BIO31570/BCH 31570 **Fall 2017**

Molecular Biology

DNA replication, part I

- **Overview of the replication process**
- **Correction mechanisms**
- **The replication machinery**
- **Similarities between bacteria and eukaryotes**
- **Histones and Replication**

Marc Ekker
Department of Biology
Gendron Hall room 280
Tel : 562-5800 ext 2605
mekker@uottawa.ca

19 September 2017

Learning Objectives

At the end of this section, you should be able to:

- Distinguish the template strand from the primer strand. Distinguish the leading (continuous) from the lagging (discontinuous) strand.

- Appreciate the role played by the energy contained in nucleotide triphosphates for DNA synthesis reaction.

- Review the concept of Okazaki fragments, something you should already be familiar with.

- Name the various factors contributing the high fidelity of DNA replication.

- Describe the various editing mechanisms taking place when DNA polymerase synthesize a new DNA molecule

- Describe the various steps of strand-directed mismatch repair.

- Distinguish the function of the three main DNA polymerases in *E. coli*.

- Describe the function of the various proteins involved in DNA replication.

- Appreciate the similarities in DNA replication of prokaryotes and eukaryotes and briefly describe a few differences between the two.

- Explain briefly how nucleosomes are formed during replication and how histone modifications are « copied » on the chromatin of the two daughter DNA molecules.

It is crucial for living organisms to maintain the integrity of their genetic information. Information contained in the DNA molecule must be faithfully copied during each cell division. In germ cells, the accurate replication of the DNA molecule will ensure the maintenance of characteristics specific to each species. In somatic cells, maintenance of genetic information is essential to survival. For example, it is known that accumulation of mistakes in the DNA molecule is responsible for numerous types of cancer.

DNA is subjected to **mutations**, due to numerous causes: radiation, chemicals present in the environment, other types of accidents. We estimate of **mutation rate** at one per billion (10^9) copied nucleotides. This rate would be roughly similar in all living organisms. To maintain a low mutation rate, the cell uses a fast and efficient machinery for DNA replication. Furthermore, as DNA is subjected to mutations attributable to a variety of causes (chemicals present in the environment, irradiation), these mutations need to be repaired efficiently. In the following lectures, we will examine mechanisms for DNA replication, DNA repair, and DNA recombination.

1. DNA replication mechanisms :

DNA replication is the copying of both of its strands, each of them used as a matrix. It involves incorporation of nucleotides through the action of an enzyme, **DNA polymerase**.

A) Overview of the replication process :

i) each strand of the double helix is used as a **template** :

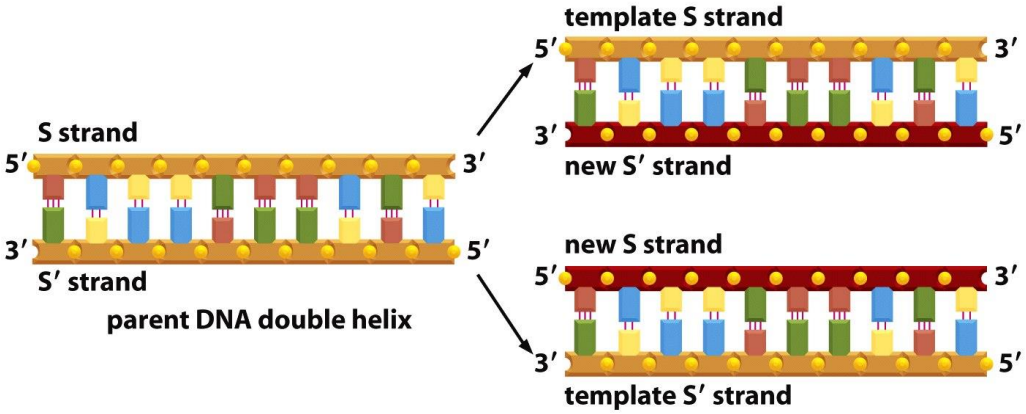


Figure 5-2 Molecular Biology of the Cell 5/e (© Garland Science 2008)

ii) The fundamental enzymatic reaction for the synthesis of DNA is the addition of a deoxyribonucleotide, entering the reaction as a dNTP, to the 3' end of DNA chain undergoing replication (**new strand** or **primer strand**). Pairing of the nucleotide to the **template strand** facilitates its insertion.

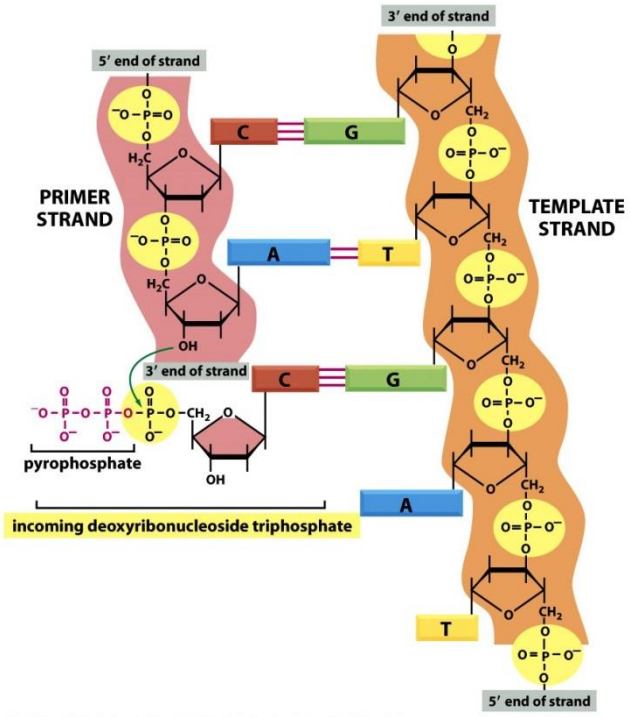


Figure 5-3 Molecular Biology of the Cell 5/e (© Garland Science 2008)

iii) This reaction is catalyzed by DNA polymerase. The new strand is synthesized in the 5' to 3' direction. The reaction is favored by the negative change in free energy produced by the hydrolysis of the tri-phosphate nucleotide (dNTP), more specifically the hydrolysis of a pyrophosphate molecule once the nucleotide has been added (see figure below). The crystallography structure of DNA polymerase resembles a right hand.

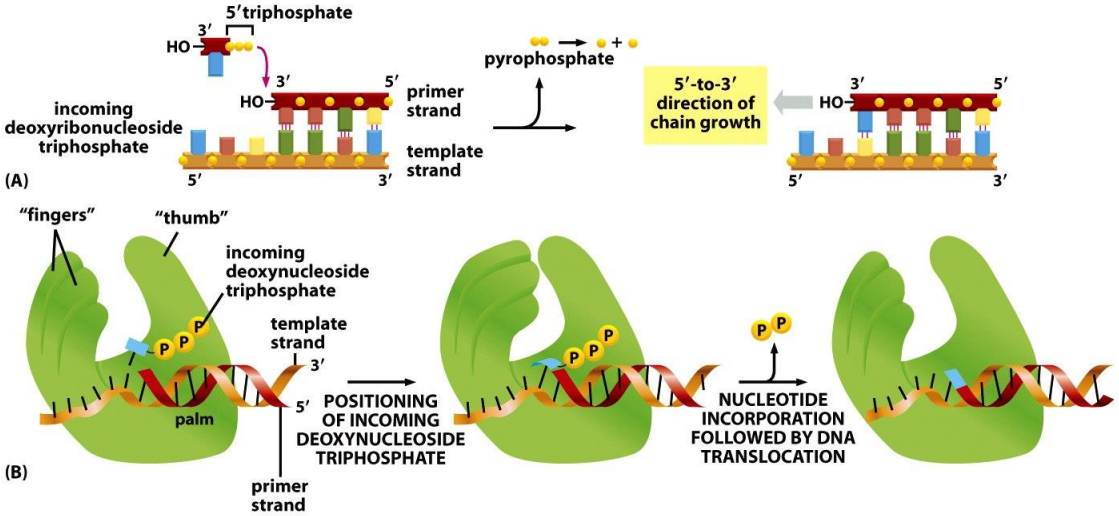


Figure 5-4 Molecular Biology of the Cell 5/e (© Garland Science 2008)

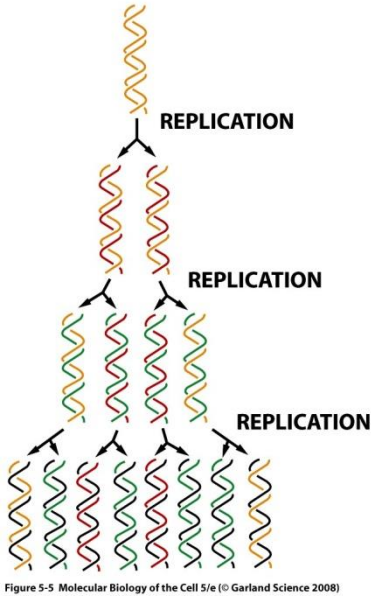


Figure 5-5 Molecular Biology of the Cell 5/e (© Garland Science 2008)

iv) DNA replication is **semi-conservative**. In each daughter cell, one strand of the DNA double helix will be one of the strands present in the mother cell; the other will be a newly-synthesized strand. This is shown on the drawing on the left.

v) The **replication fork** : DNA replication takes place in a specific region of the chromosome, shaped liked he letter Y, named the replication fork. However, because the two strand of the DNA molecule are in opposite 5' to 3' orientations and, as the fork globally moves into one direction, this creates a problem. A simple model would have DNA synthesis proceeding in the 5' to 3' (adding nucleotides at the 3' end of the new strand) direction on one strand and in the 3' to 5' direction (adding nucleotides at the 5' end of the new strand) on the other strand. However, this model is **inaccurate** as a DNA polymerase able to synthesize DNA in the 3' to 5' direction has not yet been discovered.

The way to solve this problem is to make short DNA fragments (1000-2000 nucleotides in prokaryotes, 100-200 in eukaryotes) on one of the two new strands (primer strands). These fragments are called **Okazaki fragments**. Each of these Okazaki fragments is synthesized in the 5' to 3' direction and they are subsequently linked to each other.

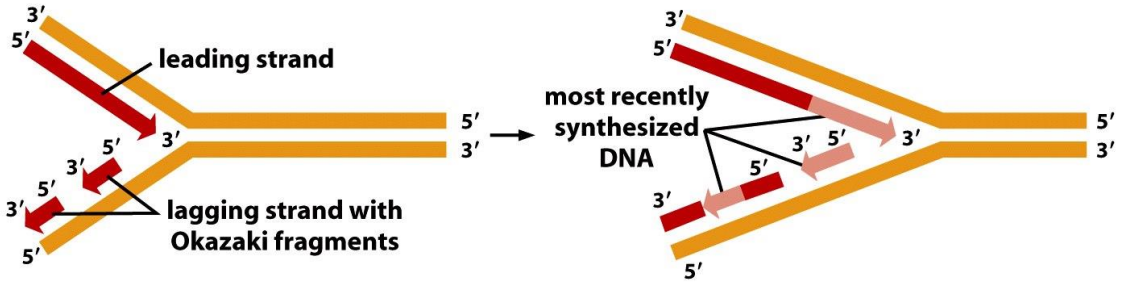


Figure 5-7 Molecular Biology of the Cell 5/e (© Garland Science 2008)

The replication fork is asymmetrical. The strand that is continuously synthesized is called the **leading strand** and the one synthesized in a discontinuous manner is called the **lagging strand**. One can see that on the lagging strand, DNA synthesis proceeds in the opposite direction compared to the overall direction of replication. DNA synthesis on the lagging strand is delayed with respect to that on the leading strand, because the template strand needs to be exposed to allow for the synthesis of Okazaki fragments. The discontinuous replication on the lagging strand is called backstitching.

2- Correction :

Fidelity of DNA copying is one error for 10^9 copied nucleotides. Correct pairing of the bases during replication cannot, by itself account for this remarkable degree of precision.

Some bases take, rarely, a tautomeric form that allow, for example, the paring of a C to an A, rather than to the usual G. Tautomeric form ratio are about one in 10^4-10^5. Therefore, one would anticipate a higher error rate in DNA replication. Proofreading mechanisms exist to eliminate errors resulting from mis-pairing.

This first correction is carried out by the enzyme DNA polymerase itself and this happens before the covalent addition of the nucleotide. This mechanisms implies: a) a return of the nucleotide to the usual tautomeric form, resulting in a less stable binding of the nucleotide to the polymerase; b) the fact that the conformation of the polymerase necessary for the synthesis of the covalent bond does not take place as efficiently when the nucleotide does not make its normal hydrogen bond with the template strand.

Even when a covalent bond that introduces an erroneous nucleotide takes place, this nucleotide can be removed by the **3'-5'exonuclease** correction activity of the polymerase. First, let's mention that the polymerase cannot link two nucleotides floating freely in solution. She needs a primer, that is a 3'OH end of a nucleotide paired to a DNA molecule. When a nucleotide is incorrectly incorporated in a DNA chain, the incorrect base pairing makes this nucleotide prone to detach from the template strand. Therefore, the polymerase has difficulties using this end to add the next nucleotide. A distinct catalytic site on the polymerase has the exonuclease activity that will remove the incorrectly incorporated nucleotide. Replication will resume once the correct nucleotide has been incorporated.

Interestingly, RNA polymerase, the enzyme responsible for the synthesis of RNA from DNA, has a much higher error rate of 1 in 10^4. Let's bear in mind that an error committed by RNA polymerase will not be transmitted to the next generations. Therefore, transcription and translation can tolerate higher error rates than DNA replication.

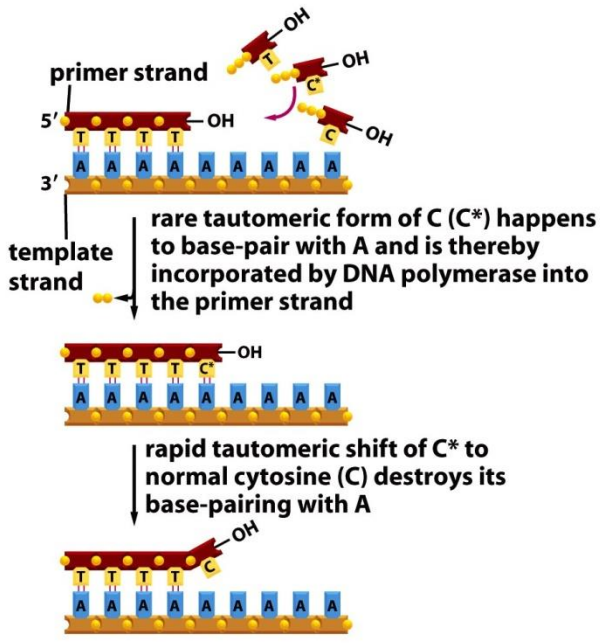


Figure 5-8 part 1 of 2 Molecular Biology of the Cell 5/e (© Garland Science 2008)

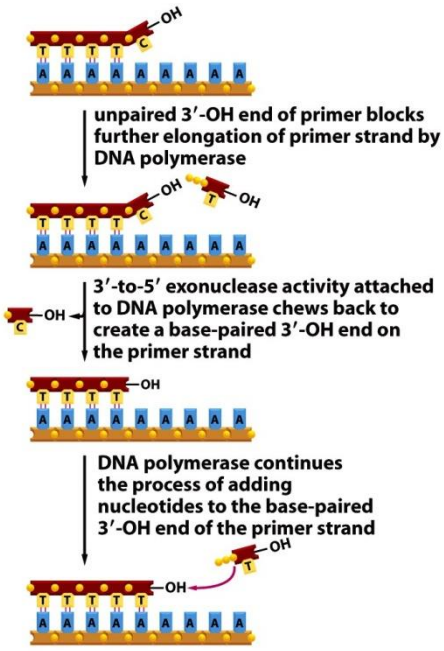


Figure 5-8 part 2 of 2 Molecular Biology of the Cell 5/e (© Garland Science 2008)

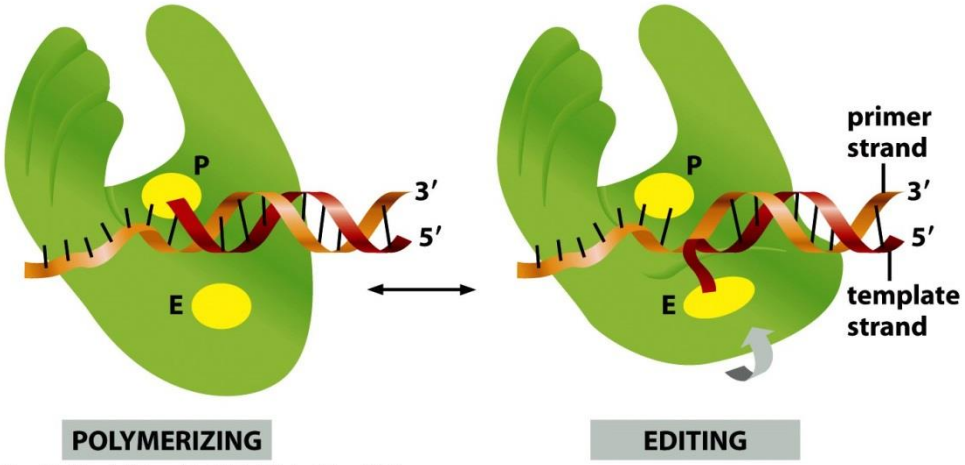


Figure 5-9 Molecular Biology of the Cell 5/e (© Garland Science 2008)

Strand-directed mismatch repair

Despite proofreading by DNA polymerase, replication errors sometimes take place and will be corrected by the strand-directed mismatch repair mechanism. This additional mechanism enables the correction of errors based on the distortion of the double helix caused by the incorrect incorporation of a nucleotide, thus the mis-pairing of a base.

The question you immediately ask yourselves is: how is the machinery going to recognize which is the right nucleotide and which is the wrong one? In fact, the correction machinery recognizes the newly-synthesized strand.

In bacteria, methylation of A residues on the GATC sequence occurs some time after DNA replication. Therefore, the newly synthesized strand does not have methyl groups on the As of the GATC sequence and this difference is recognized by one of the components of the correction system (the MutH protein).

A similar mechanisms exists in eukaryotes except that the recognition of the newly synthesized strand relies on the presence of nicks on the new strand rather than on DNA methylation. A mutation in one of the genes responsible for this repair mechanism in humans is responsible for the susceptibility to certain types of cancer.

The correction mechanism in **eukaryotes** is shown on the following figure:

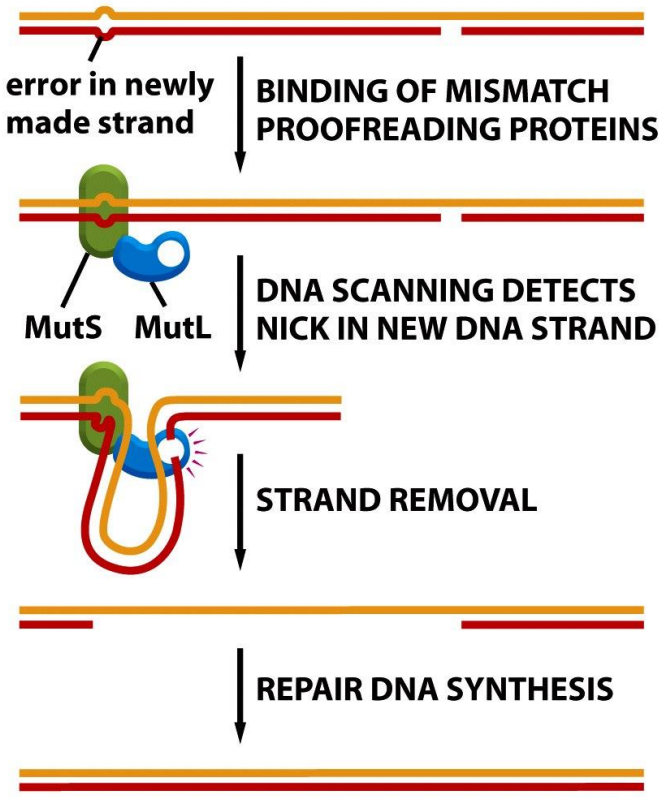


Figure 5-20a Molecular Biology of the Cell 5/e (© Garland Science 2008)

3. The replication machinery.

First of all, let's clarify something about DNA polymerase in *E. coli*. Although we are talking about DNA polymerase, there are, in fact, three DNA polymerases in *E. coli* :

a. **DNA polymerase I** : fills the gaps during DNA repair.
b. **DNA polymerase II** : Repairs a fork that is stalled because of DNA damage. Part of the SOS system (covered in the next lecture).
c. **DNA polymerase III** : is responsible for DNA replication.
d. Additional DNA polymerases exist but we won't cover them in this course.

In addition to DNA polymerase III, several proteins are necessary for the replication process

a) **Primer formation.** DNA polymerase necessitates a primer (the 3'OH end of a nucleotide paired [bound]to a DNA molecule) in order to be able to incorporate a dNTP molecule. On the **leading strand**, only one primer is necessary once a replication fork is formed. Synthesis proceeds continuously on this strand, in the 5' to 3' direction. On the lagging strand, synthesis of Okazaki fragments necessitates multiple primers. These primers are made of **RNA**, and have a size of about **10 nucleotides**. They need the action of an enzyme named **DNA primase**. The RNA primer is used by DNA polymerase to synthesize the Okazaki fragments. The RNA primers are later degraded and replaced with DNA. The Okazaki fragments are then ligated to each other by the enzyme DNA ligase (a process that necessitates ATP).

In *E. coli*, it is the exonuclease activity of DNA polymerase I that is responsible for the removal of the RNA primers. In eukaryotes, the FEN-1 (flap endonuclease) removes the RNA primers.

The following figure summarizes this process.

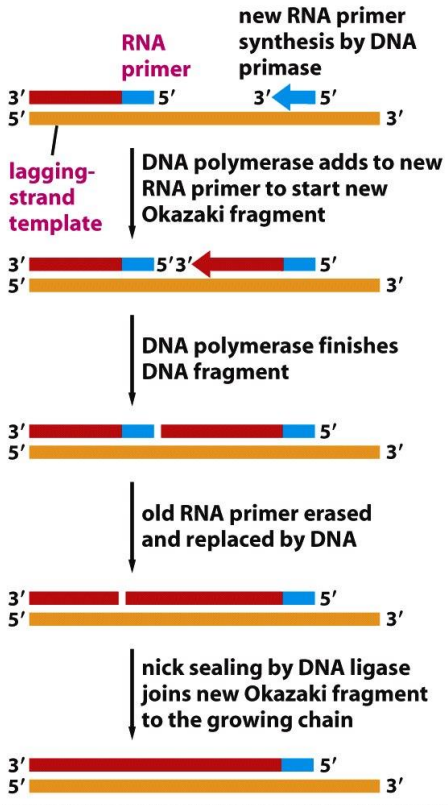
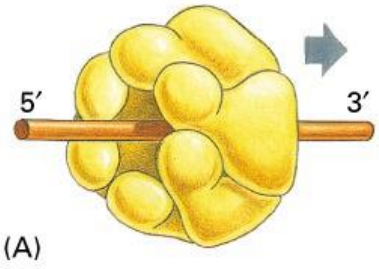


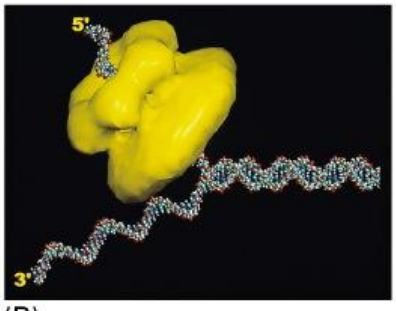
Figure 5-12 Molecular Biology of the Cell 5/e (© Garland Science 2008)

b) Opening of the double helix ahead of the replication fork.

Two types of protein are necessary to facilitate opening of the double helix: DNA helicases, and single strand binding proteins. DNA helicases use the energy provided by ATP hydrolysis to move on single strand DNA and open the double helix ahead with a speed of about 1000 nucleotide pairs per second.



(A)



(B)

Figure 5–16. Molecular Biology of

Single strand binding proteins (or *SSB proteins*) straighten the single stranded DNA in the fork and prevent the formation of hairpins.

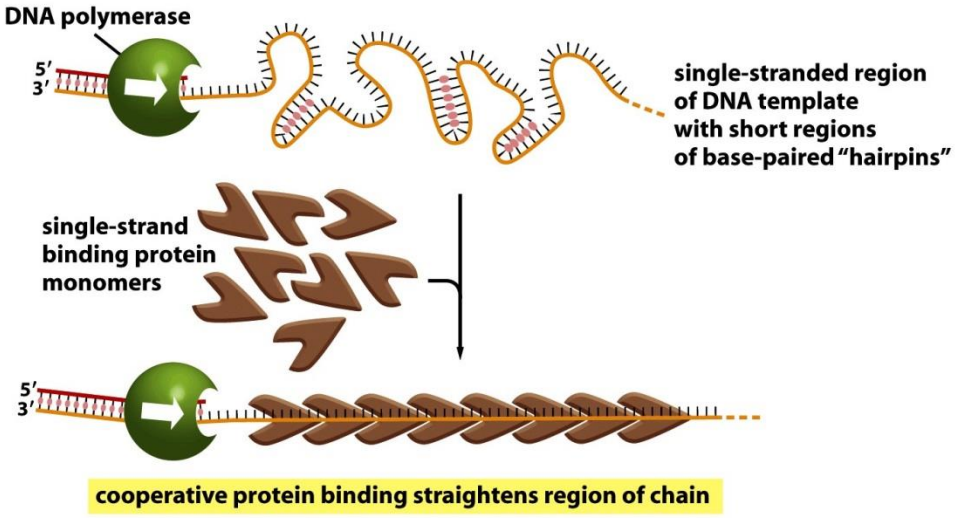


Figure 5-16 Molecular Biology of the Cell 5/e (© Garland Science 2008)

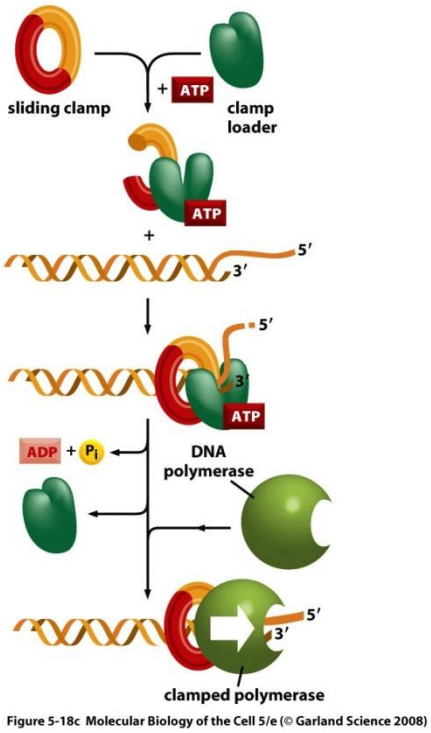


Figure 5-18c Molecular Biology of the Cell 5/e (© Garland Science 2008)

c) The « *clamp* » :

DNA polymerase has the tendency to synthesize short DNA molecules and then to rapidly dissociate itself. This property is useful when comes the time to synthesize short Okazaki fragments. However, this property is unwanted when comes the time to synthesize long DNA molecules, such as synthesis on the leading strand. DNA polymerase is assisted by the **clamp** that allows it to better stay in contact with the DNA until it meets a double strand DNA molecule (for example, the next Okazaki fragment). At this time, the polymerase stops and the clamp is released.

A **clamp loader** protein facilitates, as its name says, loading of the clamp on the DNA.

If we incorporate all the information seen so far, we obtain the following figure. Note that the association between DNA helicase and DNA primase forms a structure called the **primosome**. Also note that the lagging strand folds on itself and this gives you a more accurate view of DNA polymerase at a point where it is about to complete an Okazaki fragment and start another one.

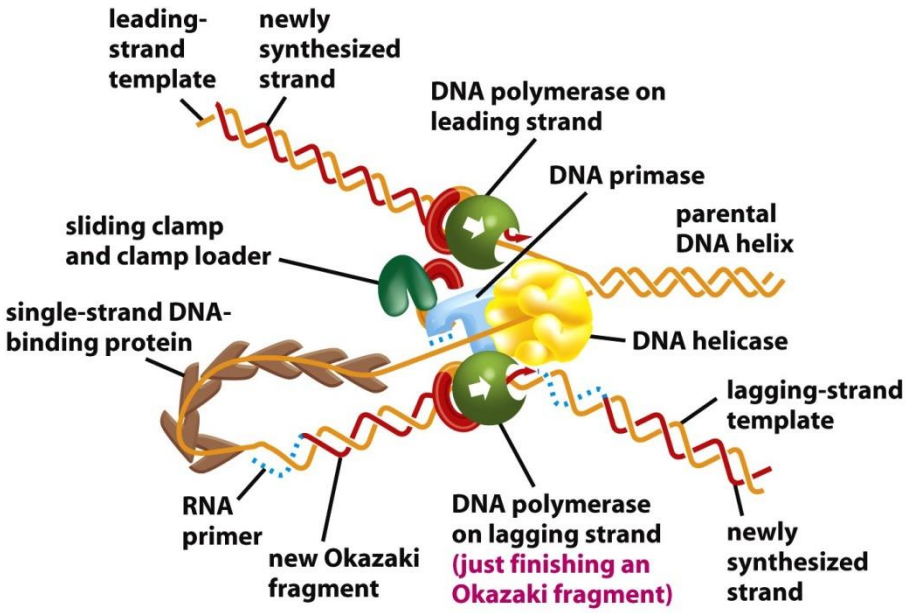


Figure 5-19a Molecular Biology of the Cell 5/e (© Garland Science 2008)

DNA topoisomerase :

The rapid movement of DNA polymerase during replication necessitates a rapid unwinding of the DNA molecule. This would require large amounts of energy. This problem is solved by the presence of topoisomerases. Topoisomerase I makes temporary breaks in the phosphodiester bonds of the DNA backbone; this breaks are rapidly repaired without a supplementary request for energy. Introduction of these breaks facilitates rotation of the DNA molecule.

4. Similarities between prokaryotes and eukaryotes

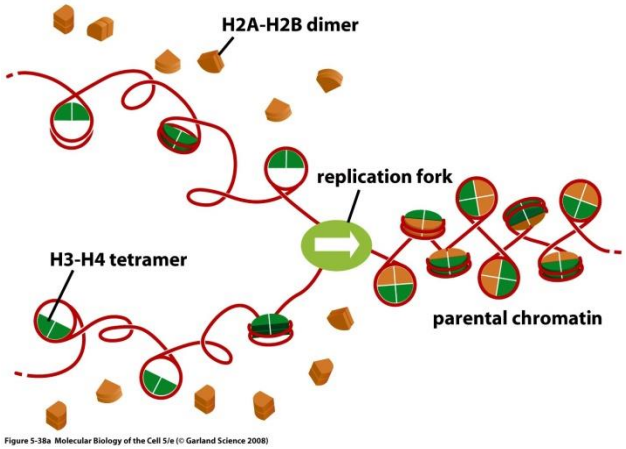
Replication mechanisms were first elucidated in bacteria, thanks to our ability to isolate mutant forms of the proteins we just described. Replication mechanisms of eukaryotes were mostly studied in yeast, again because of the possibility to obtain mutants.

Generally speaking, the replication machinery of eukaryotes involves a larger number of proteins. The SSB proteins of eukaryotes comprise three subunits whereas prokaryote SSB have one subunit. DNA primase in eukaryotes is a complex made of multiple subunits named **DNA polymerase α/primase** whose role is to begin synthesis of the Okazaki fragments before passing the hand DNA polymerases **δ** (lagging strand) and **ε** (leading strand), that carry out most of the replication. Finally, we will now see that eukaryotic DNA polymerases DNA polymerase face an additional problem in that they have to pass through nucleosomes.

5- **Histones and Replication**

DNA replication in eukaryotes faces additional challenges due to the presence of nucleosomes. These have to be made on each of the two molecules of DNA produced during one round of replication. Furthermore, as we just saw in the last few lectures, histones are often modified and these modifications are an important source of information for the cell.

During replication, H2A-H2B dimers are first released and the H3-H4 tetramers of the parent DNA molecule are distributed in a random fashion, approximately equally, on each of the two daughter DNA molecules, as shown on the figure below. Chromatin remodeling complexes (that we saw in previous lectures) destabilize interactions between histones and DNA and, thus, facilitate the work done by the replication machinery. Newly synthesized H3-H4 tetramers bind DNA to make the missing nucleosomes and H2A-H2B dimers, some of which are newly synthesized, are added afterwards.



Histone chaperones (**NAP-1** and **CAF-1** in the figure below) play a facilitating role in nucleosome formation during replication. These chaperones are targeted to the replication fork thanks to a specific interaction with the sliding clamp of eukaryotes: the **PCNA** protein.

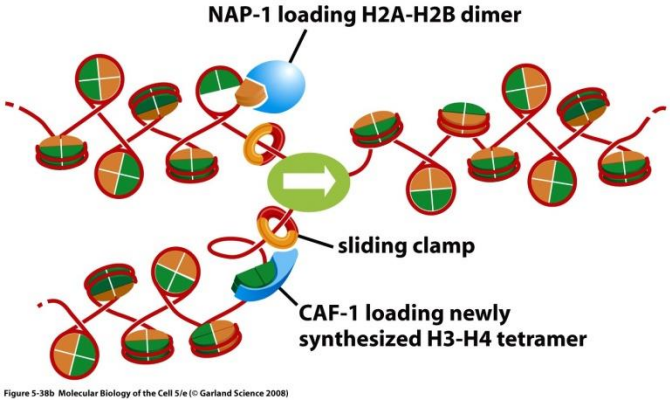


Figure 5-38b Molecular Biology of the Cell 5/e (© Garland Science 2008)

Finally, histone modifications are transmitted to the two daughter DNA molecules according to a mechanism similar to the one we saw in the last lecture when we talked about heritability of chromatin structure. This is illustrated in the figure below.

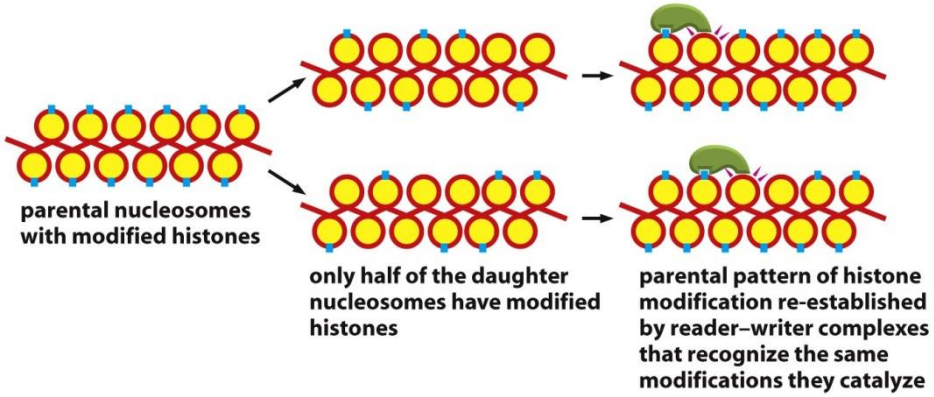


Figure 5-39 Molecular Biology of the Cell 5/e (© Garland Science 2008)