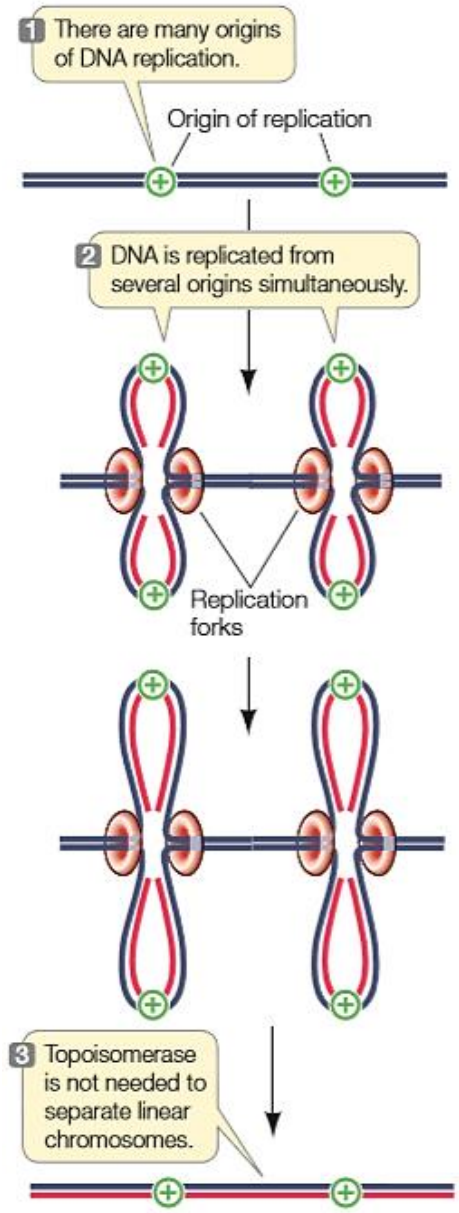


Figure 24- Replication in a bacteria having circular DNA.



Human DNA polymerases are slower than those of E. coli, and can replicate DNA at a rate of about 50 bases per second. Human chromosomes are much larger than those of bacteria (about 80 million base pairs) and linear. Large linear chromosomes such as those of humans contain hundreds of origins of replication. Numerous replication complexes bind to these sites at the same time and catalyze simultaneous replication. Thus there are many replication forks in eukaryotic DNA.

Figure 25- Larger linear chromosomes, typical of nuclear DNA in eukaryotes, have many origins of replication. Image courtesy- Sadava et al, Life: The science of Biology, 9th edition.

Activity of telomerase:

Replication of the lagging strand occurs by the addition of Okazaki fragments to RNA primers. When the terminal RNA primer is removed, no DNA can be synthesized to replace it because there is no 3'end to extend. So the new chromosome has a bit of single-stranded DNA at each end. This situation activates a mechanism for cutting off the single stranded region, along with some of the intact double-stranded DNA. This is a part of DNA repair mechanism. Thus the chromosome becomes slightly shorter with each cell division.

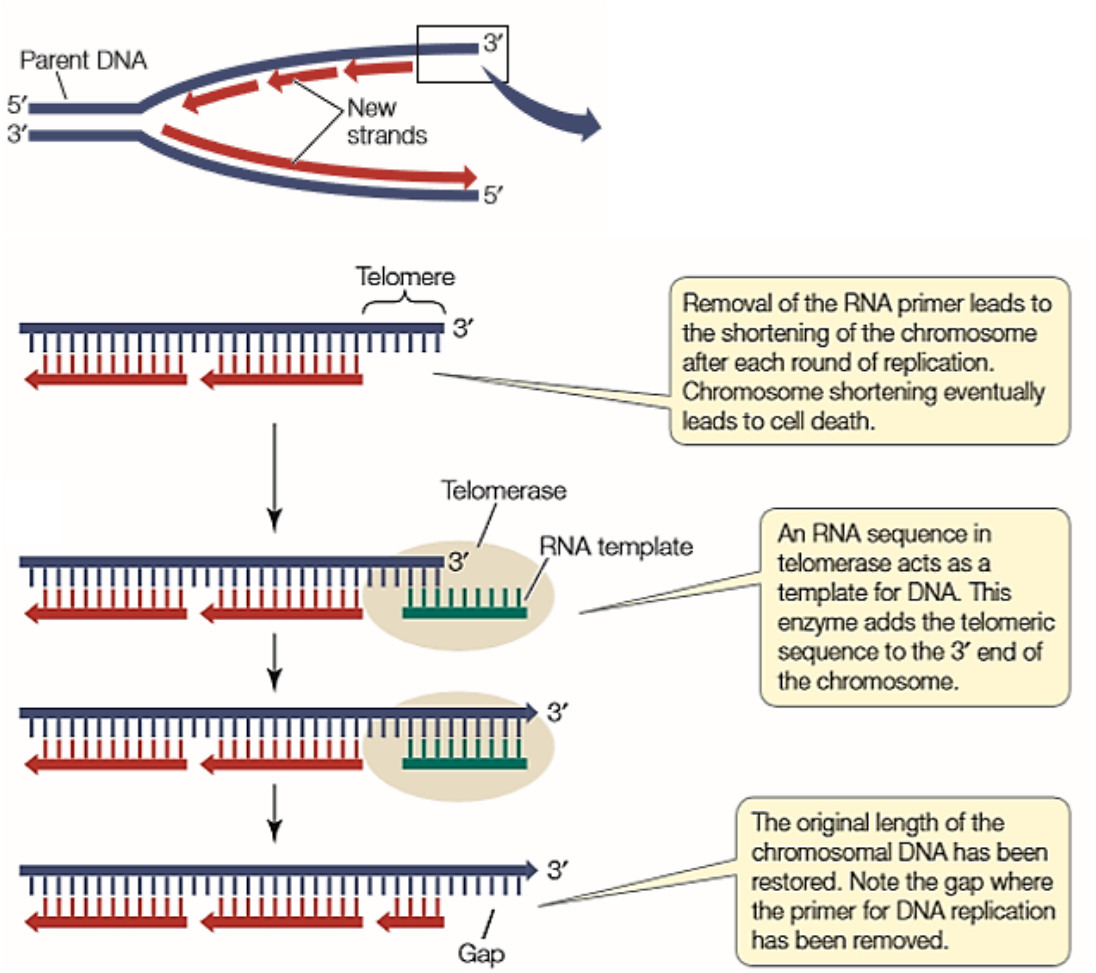


Figure 26- Removal of the RNA primer at the 3' end of the template for the lagging strand leaves a region of DNA— the telomere— unreplicated. In continuously dividing cells, the enzyme telomerase binds to the 3' end and extends the lagging strand of DNA, so the chromosome does not get shorter. Image courtesy- Sadava et al, Life: The science of Biology, 9th edition.

An enzyme, appropriately called telomerase, catalyzes the addition of any lost telomeric sequences. Telomerase contains an RNA sequence that acts as a template for the telomeric DNA repeat sequence.

DNA Replication proofreading:

DNA must be accurately replicated and faithfully maintained. The price of failure can be great; the accurate transmission of genetic information is essential for the functioning and even the life of a single cell or multicellular organism. Yet the replication of DNA is not perfectly accurate. So the problem arises in preserving life.

DNA repair mechanisms help to preserve life. DNA polymerases initially make significant numbers of mistakes in assembling polynucleotide strands. Without DNA repair, the observed error rate of one for every 10^5 bases replicated would result in about 60,000 mutations every time a human cell divided. Fortunately, our cells can repair damaged nucleotides and DNA replication errors, so that very few errors end up in the replicated DNA.

Cells have at least three DNA repair mechanisms at their disposal:

•A **proofreading** mechanism corrects errors in replication as DNA polymerase makes them.

•A **mismatch repair** mechanism scans DNA immediately after it has been replicated and corrects any base-pairing mismatches.

• An excision repair mechanism removes abnormal bases that have formed because of chemical damage and replaces them with functional bases.

Most DNA polymerases perform a proofreading function each time they introduce a new nucleotide into a growing DNA strand. When a DNA polymerase recognizes a mispairing of bases, it removes the improperly introduced nucleotide and tries again. The error rate for this process is only about 1 in 10,000 repaired base pairs, and it lowers the overall error rate for replication to about one error in every 10^{10} bases replicated.

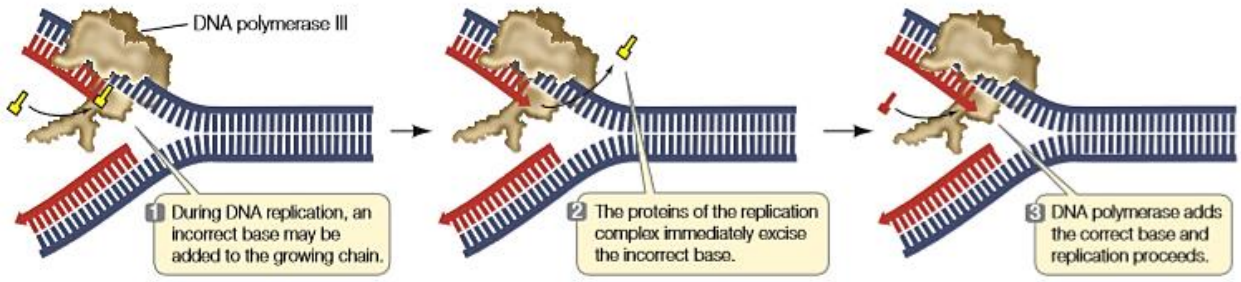


Figure 27- Proofreading activity during DNA replication. Image courtesy- Sadava et al, Life: The science of Biology, 9th edition.

After the DNA has been replicated, a second set of proteins surveys the newly replicated molecule and looks for mismatched base pairs that were missed in proofreading. **Mismatch repair** mechanism might detect an A-C base pair instead of an A-T pair. The repair mechanism "knows" whether the A-C pair should be repaired by removing the C and replacing it with T or by removing the A and replacing it with G (it can detect the "wrong" base)because a DNA strand is chemically modified some time after replication. In prokaryotes, methyl groups (—CH_3) are added to some adenines. In eukaryotes, cytosine bases are methylated. Immediately after replication, methylation has not yet occurred on the newly replicated strand, so the new strand is "marked" (distinguished by being unmethylated) as the one in which errors

should be corrected. When mismatch repair fails, DNA sequences are altered. One form of colon cancer arises in part from a failure of mismatch repair.

Excision repair mechanisms deal with damage due to high-energy radiation, chemicals from the environment, and random spontaneous chemical reactions. Individuals who suffer from a condition known as xeroderma pigmentosum lack an excision repair mechanism that normally corrects the damage caused by ultraviolet radiation. They can develop skin cancers after even a brief exposure to sunlight.

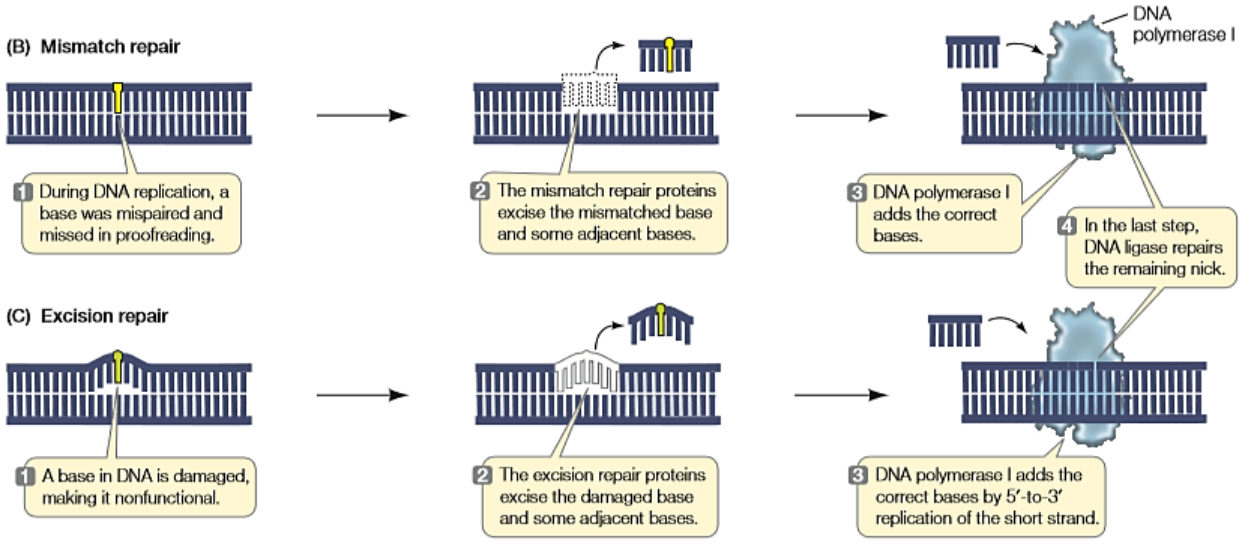


Figure 28- DNA repair mechanisms. Image courtesy- Sadava et al, Life: The science of Biology, 9th edition.