DNA

The Search to Identify the Genetic Material

Discovery of Nucleic Acids - Friedrich Miescher, 1869

Miescher isolated the nuclei of white blood cells obtained from pus cells. His experiments revealed that nuclei contained a chemical that contained nitrogen and phosphorus but no sulfur. He called the chemical nuclein because it came from nuclei. It later became known as nucleic acid.

Proteins Produce Genetic Traits - Archibald Garrod, 1909

Garrod noticed that people with certain genetic (inherited) abnormalities lacked certain enzymes. This observation linked proteins (enzymes) to genetic traits.

Genetic Material can Transform Bacteria - Frederick Griffith, 1931

When Streptococcus pneumoniae (pneumococcus) bacteria are grown on a culture plate, some produce smooth shiny colonies (S) while others produce rough colonies (R). This is because the S strain bacteria have a mucous (polysaccharide) coat, while R strain does not.

Mice infected with the S strain die from pneumonia infection but mice infected with the R strain do not develop pneumonia.

S strain → Inject into mice → Mice die

R strain → Inject into mice → Mice live

Griffith was able to kill bacteria by heating them. He observed that heat-killed S strain bacteria injected into mice did not kill them. When he injected a mixture of heat-killed S and live R bacteria, the mice died. Moreover, he recovered living S bacteria from the carcasses.

S strain (heat killed) → Inject into mice → Mice live

S strain (heat killed) + R strain (live) → Inject into mice → Mice die

He concluded that some substance needed to produce the mucous coat was passed from the dead bacteria (S strain) to the live ones (R strain); they became ***transformed***.

This must be due to a change in the genotype associated with the transfer of the genetic material.

The transforming material is DNA - Oswald Avery, Colin MacLeod, and Maclyn McCarty,

1944

Prior to the work of Avery, MacLeod, and McCarty, the genetic material was thought to be protein. Avery, MacLeod, and McCarty worked to determine what the transforming substance was in Griffith's experiment (above).

They purified chemicals from the heat-killed S cells to see which ones could transform live R cells into S cells. They discovered that DNA alone from S bacteria caused R bacteria to become transformed.

They also discovered that protein-digesting enzymes (proteases) and RNA-digesting enzymes did not affect transformation, so the transforming substance was not a protein or RNA. Digestion with a DNA-digesting enzyme did inhibit transformation, so DNA caused transformation.

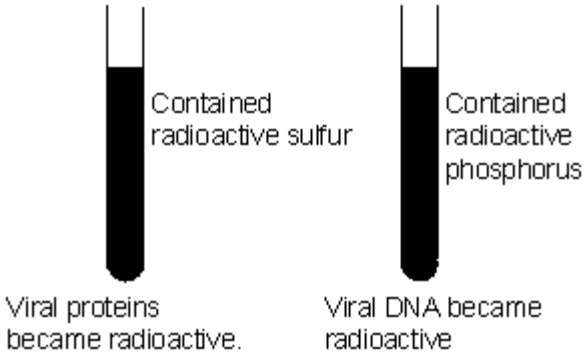
They concluded that DNA is the hereditary material, but not all biologists were convinced.

More Evidence: The Genetic Material is DNA - Alfred D. Hershey and Martha Chase, 1952

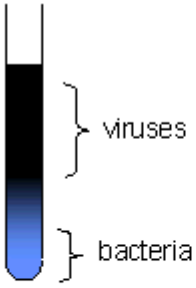
Hershey and Chase worked with viruses that infect bacteria called bacteriophages.

The bacteriophage becomes attached to the bacteria and its genetic material then enters the bacterial cell. The bacterial cell treats the viral genetic material as if it was its own and subsequently manufactures more virus particles. Hershey and Chase worked to discover whether it was protein or DNA from the viruses that entered the bacteria.

They grew a virus population in medium that contained radioactive phosphorus and another in medium that contained radioactive sulfur. Viruses grown in the presence of radioactive phosphorus contained radioactive DNA but not radioactive protein because DNA contains phosphorus but protein does not. Similarly, viruses grown on radioactive sulfur contained radioactive protein but not radioactive DNA because DNA does not contain sulfur.



Radioactive bacteriophages were allowed to attach to E. coli bacteria. Then as the infection proceeded, the viral coats were removed from the bacteria by agitating them in a blender. The viruses particles were separated from the bacteria by spinning them in a centrifuge.



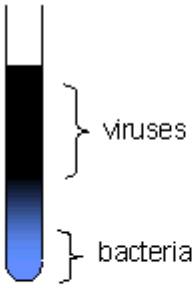
These viruses contained radioactive protein (sulfur).

The bacteria were not radioactive.

These viruses contained radioactive DNA (phosphorus).

The bacteria were radioactive.

Bacteria that were infected with viruses that had radioactive DNA were radioactive, indicating that DNA was the material that passed from the virus to the bacteria. Bacteria that were infected with viruses that had radioactive proteins were not radioactive. This indicates that proteins did not enter the bacteria from the viruses. DNA is therefore the genetic material that is passed from virus to bacteria.



These viruses contained radioactive protein (sulfur).

The bacteria were not radioactive.

These viruses contained radioactive DNA (phosphorus).

The bacteria were radioactive.

Discovery of the Structure of DNA

Erwin Chargaff, 1940's and early 50's

DNA was thought to contain equal amounts of A, T, T, and C. Chargaff found that the base composition of DNA differs among species.

His data showed that in each species, the percent of A equals the percent of T, and the percent of G equals the percent of C. so that 50% of the bases were purines (A + G) and 50% were pyrimidines (T + C)

Chargafff's rules: 1) The percent of each base varies from species to species. 2) Within a species, the amount of A = T and the amount of G = C.

M.H.F. Wilkins and Rosalind Franklin, early 50's

Wilkins and Franklin studied the structure of DNA crystals using X-rays.

They found that the crystals contain regularly repeating subunits.

Structures that are close together cause the x-ray to bend more than structures that are further apart. The X pattern produced by DNA suggested that DNA contains structures with dimensions of 2 nm, 0.34 nm, and 3.4 n. The dark structures at the top and bottom of their X-ray photograph of DNA indicate that DNA contains repeating units, suggesting a helix.

James Watson and Francis H.C. Crick, 1953

Watson and Crick used Chargaff's base data and Franklin's X-ray diffraction data to construct a model of DNA.

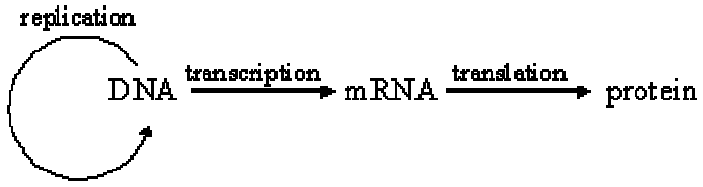
The model showed that DNA is a double helix with sugar-phosphate backbones on the outside and the paired nucleotide bases on the inside, in a structure that fit the spacing estimates from the X-ray diffraction data.

Chargaff's rules showed that A = T and G = C, so there was complementary base pairing of a purine with a pyrimidine, giving the correct width for the helix.

The paired bases can occur in any order, giving an overwhelming diversity of sequences.

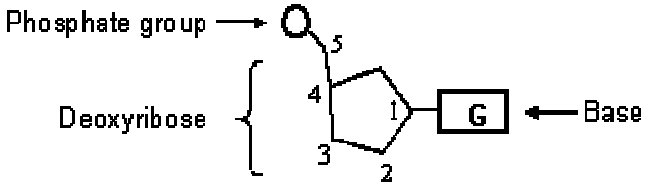
Properties of Genetic Material

DNA is an ideal genetic material because it can store information, is able to replicate, and is able to undergo changes (mutate).

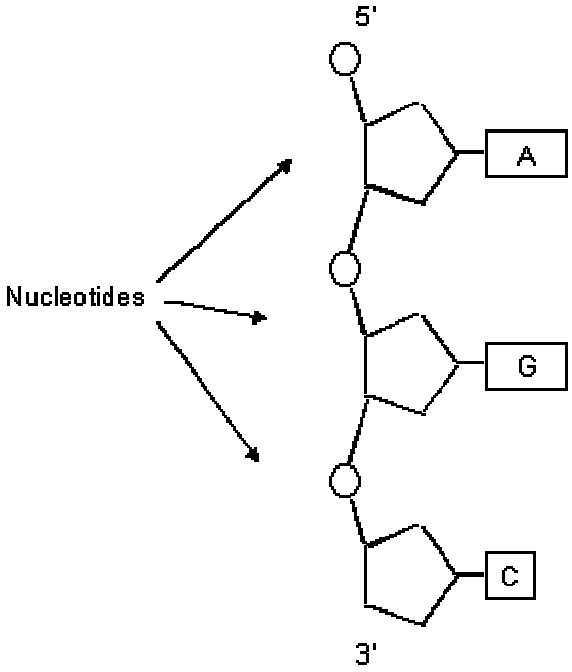


Structure of DNA

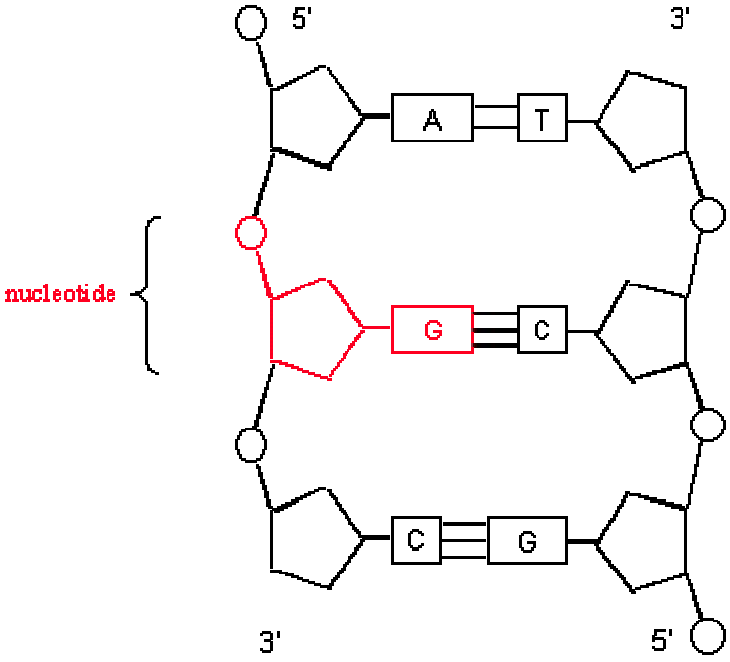
DNA is composed of units called nucleotides. Each nucleotide contains a phosphate group, a deoxyribose sugar, and a nitrogenous base.



The nucleotides joined together to form a chain. The phosphate end of the chain is referred to as the 5' end. The opposite end is the 3' end.



DNA is composed of two chains of nucleotides linked together in a ladder-like arrangement with the sides composed of alternating deoxyribose sugar and phosphate groups and the rungs being the nitrogenous bases as indicated by the diagram below.



The "A" of one strand is always paired with a "T" on the other. Similarly, the "G" of one strand is paired with a "C" on the other.

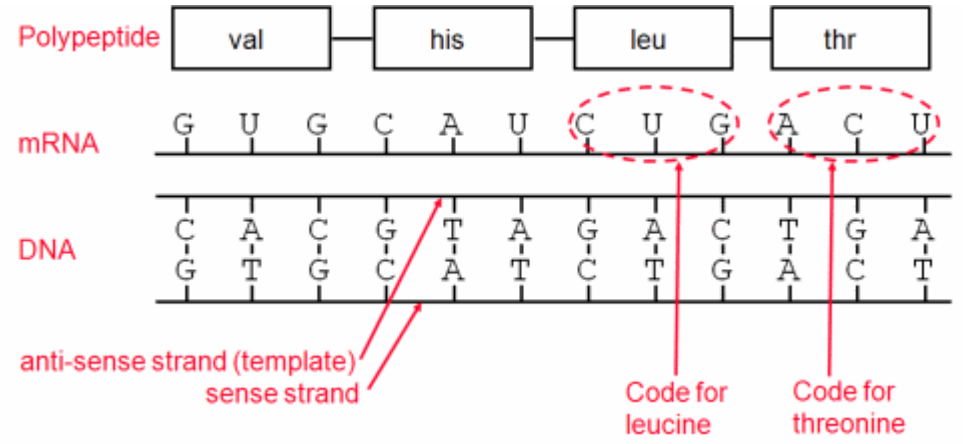
The two strands are held together by hydrogen bonds (electrostatic attraction). Two hydrogen bonds hold adenine to thymine. Three bonds attach cytosine to guanine as indicated in the diagram above.

During the process of cell division, the DNA and associated proteins become tightly coiled producing visible chromosomes.

How is Information Stored?

The diagram below shows that one strand of the DNA double-helix serves as a template for the construction of mRNA. The sequence of nucleotides in this DNA strand is complimentary (opposite) the sequence in mRNA. The diagram also shows that the sequence of nucleotides in mRNA determines the amino acids in the protein. For example GUG in mRNA (or CAC in DNA) codes for valine (see below).

The strand of DNA that contains the genetic code is called the anti-sense strand. It is often referred to as the template strand. The other strand (the sense strand) is not used. Notice that the sense strand has the same base sequence as mRNA except that mRNA has U instead of T.



The codes in DNA are copied to produce mRNA. Each three-letter code in mRNA (called a ***codon***) codes for one amino acid. The sequence of amino acids in proteins is therefore most directly determined by the sequence of bases in mRNA, which in turn, are determined by the sequence of bases in DNA.

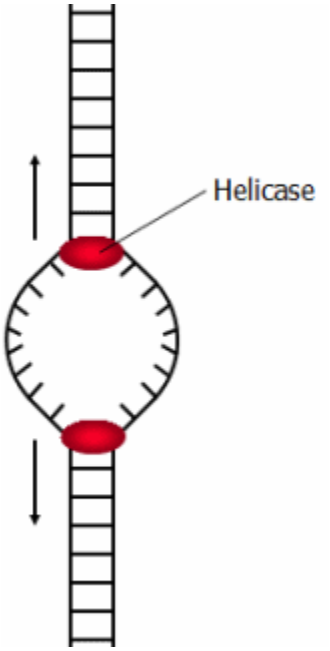
There are four letters in the genetic alphabet (A, T, G, and C) and each codon contains three letters. It is therefore possible to have 64 different codons. Because there are only 20 different amino acids and 64 possible codons, some amino acids have several different codons.

Terminators are codes that indicate the end of a genetic message (gene).

An initiator codon (usually AUG) indicates where the genetic information begins.

DNA replication

A ***replication bubble*** forms at a point in the DNA called the ***origin of replication***. Eukaryotic cells have thousands of origins of replication but prokaryotic cells have only one.

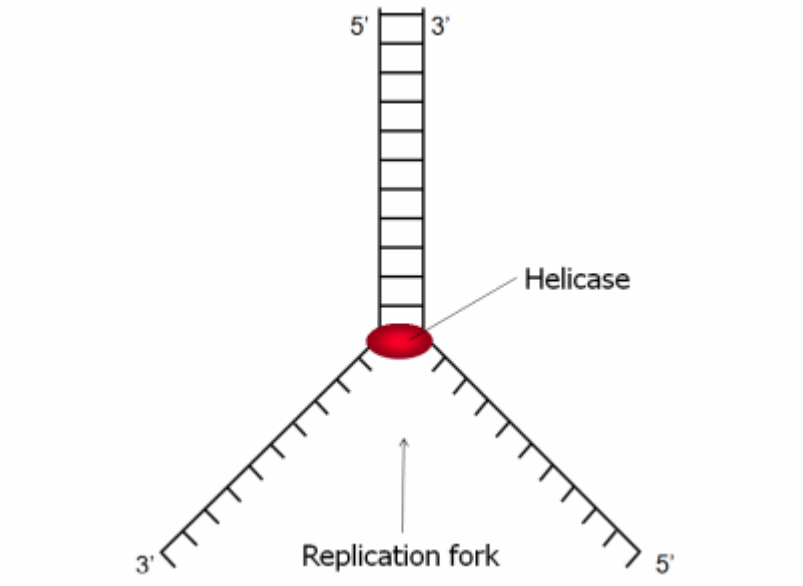


The enyme ***helicase*** unwinds the DNA molecule by breaking hydrogen bonds between the bases.

Because DNA is coiled, as helicase advances, the DNA in front of it becomes more tightly twisted. Topoisomerases cut the DNA to relieve the excess winding and then rejoin the cut pieces.

DNA synthesis occurs at both ends of the replication bubble. However, the remaining diagrams show only half of this replication bubble.

In the diagrams, the top half of the replication bubble looks like an upside down "Y". This area is called a ***replication fork***. The other side of the replication bubble also has a replication fork but this is not shown in the diagrams below.



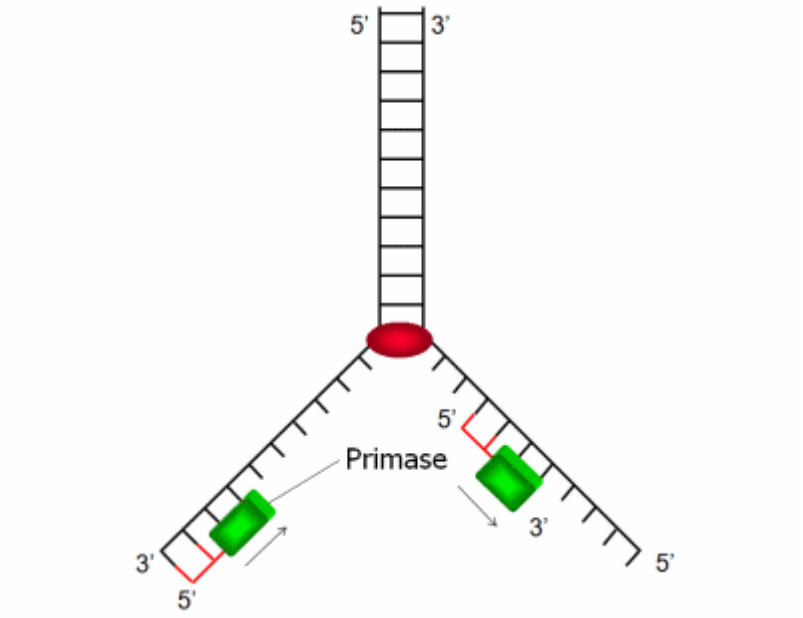
Single-strand binding proteins attach to the separated strands and keep them separate. The

separated strands serve as templates for the synthesis of new DNA.

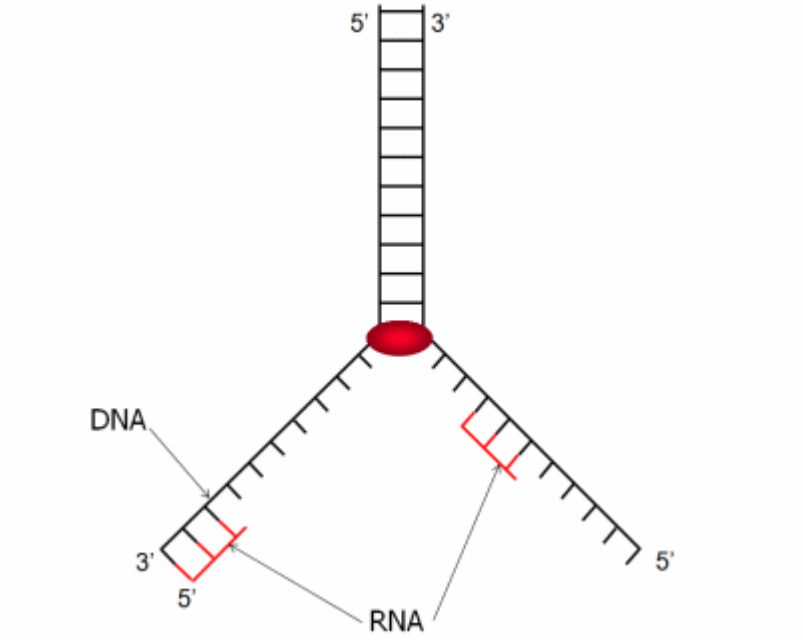
DNA polymerase creates a new strand by pairing complimentary bases to the separated strands. However, it cannot initiate a new strand; it functions to lengthen an existing strand. Another enzyme,***primase***, must first add a short segment of nucleotides called a ***primer*** to the 3' end of the template strand. The primer is composed of RNA nucleotides.

Both Primase and DNA Polymerase add nucleotides to the 3' end of the new strand. The direction of synthesis is therefore 5' to 3'.

The red lines in the diagram represent the RNA primer.



The primers are approximately 10 nucleotides in length and are complimentary to the template strand. The RNA nucleotides containing A, U, G, and C are bonded to DNA nucleotides containing T, A, C, and G respectively.

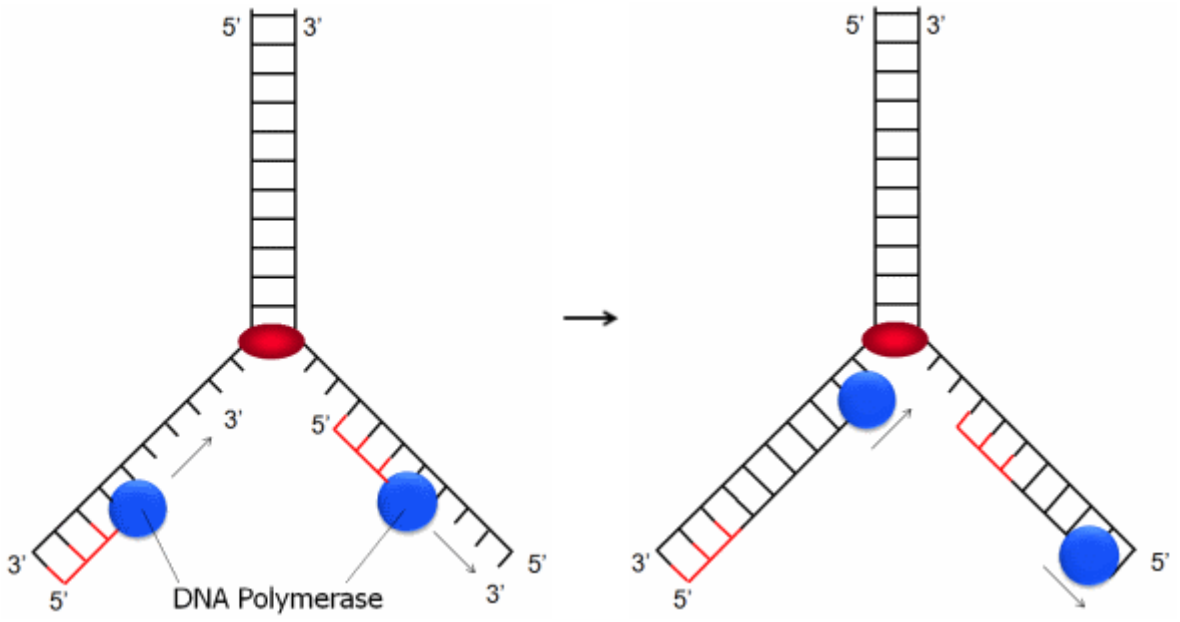


After the primer is created, DNA polymerase III attaches and continues to elongate the new strand beginning at the primer and progressing along the template strand as it adds complimentary nucleotides- A with T and G with C.

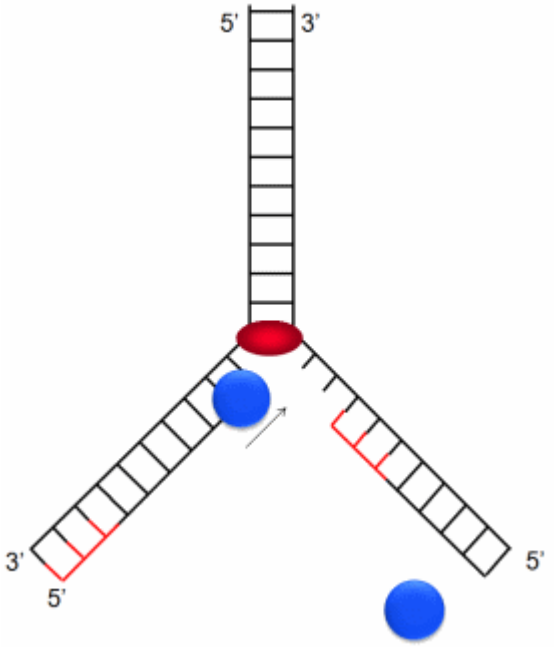
The direction of synthesis- 5' to 3'.

Clamp proteins are involved in attaching DNA polymerase to the template strand. They circle the DNA and slide along with polymerase as replication proceeds.

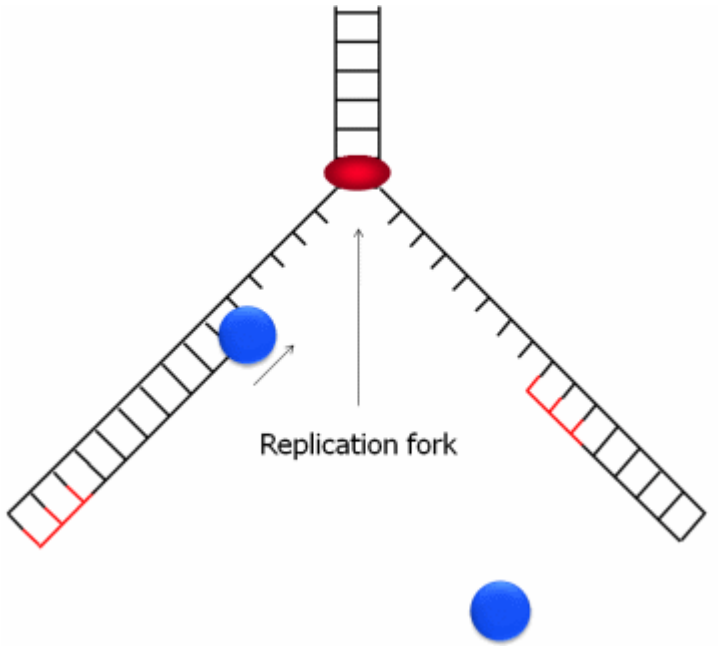
DNA polymerase proofreads the new strand as it is being synthesized. Incorrectly paired bases are removed and the correct one is inserted (discussed later).



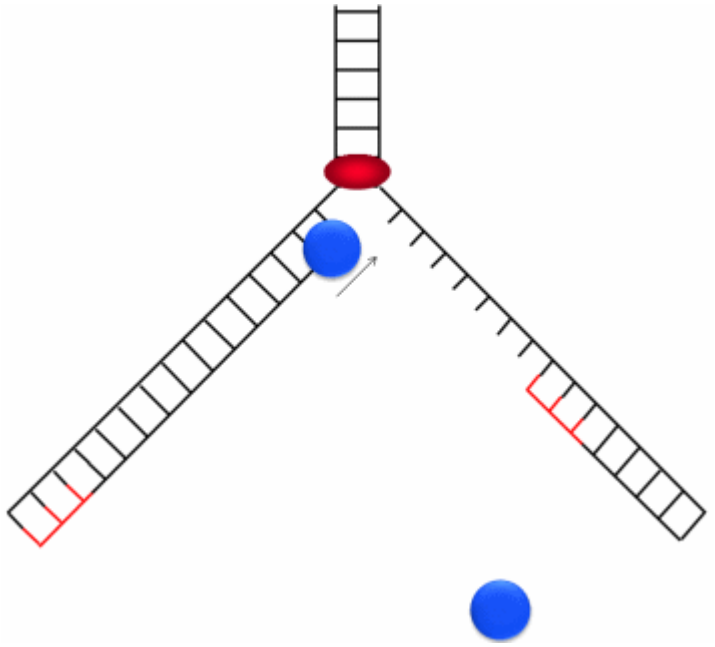
The DNA polymerase molecule on the right cannot proceed further.



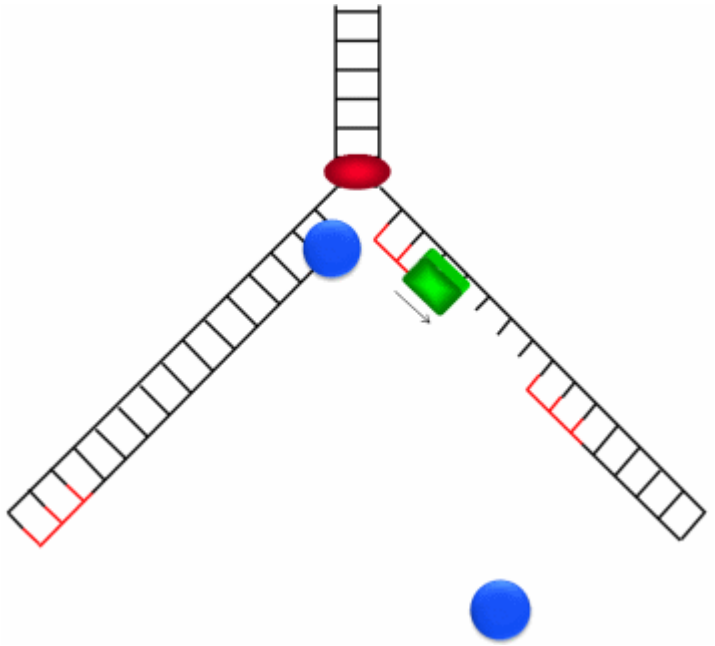
This diagram shows that helicase continues to unwind the DNA.



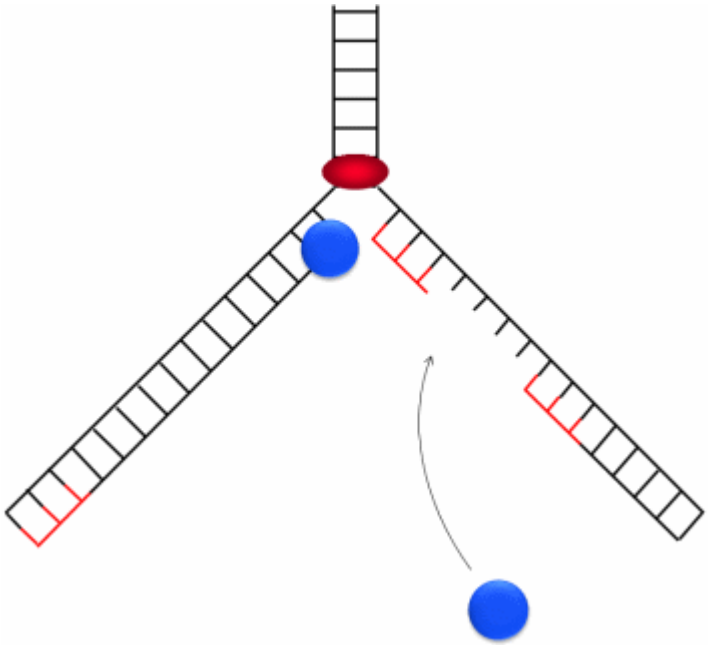
DNA polymerase on the left side is able to continue lengthening the strand continuously.



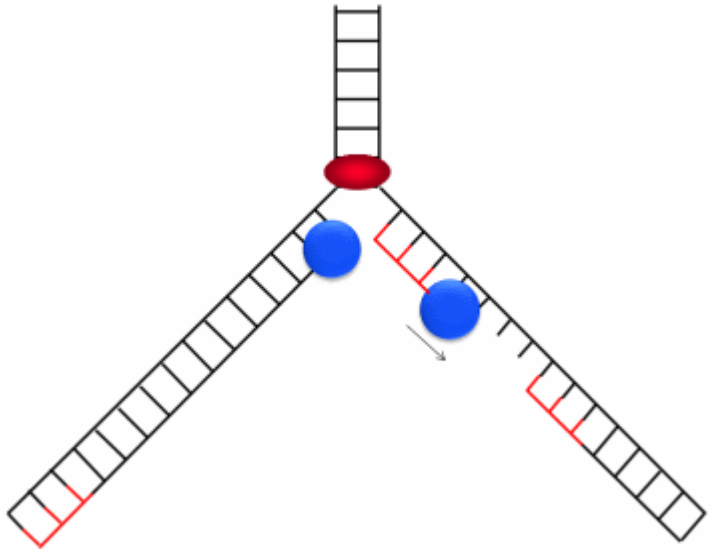
DNA polymerase cannot synthesize DNA on the right side of the diagram because it needs a primer. Recall that DNA polymerase can lengthen a strand but it cannot initiate synthesis. A new RNA primer is synthesized by primase, indicated by the green structure in the diagram.



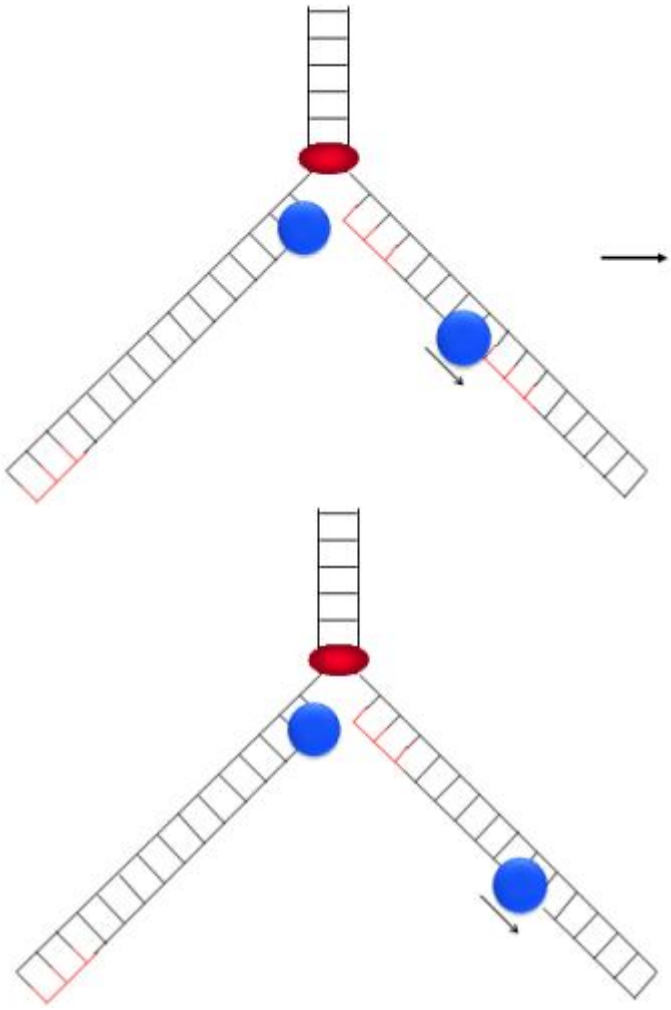
DNA polymerase can now attach and add nucleotides.



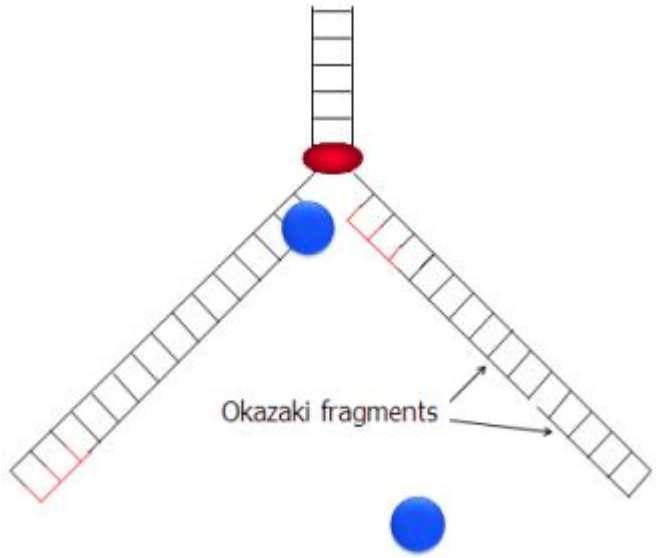
After the primer is synthesized (red lines), DNA polymerase can lengthen the strand.



Another molecule of DNA polymerase, called DNA polymerase I replaces DNA polymerase III. It removes the RNA primer and replaces it with DNA.



Notice that there is a gap between the two DNA strands where the primer was removed. DNA polymerase does not form the final bond needed to attach the two fragments.

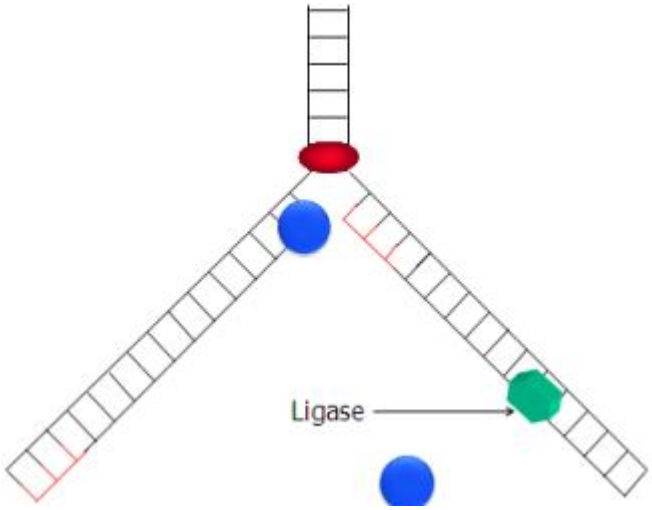


Because the direction of synthesis is from 5' to 3', the strand on the left in the diagram is synthesized continuously but the strand on the right is synthesized in fragments. The strand that is synthesized continuously is called the ***leading strand***. The strand that is synthesized in

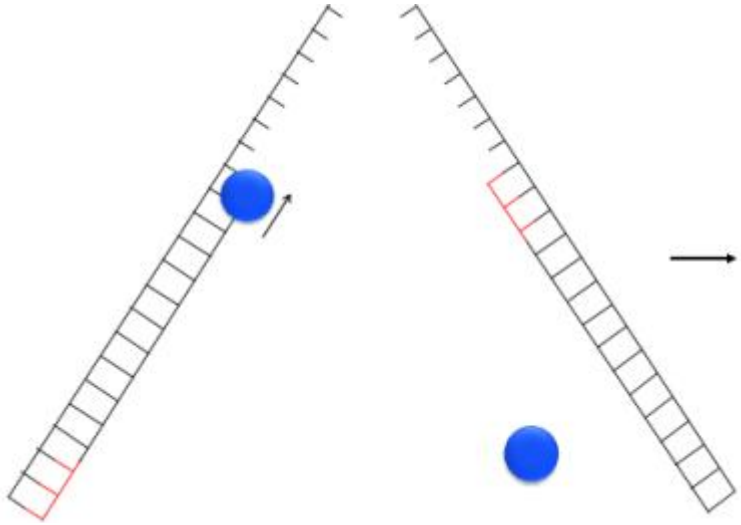
fragments is called the ***lagging strand***.

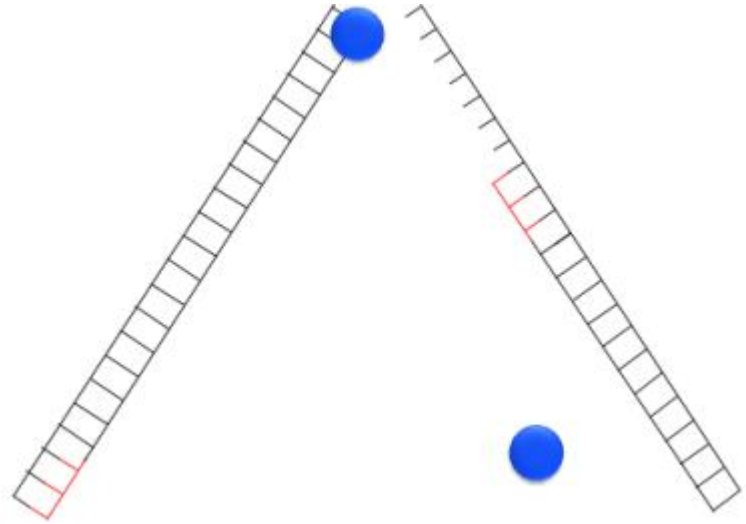
The fragments are called ***Okazaki fragments***.

Ligase catalyzes the formation of covalent bonds between the Okazaki fragments.

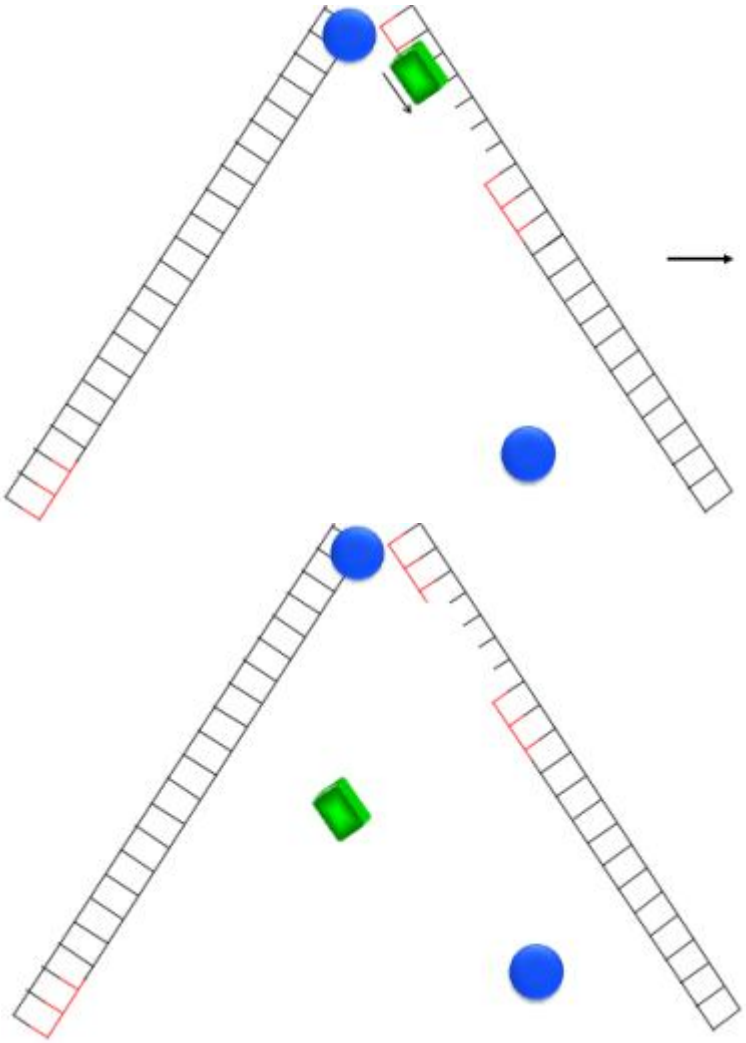


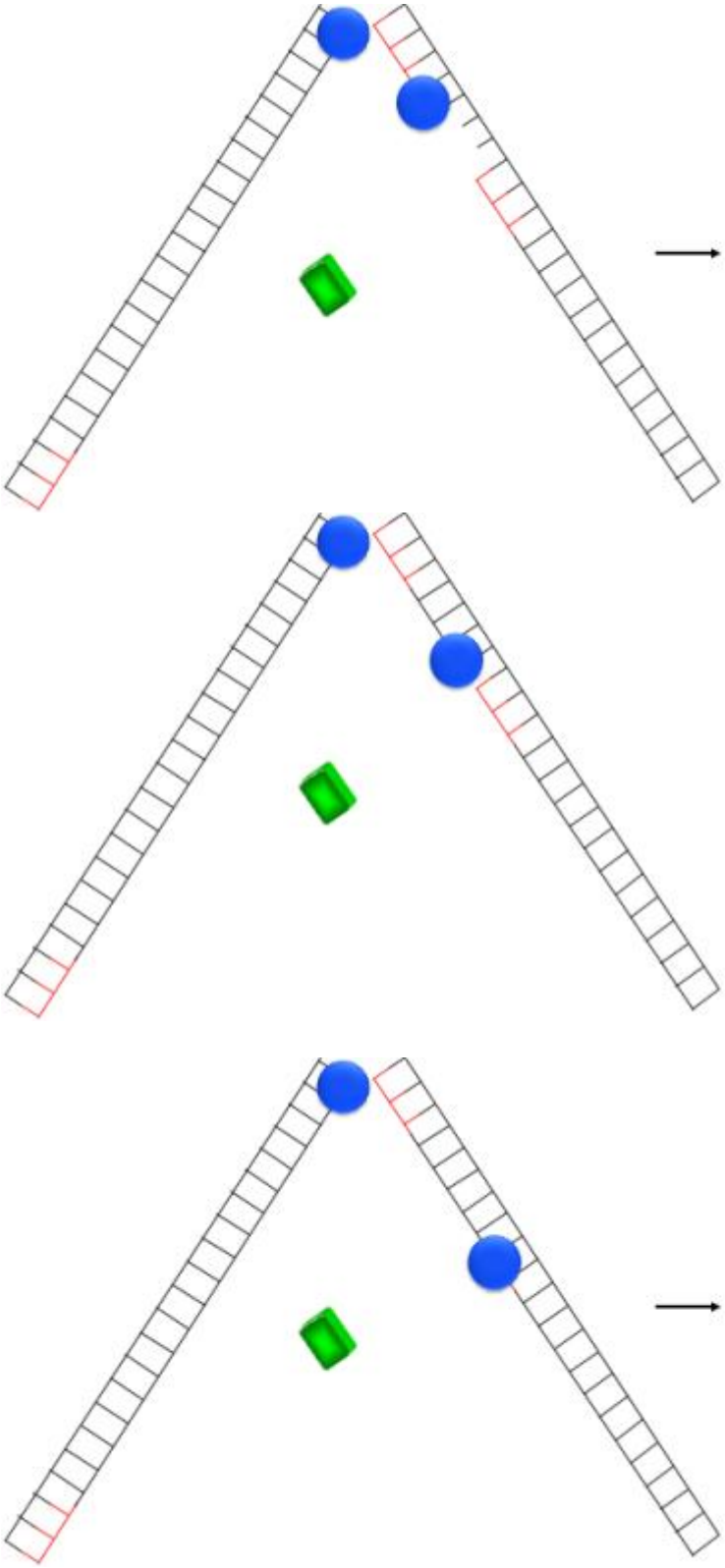
Helicase continues to separate the strands and the new DNA strand shown on the left continues to be lengthened.

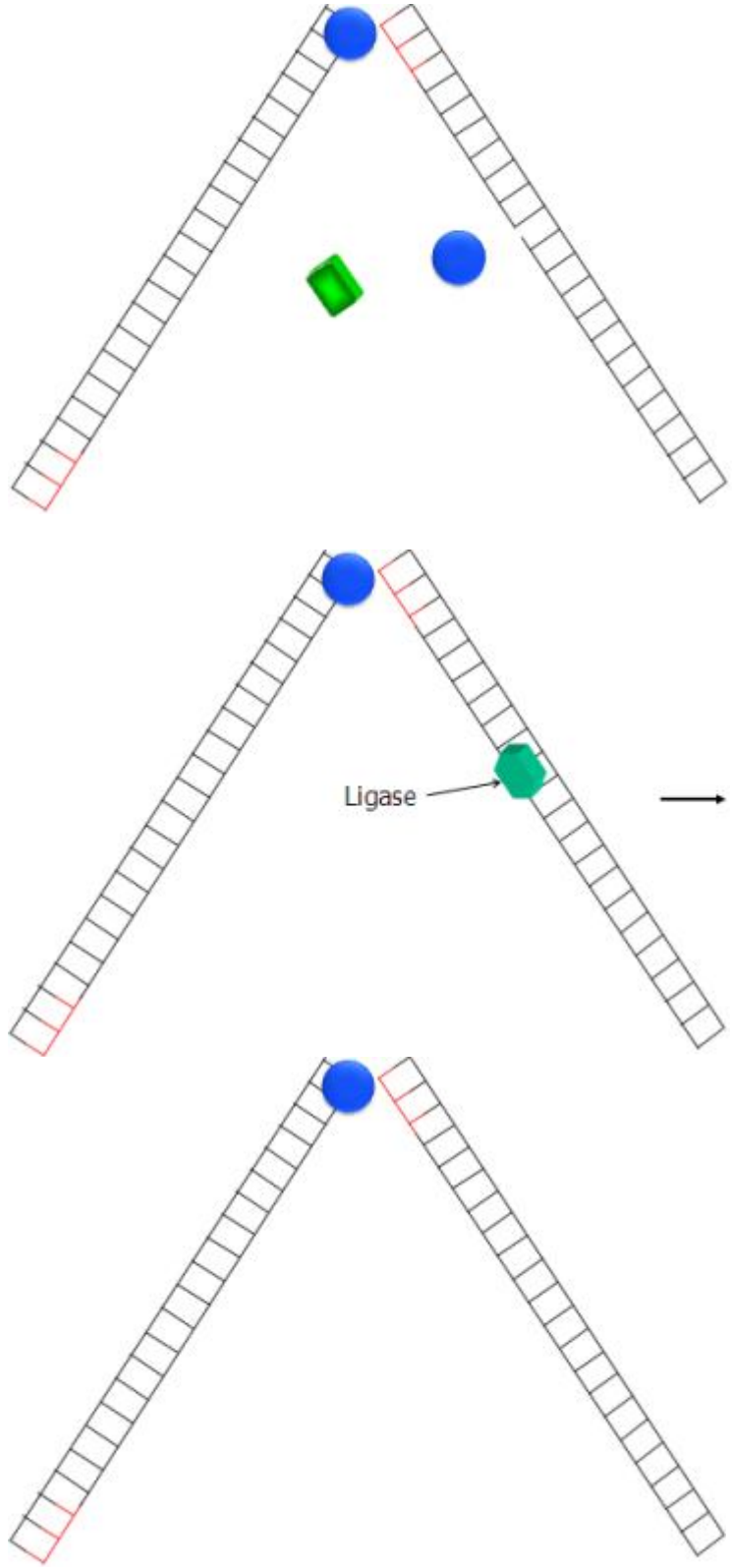




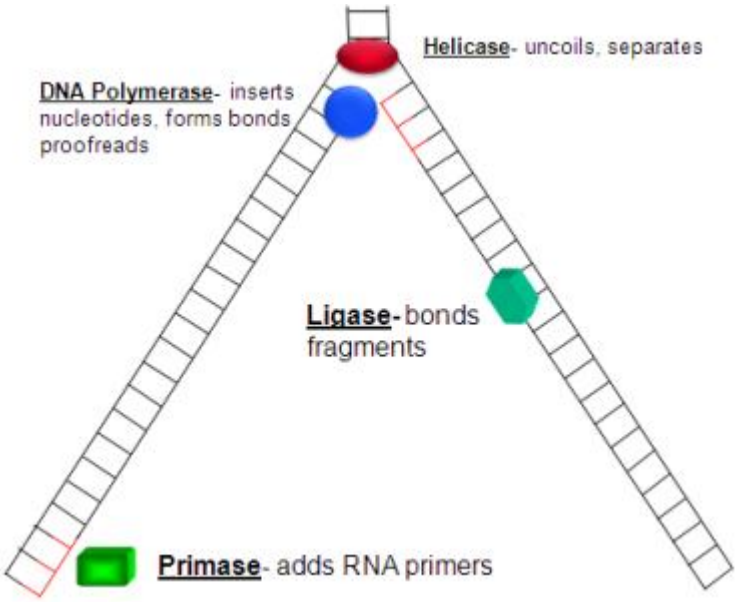
Another RNA primer is now needed because DNA polymerase III cannot initiate synthesis.







This diagram summerizes the enzymes that function in DNA synthesis.



Link

The link below may be a helpful summary.

http://www.johnkyrk.com/DNAreplication.html

Repair of damaged DNA

Changes in the DNA code are called ***mutations.*** Repair enzymes repair most of the errors that occur in DNA. There are three different classes of repair mechanisms.

1. Proofreading corrects errors made during the DNA replication process.
2. ***Mismatch repair*** corrects base pair mismatching (A-T and G-C).
3. Excision repair removes and replaces small segments of damaged DNA.

Errors corrected as a result of DNA synthesis

Proofreading

The overall error rate during DNA replication in E.coli is one nucleotide in 100,000.

DNA polymerase proofreads the new strand of DNA as it is synthesized and it removes mismatched bases and replaces them with the correct bases.

After proofreading, the error rate is 1 in 10 billion nucleotides.

Mismatch Repair Immediately after Replication

After DNA is replicated, repair enzymes function to locate mismatched base pairs and replace the segment with the correct nucleotides.

In order to repair the mismatched bases, repair enzymes need to be able to identify which strand is the new strand. The mechanism of strand identification used by repair enzymes is not well-understood.

DNA Repair Enzymes

In addition to incorrectly paired bases during replication, a number of environmental agents such as radiation (UV, X-rays, radioactive elements) and chemicals (pesticides, cigarette smoke) can cause changes in the DNA.

Repair enzymes monitor the DNA and repair these changes. For example, excision repair occurs when a mutated segment of DNA is removed and replaced with a new segment. The new segment is then sealed to the original strand by ligase.

A common type of error caused by ultraviolet radiation occurs when two thymines become bonded to each other, forming a kink in the DNA molecule. This type of mutation, called a ***thymine dimer***, can result in incorrect nucleotides being paired with it when the strand is replicated. To repair this mutation, an enzyme removes a segment of DNA that contains the dimer and replaces the removed nucleotides with nucleotides complimentary to the opposite strand. The new DNA is then bonded to the original strand with DNA ligase.

Mutation

Changes in the DNA which are not repaired correctly are called ***mutations***.

Most mutations are harmful. Occasionally, a mutation is beneficial and spreads by natural selection.

Mutations are the source of variation needed for natural selection to operate.

Organization of DNA

Chromosome Structure

Chromosomes are structures composed of condensed DNA and associated proteins. When DNA condenses, the molecule becomes wrapped around proteins called ***histones*** forming a bead-like structure called a ***nucleosome***. DNA and its associated nucleosomes form a beaded thread-like material that is 10 nm in diameter.

The histones are then compacted together to produce a larger fiber 30 nm in diameter.

The 30 nm fibers are further compacted by looping and attachment to a protein scaffolding to produce ***looped domains*** (300 nm).

When the cell is not dividing (interphase), 30 nm fibers are abundant. The 10 nm fibers (DNA and nucleosomes) as well as looped domains (300 nm) are also present. Larger (700 nm) chromosomes form during cell division (mitosis and meiosis).

Evidence indicates that each chromosome occupies a specific area of the nucleus.

Heterochromatin and Euchromatin

Chromatin is DNA and its associated protein.

During interphase, looped domains may be attached to protein supporting structures on the inside of the

nuclear membrane.

Some of the chromatin in the nucleus, called ***heterochromatin***, is tightly coiled and condensed. In this state, it is not transcribed.

Euchromatin is a form of chromatin that is less condensed and is actively transcribed.

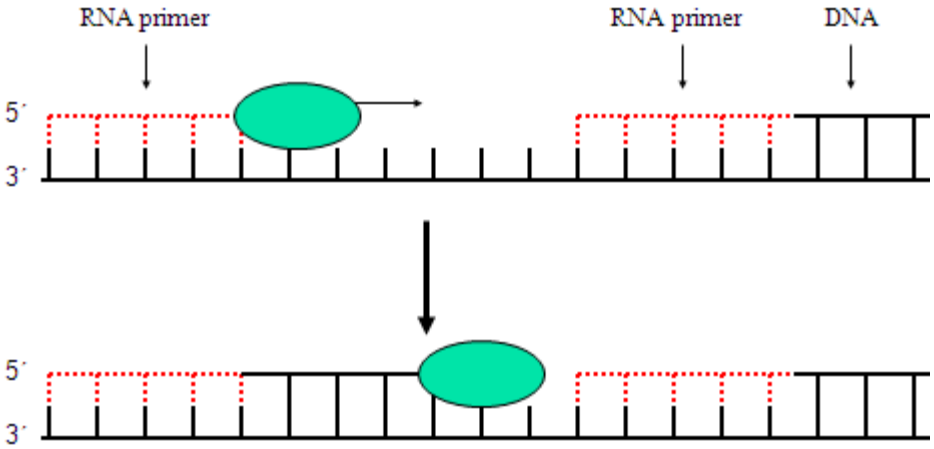
Unknown Function

Less than 5% of eukaryotic DNA functions to code for proteins. Approximately 1.5% of human DNA codes for protein. The function of the remaining DNA is not known but perhaps much of it has no function. In the past, noncoding DNA was sometimes called "junk DNA" because its function was not well understood.

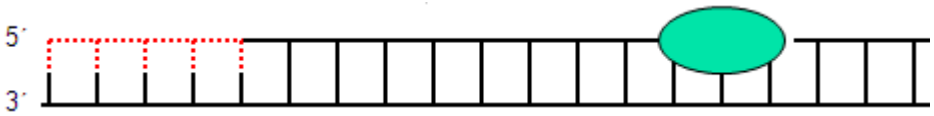
Some parts of the DNA contain more genes than other parts. The gene-rich portions are rich in G and C while the noncoding DNA is rich in A and T. The light bands on chromosomes are gene-rich regions.

Telomeres

DNA polymerase is not capable of initiating the synthesis of DNA; it can only elongate a strand that has already been started. Normally, an RNA primer functions to begin the process, allowing DNA polymerase to attach and finish synthesizing the strand.



RNA primers that are not at the ends of the DNA will be replaced with DNA nucleotides.



RNA Primers located on the 5' end of a DNA strand cannot be replaced because DNA polymerase cannot begin at the end of a strand. It can only add to an existing strand.

The new DNA strand is shorter than the template strand. As a result of the inability of DNA polymerase to initiate synthesis, the DNA molecule becomes shorter with each replication.

Human chromosomes have the sequence "TTAGGG" repeated 100 to 1500 times at each end of the DNA strand. This repeated sequence, called a telomere, becomes shorter each time the DNA

replicates.

Telomeres do not contain any genetic information.

Because DNA replication is part of the cycle of cell reproduction, the telomeres become shorter each time the cell reproduces. Short telomeres may prevent a cell from dividing. The length of telomeres, therefore, limits the number of times a cell can divide.

Telomerase is an enzyme that restores the length of telomeres. The active enzyme is normally not found in somatic (body) cells but is found in germ cells.

Prokaryotic chromosomes do not have telomeres. The chromosomes are circular and thus do not become shortened with each replication because RNA primers do not occur at the ends of the DNA.

Prokaryotes

The DNA of prokaryotes is not condensed into chromosomes as in eukaryotes. Their chromosome consists of a single loop of DNA.

Replication begins at a single origin and proceeds in both directions. It typically requires 40 minutes.

Another round of replication may begin before the previous one has finished, thus some cells are able to divide every 20 minutes.