DNA Replication

DNA replication involves separation of the two original strands and production of two new strands with the original strands as templates. Each new DNA molecule contains one strand from the original DNA and one newly synthesized strand. This situation is called **semiconservative replication.** Semiconservative replication of DNA was established in the late 1950s by experiments performed by Matthew Meselson and Franklin Stahl. *E. coli* bacteria were grown with $^{15}NH_4Cl$ as the sole nitrogen source, ^{15}N being a heavy isotope of nitrogen. (The usual isotope of nitrogen is ^{14}N.) In such a medium, all newly formed nitrogen compounds, including purine and pyrimidine nucleobases, become labeled with ^{15}N. The ^{15}N-labeled DNA has a higher density than unlabeled DNA, which contains the usual isotope, ^{14}N. In this experiment, the ^{15}N-labeled cells were then transferred to a medium that contained only ^{14}N. The cells continued to grow in the new medium. With every new generation of growth, a sample of DNA was extracted and analyzed by the technique of **density-gradient centrifugation** (Figure). This technique depends on the fact that heavy 15N DNA (DNA that contains ^{15}N alone) forms a band at the bottom of the tube; light ^{14}N DNA (containing ^{14}N alone) appears at the top of the tube. DNA containing a 50–50 mixture of ^{14}N and ^{15}N appears at a position halfway between the two bands. In the actual experiment, this 50–50 hybrid DNA was observed after one generation, a result to be expected with semiconservative replication. After two generations in the lighter medium, half of the DNA in the cells should be the 50–50 hybrid and half should be the lighter ^{14}N DNA. This prediction of the kind and amount of DNA that should be observed was confirmed by the experiment.

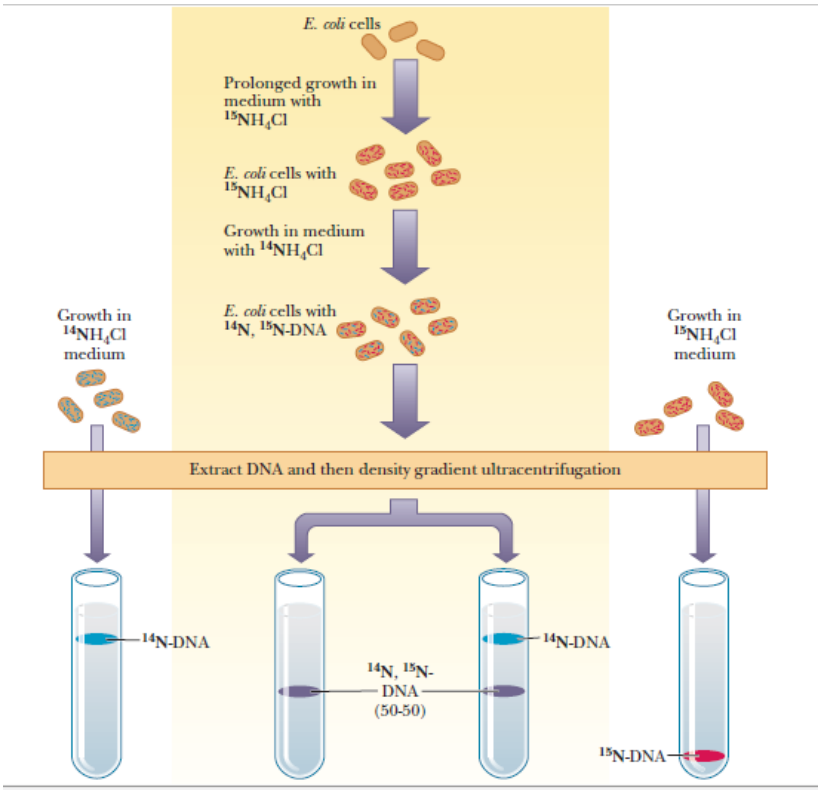
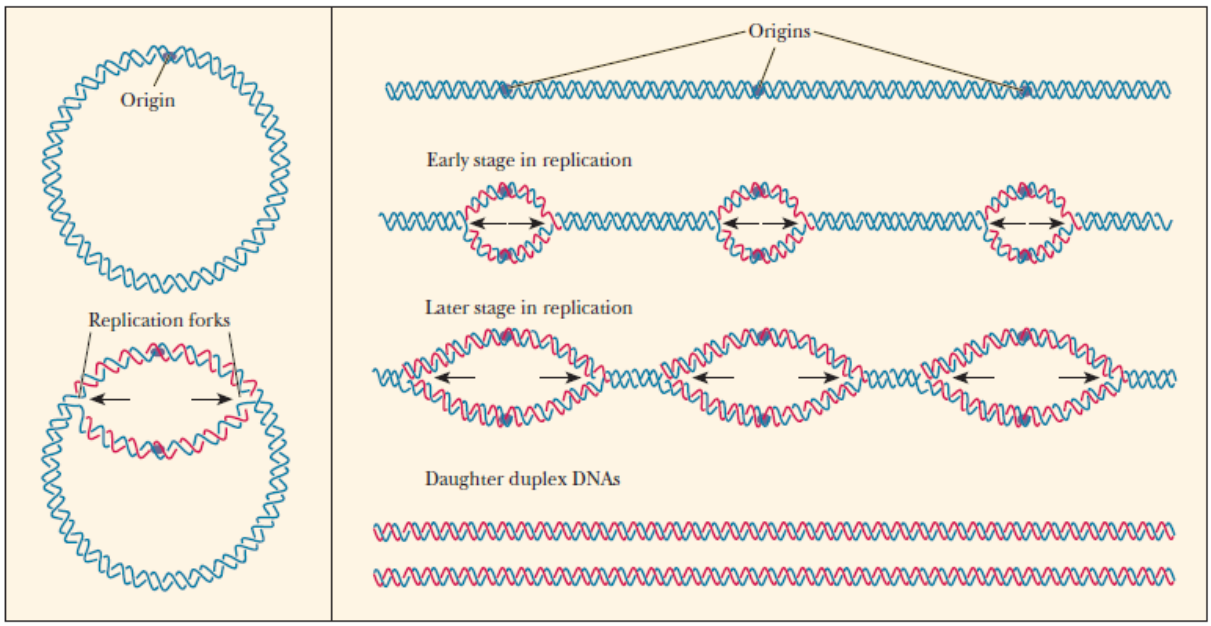


Figure: Experiments performed by Matthew Meselson and Franklin Stah

Origin of replication

During replication, the DNA double helix unwinds at a specific point called the **origin of replication** (OriC in *E. coli*). New polynucleotide chains are synthesized using each of the exposed strands as a template. Two possibilities exist for the growth of the new strands: synthesis can take place in both directions from the origin of replication, or in one direction only. It has been established that DNA synthesis is bidirectional in most organisms, with the exception of a few viruses and plasmids. (Plasmids are rings of DNA that are found in bacteria and that replicate independently from the regular bacterial genome. For each origin of replication, there are two points **(replication forks)** at which new polynucleotide chains are formed. A "bubble" (also called an "eye") of newly synthesized DNA between regions of the original DNA is a manifestation of the advance of the two replication forks in opposite directions. This feature is also called a θ structure because of its resemblance to the lowercase Greek letter theta. One such bubble (and one origin of replication) exists in the circular DNA of prokaryotes (Figure). In eukaryotes, several origins of replication, and thus several bubbles, exist (Figure). The bubbles grow larger and eventually merge, giving rise to two complete daughter DNAs. This bidirectional growth of both new polynucleotide chains represents *net chain growth.* Both new polynucleotide chains are synthesized in the 5' to 3' direction.



A Replication of the chromosome of *E. coli*, a typical prokaryote. There is one origin of replication, and there are two replication forks.

B Replication of a eukaryotic chromosome. There are several origins of replication, and there are two replication forks for each origin. The "bubbles" that arise from each origin eventually coalesce.

Stages of DNA replication

DNA replication can be divided into three stages; **Initiation, Elongation, Termination**

1. Initiation

DNA synthesis is initiated at particular points within the DNA strand known as '**origins**', which are specific coding regions. These origins are targeted by initiator proteins, which go on to recruit more proteins that help aid the replication process, forming a replication complex

around the DNA origin. There are multiple origin sites, and when replication of DNA begins, these sites are referred to as **Replication Forks**.

Within the replication complex is the enzyme **DNA Helicase**, which unwinds the double helix and exposes each of the two strands, so that they can be used as a template for replication. It does this by hydrolysing the ATP used to form the bonds between the nucleobases, therefore breaking the bond between the two strands. DNA can only be extended via the addition of a free nucleotide triphosphate to the **3'- end** of a chain. As the double helix runs **antiparallel,** but DNA replication only occurs in one direction, it means growth of the two new strands is very different (and will be covered in Elongation). DNA Primase is another enzyme that is important in DNA replication. It synthesises a small **RNA primer**, which acts as a 'kick-starter' for **DNA Polymerase.** DNA Polymerase is the enzyme that is ultimately responsible for the creation and expansion of the new strands of DNA.

2. Elongation

Once the DNA Polymerase has attached to the original, unzipped two strands of DNA (i.e. the **template** strands), it is able to start synthesising the new DNA to match the templates. This enzyme is only able to extend the primer by adding free nucleotides to the **3'-end** of the strand, causing difficulty as one of the template strands has a 5'-end from which it needs to extend from.

One of the templates is read in a 3' to 5' direction, which means that the new strand will be formed in a 5' to 3' direction (as the two strands are antiparallel to each other). This newly formed strand is referred to as the Leading Strand. Along this strand, DNA Primase only needs to synthesise an **RNA primer** once, at the beginning, to help initiate DNA Polymerase to continue extending the new DNA strand. This is because DNA Polymerase is able to extend the new DNA strand normally, by adding new nucleotides to the 3' end of the new strand (how DNA Polymerase usually works).

However, the other template strand is antiparallel, and is therefore read in a **5' to 3'** direction, meaning the new DNA strand being formed will run in a 3' to 5' direction. This is an issue as DNA Polymerase doesn't extend in this direction. To counteract this, DNA Primase synthesises a new RNA primer approximately every 200 nucleotides, to prime DNA synthesis to continue extending from the 5' end of the new strand. To allow for the continued creation of RNA primers, the new synthesis is delayed and is such called the **Lagging Strand.**

The **leading strand** is one complete strand, while the lagging strand is **not.** It is instead made out of multiple 'mini-strands', known of **Okazaki fragments**. These fragments occur due to the fact that new primers are having to be synthesised, therefore causing multiple strands to be created, as opposed to the one initial primer that is used with the leading strand.

3. Termination

The process of expanding the new DNA strands continues until there is either no more DNA template left to replicate (i.e. at the end of the chromosome), or two replication forks meet and subsequently **terminate.** The meeting of two replication forks is not regulated and happens randomly along the course of the chromosome. Once DNA synthesis has finished, it is important that the newly synthesised strands are bound and stabilized. With regards to the lagging strand, two enzymes are needed to achieve this; **RNAase H** removes the RNA primer

that is at the beginning of each Okazaki fragment, and **DNA Ligase** joins two fragments together creating one complete strand. Now with two new strands being finally finished, the DNA has been successfully replicated, and will just need other intrinsic cell systems to 'proof-read' the new DNA to check for any errors in replication, and for the new single strands to be stabilized.

Enzyme	Function in DNA replication
DNA helicase	Also known as helix destabilizing enzyme. Helicase separates the two strands of DNA at the Replication Fork behind the topoisomerase.
DNA polymerase	The enzyme responsible for catalyzing the addition of nucleotide substrates to DNA in the 5′ to 3′ direction during DNA replication. Also performs proof-reading and error correction. There exist many different types of DNA Polymerase, each of which perform different functions in different types of cells.
DNA clamp	A protein which prevents elongating DNA polymerases from dissociating from the DNA parent strand.
Single-strand DNA-binding protein	Bind to ssDNA and prevent the DNA double helix from re-annealing after DNA helicase unwinds it, thus maintaining the strand separation, and facilitating the synthesis of the nascent strand.
Topoisomerase	Relaxes the DNA from its super-coiled nature.
DNA gyrase	Relieves strain of unwinding by DNA helicase; this is a specific type of topoisomerase
DNA ligase	Re-anneals the semi-conservative strands and joins Okazaki Fragments of the lagging strand.
Primase	Provides a starting point of RNA (or DNA) for DNA polymerase to begin synthesis of the new DNA strand.
Telomerase	Lengthens telomeric DNA by adding repetitive nucleotide sequences to the ends of eukaryotic chromosomes. This allows germ cells and stem cells to avoid the Hayflick limit on cell division.

Replication fork

The replication fork is a structure that forms within the long helical DNA during DNA replication. It is created by helicases, which break the hydrogen bonds holding the two DNA strands together in the helix. 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.

DNA is read by DNA polymerase in the 3′ to 5′ direction, meaning the nascent strand is synthesized in the 5' to 3' direction**.** Since the leading and lagging strand templates are oriented in opposite directions at the replication fork, a major issue is how to achieve synthesis of nascent (new) lagging strand DNA, whose direction of synthesis is opposite to the direction of the growing replication fork.

Leading strand

The leading strand is the strand of nascent DNA which is synthesized in the same direction as the growing replication fork. This sort of DNA replication is continuous.

Lagging strand

The lagging strand is the strand of nascent DNA whose direction of synthesis is opposite to the direction of the growing replication fork. Because of its orientation, replication of the lagging strand is more complicated as compared to that of the leading strand. As a consequence, the DNA polymerase on this strand is seen to "lag behind" the other strand. The lagging strand is synthesized in short, separated segments. On the lagging strand *template*, a primase "reads" the template DNA and initiates synthesis of a short complementary RNA primer. A DNA polymerase extends the primed segments, forming Okazaki fragments. The RNA primers are then removed and replaced with DNA, and the fragments of DNA are joined together by DNA ligase.

In all cases the helicase is composed of six polypeptides that wrap around only one strand of the DNA being replicated. The two polymerases are bound to the helicase heximer. In eukaryotes the helicase wraps around the leading strand, and in prokaryotes it wraps around the lagging strand.

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 forms a torsional resistance that would eventually halt the progress of the replication fork. Topoisomerases are enzymes that temporarily break the strands of DNA, relieving the tension caused by unwinding the two strands of the DNA helix; topoisomerases (including DNA gyrase) achieve this by adding negative supercoils to the DNA helix. Bare single-stranded DNA tends to fold back on itself forming 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.

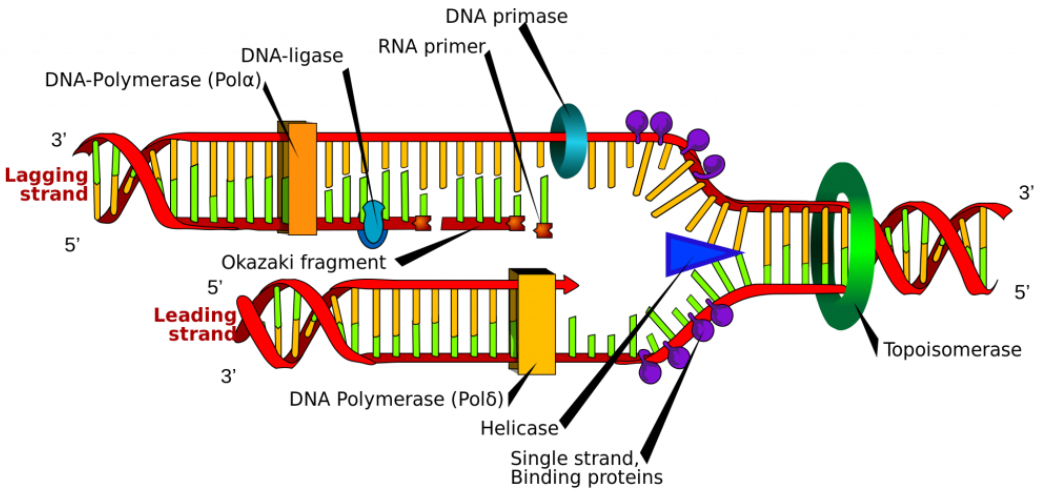


Fig: Diagrammatic representation of DNA replication

All synthesis of nucleotide chains occurs in the 5' 3 3' direction from the perspective of the chain being synthesized. This is due to the nature of the reaction of DNA synthesis. The last nucleotide added to a growing chain has a 3'-hydroxyl on the sugar. The incoming nucleotide has a 5'-triphosphate on its sugar. The 3'-hydroxyl group at the end of the growing chain is a nucleophile. It attacks the phosphorus adjacent to the sugar in the nucleotide to be added to the growing chain, leading to the elimination of the pyrophosphate and the formation of a new phosphodiester bond (Figure).

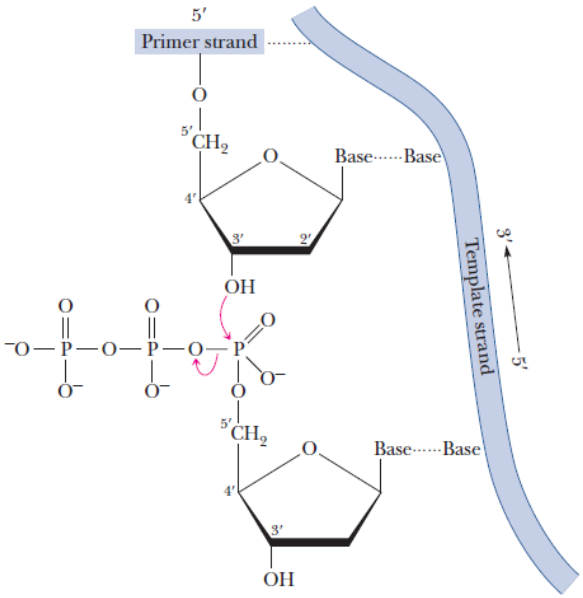


Figure: The addition of a nucleotide to a growing DNA chain. The 3'-hydroxyl group at the end of the growing DNA chain is a nucleophile. It attacks at the phosphorus adjacent to the sugar in the nucleotide, which is added to the growing chain. Pyrophosphate is eliminated and a new phosphodiester bond is formed.

Prokaryotic DNA replication

The first DNA polymerase discovered was found in *E. coli.* **DNA polymerase** catalyzes the successive addition of each new nucleotide to the growing chain. At least five DNA polymerases are present in *E. coli.* Three of them have been studied more extensively, and some of their properties are listed in Table 10.1. DNA polymerase I (Pol I) was discovered first, with the subsequent discovery of polymerases II (Pol II) and polymerase III (Pol III). Polymerase I consists of a single polypeptide chain, but polymerases II and III are multisubunit proteins that share some common subunits. Polymerase II is not required for replication; rather, it is strictly a repair enzyme. Recently, two more polymerases, Pol IV and Pol V, were discovered. They, too, are repair enzymes, and both are involved in a unique repair mechanism called the SOS response. Two important considerations regarding the polymerases are the speed of the synthetic reaction (turnover number) and the **processivity,** which is the number of nucleotides joined before the enzyme dissociates from the template.

Properties of DNA Polymerases of *E. coli*

Property	Pol I	Pol II	Pol III
Mass (kDa)	103	90	830
Turnover number (min^{-1})	600	30	1200
Processivity	200	1500	≥500,000
Number of subunits	1	≥4	≥10
Structural gene	*pol*A	*pol*B*	*pol*C*
Polymerization 5'→ 3'	Yes	Yes	Yes
Exonuclease 5' → 3'	Yes	No	No
Exonuclease 3' → 5'	Yes	Yes	Yes

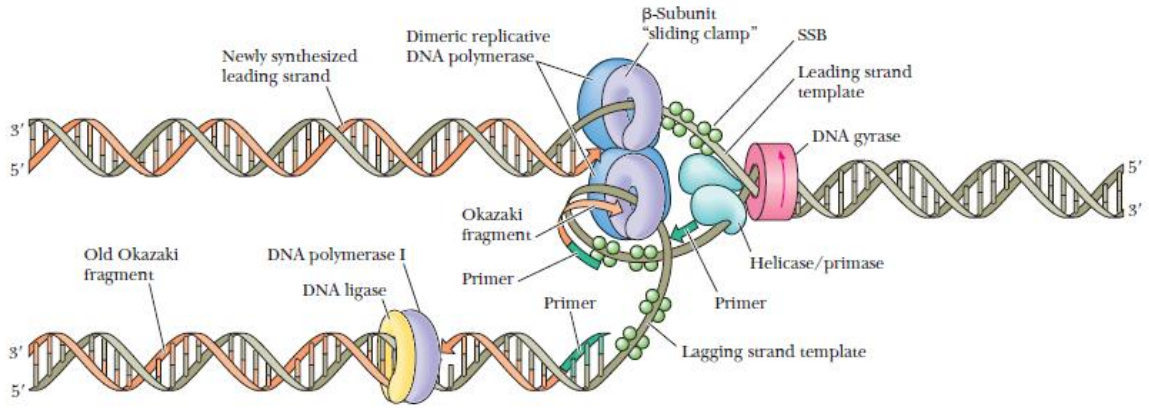
* Polymerization subunit only. These enzymes have multiple subunits, and some of them are shared between both enzymes.

The Subunits of *E. coli* DNA Polymerase III Holoenzyme

Subunit	Mass (kDa)	Structural Gene	Function
α	130.5	*polC* (*dnaE*)	Polymerase
ε	27.5	*dnaQ*	3'-exonuclease
θ	8.6	*holE*	α, ε assembly?
τ	71	*dnaX*	Assembly of holoenzyme on DNA
β	41	*dnaN*	Sliding clamp, processivity
γ	47.5	*dnaX(Z)*	Part of the γ complex*
δ	39	*holA*	Part of the γ complex*
δ	37	*holB*	Part of the γ complex*
χ	17	*holC*	Part of the γ complex*
ψ	15	*holD*	Part of the γ complex*

* Subunits γ, δ, δ , χ, and ψ form the so-called γ complex, which is responsible for the placement of the β-subunits (the sliding clamp) on the DNA. The γ complex is referred to as the clamp loader. The δ and τ subunits are encoded by the same gene.

The synthesis of two new strands of DNA is begun by DNA polymerase III. The newly formed DNA is linked to the 3'-hydroxyl of the RNA primer, and synthesis proceeds from the 5' end to the 3' end on both the leading and the lagging strands. Two molecules of Pol III, one for the leading strand and one for the lagging strand, are physically linked to the *primosome.* The resulting multiprotein complex is called the **replisome.** As the replication fork moves, the RNA primer is removed by polymerase I, using its exonuclease activity. The primer is replaced by deoxynucleotides, also by DNA polymerase I, using its polymerase activity. (The removal of the RNA primer and its replacement with the missing portions of the newly formed DNA strand by polymerase I are the repair function we mentioned earlier.) None of the DNA polymerases can seal the nicks that remain; DNA ligase is the enzyme responsible for the final linking of the new strand.



General features of a Prokaryotic replication fork. The DNA duplex is unwound by the action of DNA gyrase and helicase, and the single strands are coated with SSB (ssDNA-binding protein). Primase periodically primes synthesis on the lagging strand. Each half of the dimeric replicative polymerase is a holoenzyme bound to its template strand by a β-subunit sliding clamp. DNA polymerase I and DNA ligase act downstream on the lagging strand to remove RNA primers, replace them with DNA, and ligate the Okazaki fragments.

Summary of DNA replication in Prokaryotes

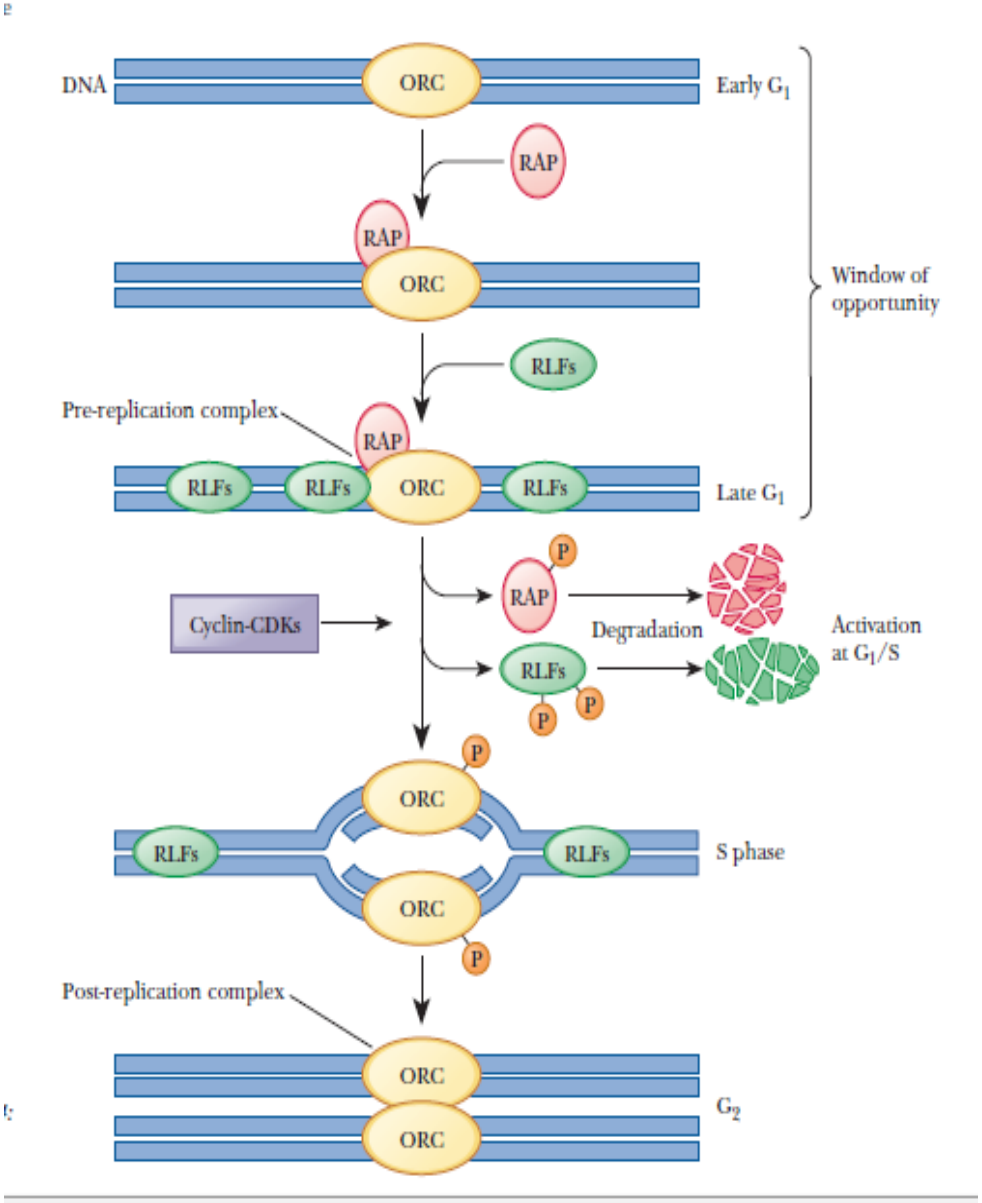
- DNA synthesis is bidirectional. Two replication forks advance in opposite directions from an origin of replication.
- The direction of DNA synthesis is from 5' end to the 3' end of the newly formed strand. One strand (the leading strand) is formed continuously, while the other strand (the lagging strand) is formed discontinuously. On the lagging strand, small fragments of DNA (Okazaki fragments) are subsequently linked.
- Five DNA polymerases have been found in *E. coli*. Polymersase III is primarily responsible for the synthesis of new strands. The first polymerase enzyme discovered, polymerase I, is involved in synthesis, proofreading, and repair. Polymerase II, IV, and V function as repair enzymes under unique conditions.
- DNA gyrase introduces a swivel point in advance of the movement of the replication fork. A helix-destabilizing protein, a helicase, binds at the replication fork and promotes unwinding. The exposed single-stranded regions of the template DNA are stabilized by a DNA-binding protein.
- Primase catalyzes the synthesis of an RNA primer.

- The synthesis of new strands is catalyzed by Pol III. The primer is removed by Pol I, which also replaces the primer with deoxynucleotides. DNA ligase seals theremaining nicks.

Eukaryotic DNA Replication:

Eukaryotic replication is more complicated in three basic ways: there are multiple origins of replication, the timing must be controlled to that of cell divisions, and more proteins and enzymes are involved. In a human cell, a few billion base pairs of DNA must be replicated once, and only once, per cell cycle. Cell growth and division are divided into phases—M, G1, S, and G2. DNA replication takes place during a few hours in the S phase, and pathways exist to make sure that the DNA is replicated only once per cycle. Eukaryotic chromosomes accomplish this DNA synthesis by having replication begin at multiple origins of replication, also called **replicators.** These are specific DNA sequences that are usually between gene sequences. An average human chromosome may have several hundred replicators. The zones where replication is proceeding are called **replicons,** and the size of these varies with the species. In higher mammals, replicons may span 500 to 50,000 base pairs.

How is replication tied to cell division?
The best-understood model for control of eukaryotic replication is from yeast cells (Figure). Only chromosomes from cells that have reached the G1 phase are competent to initiate DNA replication. Many proteins are involved in the control of replication and its link to the cell cycle. As usual, these proteins are usually given an abbreviation that makes them easier to say, but more difficult for the uninitiated to comprehend at first glance. The first proteins involved are seen during a window of opportunity that occurs between the early and late G1 phase (see Figure). Replication is initiated by a multisubunit protein called the **origin recognition complex (ORC),** which binds to the origin of replication. This protein complex appears to be bound to the DNA throughout the cell cycle, but it serves as an attachment site for several proteins that help control replication. The next protein to bind is an activation factor called the **replication activator protein (RAP).** After the activator protein is bound, **replication licensing factors (RLFs)** can bind. Yeast contains at least six different RLFs. They get their name from the fact that replication cannot proceed until they are bound. One of the keys to linking replication to cell division is that some of the RLF proteins have been found to be cytosolic. Thus, they have access to the chromosome only when the nuclear membrane dissolves during mitosis. Until they are bound, replication cannot occur. After RLFs bind, the DNA is then competent for replication. The combination of the DNA, ORC, RAP, and RLFs constitutes what researchers call the **pre-replication complex (pre-RC)** proteins. One of the great discoveries in this field was the existence of **cyclins,** which are proteins that are produced in one part of a cell cycle and degraded in another. Cyclins are able to combine with specific protein kinases, called **cyclin dependent protein kinases** (CDKs). When these cyclins combine with CDKs, they can activate DNA replication and also block reassembly of a pre-RC after initiation. The state of activity of the CDKs and the cyclins determines the window of opportunity for DNA synthesis. Cyclin–CDK complexes phosphorylate sites on the RAP, the RLFs, and the ORC itself. Once phosphorylated, the RAP dissociates from the pre-RC, as do the RLFs. Once phosphorylated and released, the RAP and the RLFs are degraded (Figure *middle*). Thus, the activation of cyclin–CDKs serves both to initiate DNA replication and to prevent formation of another pre-RC. In the G2 phase, the DNA has been replicated. During mitosis, the DNA is separated into the daughter cells. At the same time, the dissolved nuclear membrane allows entrance of the licensing factors that are produced in the cytosol so that each daughter cell can initiate a new round of replication.



Model for initiation of the DNA replication cycle in eukaryotes. ORC is present at the replicators throughout the cell cycle. The prereplication complex (pre-RC) is assembled through sequential addition of the RAP (replication activator protein) and RLFs (replication licensing factors) during a window of opportunity defined by the state of cyclin–CDKs. Phosphorylation of the RAP, ORC, and RLFs triggers replication. After initiation, a post-RC state is established, and the RAP and RLFs are degraded. *(Adapted from Figure 2 in Stillman, B., 1996. Cell Cycle Control of DNA Replication.* Science ***274:****1659 1663. © 1996 AAAS. Used bypermission.)*

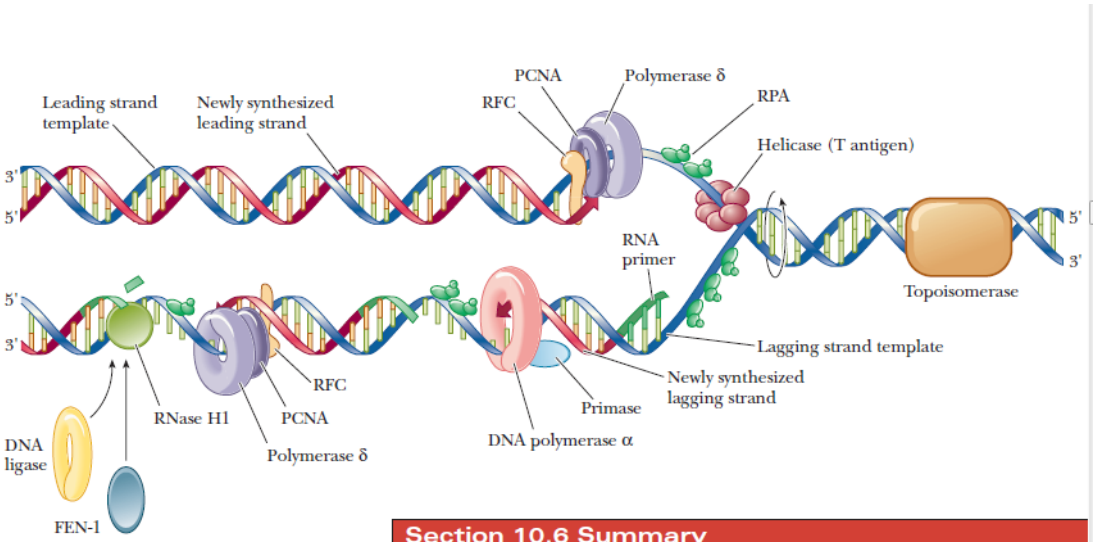
Eukaryotic DNA Polymerases

At least 15 different polymerases are present in eukaryotes, of which 5 have been studied more extensively (Table). The use of animals rather than plants for study avoids the complication of any DNA synthesis in chloroplasts. The five best-studied polymerases are called α, β, γ, δ, and ε. The α, β, δ, and ε enzymes are found in the nucleus, and the γ form occurs in mitochondria.

The Biochemical Properties of Eukaryotic DNA Polymerases					
	α	δ	ε	β	γ
Mass (kDa)					
Native	>250	170	256	36-38	160-300
Catalytic core	165-180	125	215	36-38	125
Other subunits	70, 50, 60	48	55	None	35, 47
Location	Nucleus	Nucleus	Nucleus	Nucleus	Mitochondria
Associated functions					
3' → 5' exonuclease	No	Yes	Yes	No	Yes
Primase	Yes	No	No	No	No
Properties					
Processivity	Low	High	High	Low	High
Fidelity	High	High	High	Low	High
Replication	Yes	Yes	Yes	No	Yes
Repair	No	?	Yes	Yes	No

The Eukaryotic Replication Fork

The general features of DNA replication in eukaryotes are similar to those in prokaryotes. Table summarizes the differences. As with prokaryotes, DNA replication in eukaryotes is semiconservative. There is a leading strand with continuous synthesis in the 5'→3' direction and a lagging strand with discontinuous synthesis in the 5'→3' direction. An RNA primer is formed by a specific enzyme in eukaryotic DNA replication, as is the case with prokaryotes, but in this case the primase activity is associated with Pol α. The structures involved at the eukaryotic replication fork are shown in Figure. The formation of Okazaki fragments (typically 150 to 200 nucleotides long in eukaryotes) is initiated by Pol α. After the RNA primer is made and a few nucleotides are added by Pol α, the polymerase dissociates and is replaced by Pol δ and its attached PCNA protein. Another protein, called *RFC* (replication factor C), is involved in attaching PCNA to Pol δ. The RNA primer is eventually degraded, but, in the case of eukaryotes, the polymerases do not have the 5'→3' exonuclease activity to do it. Instead, separate enzymes, FEN-1 and RNase H1, degrade the RNA. Continued movement of Pol δ fills in the gaps made by primer removal. As with prokaryotic replication, topoisomerases relieve the torsional strain from unwinding the helix, and a single-stranded binding protein, called RPA, protects the DNA from degradation. Finally, DNA ligase seals the nicks that separate the fragments.



Section 10.6 Summary

The basics of the eukaryotic replication fork. The primase activity is associated with DNA polymerase α. After a few nucleotides are incorporated, DNA polymerase δ, with its associated proteins called PCNA and RFC, bind and do the majority of the synthesis. The enzymes FEN-1 and RNase H1 degrade the RNA primers in eukaryotic replication. *(From* Cellular and Molecular Biology *by Karp, Figure 13–22. Used by permission of John Wiley & Sons, Inc.)*

Summary of eukaryotic DNA replication

- Replication in eukaryotes follows the same general outline as replication in prokaryotes, with the most important difference being the presence of histone proteins complexed to eukaryotic DNA.
- Different proteins are used, and the system is more complex than it is in prokaryotes. Replication is controlled so that it occurs only once during a cell-division cycle, during the S phase.
- Five different DNA polymerases are present in eukaryotes: α, β, γ, δ, and ε. Polymerase δ is the principal synthesizer of DNA and is the equivalent of Pol III in prokaryotes.

Differences in DNA Replication in Prokyryotes and Eukaryotes

Prokaryotes	Eukaryotes
Five polymerases (I, II, III, IV, V)	Five polymerases (α, β, γ, δ, ε)
Functions of polymerase:	Functions of polymerases:
I is involved in synthesis, proofreading, repair, and removal of RNA primers	α: a polymerizing enzyme
II is also a repair enzyme	β: is a repair enzyme
III is main polymerizing enzyme	γ: mitochondrial DNA synthesis
IV, V are repair enzymes under unusual conditions	δ: main polymerizing enzyme ε: function unknown
Polymerases are also exonucleases	Not all polymerases are exonucleases
One origin of replication	Several origins of replication
Okazaki fragments 1000-2000 residues long	Okazaki fragments 150-200 residues long
No proteins complexed to DNA	Histones complexed to DNA