Chapter 14: DNA Replication

The only way to make new cells is by the division of pre-existing cells. This means that all organisms depend on cell division for their continued existence. DNA, as you know, carries the genetic information that each cell needs. Each time a cell divides, all of its DNA must be copied faithfully (or near perfectly) so that a copy of this information can be passed on to the daughter cell. This process is called DNA replication. Before examining the actual process of DNA replication, it is useful to think about what it takes to accomplish this task successfully. Consider the challenges facing a cell in this process:

* The sheer number of nucleotides to be copied is enormous: e.g., in human cells, on the order of several billion.
* A double-helical parental DNA molecule must be unwound to expose single strands of DNA that can serve as templates for the synthesis of new DNA strands.
* This unwinding must be accomplished without introducing significant strain into the molecule.
* The unwound single strands of DNA must be kept from coming back together and hydrogen bonding long enough for the new strands to be synthesized.
* DNA polymerases cannot begin synthesis of a new DNA strand *de novo* and require a free 3' OH to which they can add DNA nucleotides.
* The use of RNA primers requires that the RNA nucleotides must be removed and replaced with DNA nucleotides, and the resulting DNA fragments must be joined.
* DNA polymerases can only extend a strand in the 5' to 3' direction. The 5' to 3' extension of both new strands at a single replication fork means that one of the strands is made in pieces.
* Eukaryotic DNA is linear and has ends that must be replicated.
* Ensuring accuracy in the copying of so much genetic information.

Eukaryotic genomes are much more complex and larger in size than prokaryotic genomes and their genomes are often contained within multiple linear chromosomes. The human genome contains 6 billion base pairs that are replicated during the S phase of the cell cycle. Eukaryotes have multiple origins of replication on each chromosome; humans can have up to 100,000 origins of replication across the genome. Eukaryotic replication occurs at approximately 100 nucleotides per second, much slower than prokaryotic replication.

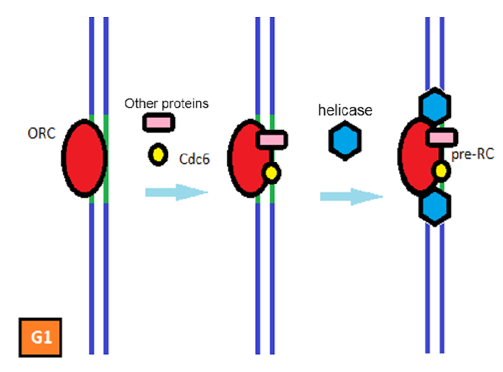
DNA polymerase is the enzyme responsible for generating a newly synthesized strand of DNA during S phase/DNA replication. The number of DNA polymerases in eukaryotes is much more than in prokaryotes: 14 are known, of which, five have major roles during S phase They are known as pol α, pol β, pol γ, pol δ, and pol ε.

# Steps of DNA replication

Before replication can start, the DNA has to be unpackaged from histones so that the DNA replication machinery can access the DNA. Histones must be removed and then replaced during the replication process, which helps to account for the lower replication rate in eukaryotes.

## Replication Initiation

A protein complex called the **Origin Recognition Complex (ORC)** binds at the origin of replication sequence in the DNA during late M phase/G1 phase of the cell cycle. The ORC interacts with a few other proteins, one known as **Cdc6** works with another protein called **Cdt1** to help load a complex of proteins known as helicase onto the DNA. (Remember that DNA helicases use ATP to unwind the DNA helix by breaking the hydrogen bonds between paired bases.) The ORC/Cdc6/Cdt1/Helicase complex along with other associated proteins is known as the pre-replicative complex or pre-RC (Figure 14-1). The pre-RC persists until the G1/S transition, which occurs once the cell receives a signal to divide. At the G1/S transition, it is thought that Cdc6 and Cdt1 unbind from the pre-RC and other proteins that are responsible for recruiting additional DNA replication proteins associate with the helicase to form the pre-initiation complex. Interestingly, the ORC is inhibitory to replication and must be removed before replication can proceed.



**Figure 14-1: Licensing of Replication.** In late M phase/early G1, origin recognition (ORC) proteins recognize the origin of replication (Ori), shown in green. These recruit many proteins, one is Cdc6, which helps recruit helicase. This establishes the pre-initiation replication complex (Pre-RC). Activation of replication occurs during early S phase, once the ORC, Cdc6, the helicase and several other proteins (not shown) are phosphorylated by the S-phase CDK/cyclin. Phosphorylation of ORC and Cdc6 cause the proteins to be deactivated, while phosphorylation of the helicase causes the protein to be activated. (Attributed to Devin A. King)

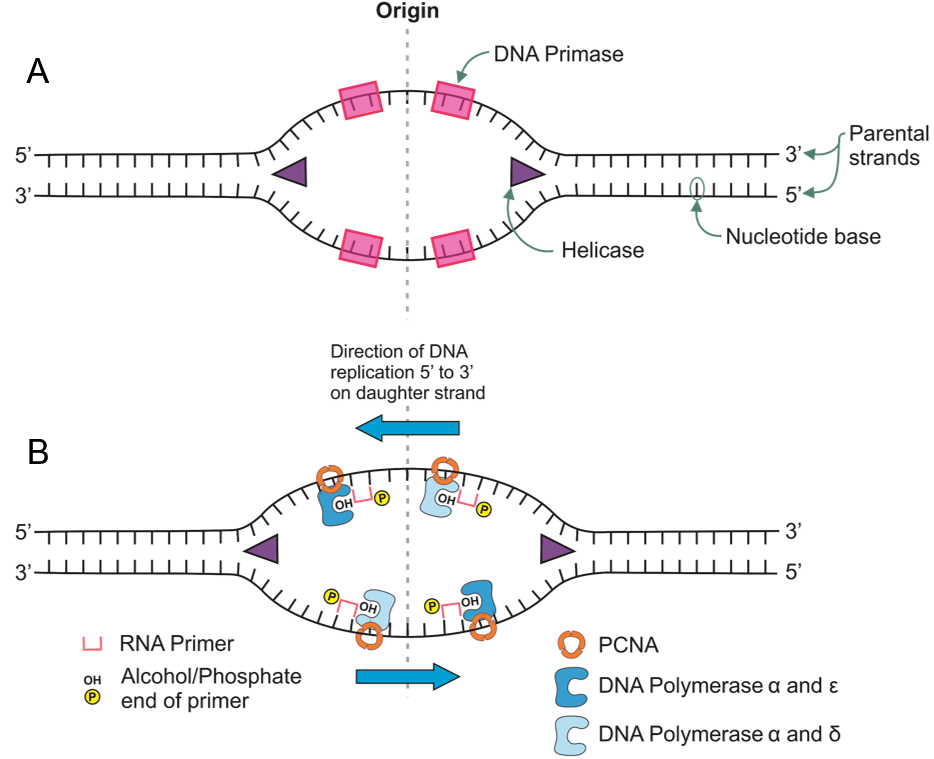
At this point, the activated S-phase CDK/cyclin phosphorylates several proteins to start or ‘fire’ replication. One of the proteins the S-phase CDK/cyclin phosphorylates is Cdc6, which helps target this protein for degradation. Additionally, the ORC is phosphorylated to prevent it from being used a second time (or rebinding the origin of replication) and phosphorylation is thought reduce the binding affinity of ORC for the origin. The helicase is also phosphorylated to activate it and encourage it to begin unwinding the DNA. Thus, these three phosphorylation events, and many others not described here, help facilitate the start of DNA replication.

The opening of the double helix by helicase causes over-winding, or supercoiling, in the DNA ahead of the replication fork. These are resolved by the action of **topoisomerases**, which cut the phosphodiester backbone to release torsional strain and then re-ligate the strand back together. Additionally, opening the helix makes the DNA single-stranded. To prevent the DNA strands from reannealing before being replicated and to protect them from chemical modifications from their environment, proteins called **single-stranded DNA binding proteins** **(SSBs)** are associated with the DNA.

## Replication Elongation

After origin ‘firing’, helicase begins to unwind the DNA helix and assembly of the replication machinery happens. Where helicase is unwinding the DNA helix is called the **replication fork**. The following proteins are involved in DNA replication and will be discussed in this section: Helicase, DNA primase, DNA polymerases α, δ and ε, PCNA, FEN-1 and RNase H2.

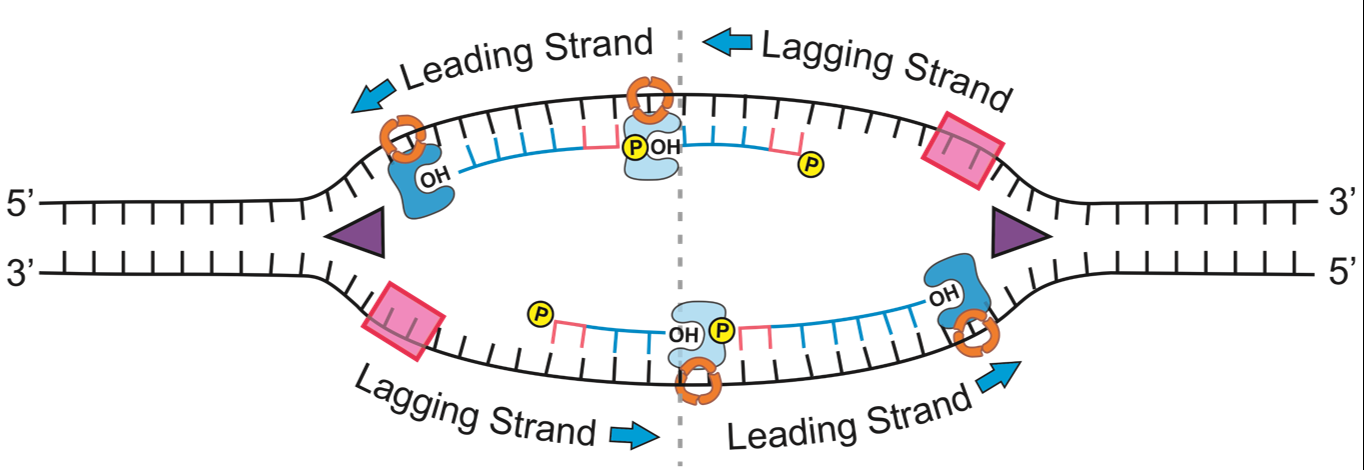
DNA polymerases are unable to make a new strand of DNA without a short nucleic acid sequence called a primer that can provide a free 3’-hydroxyl group. Alternatively, during transcription, RNA polymerases can bind single-stranded DNA molecules, ‘read’ the DNA sequence, and immediately start building an RNA molecule. Because RNA polymerases can synthesize nucleic acids *de novo* an RNA polymerase called **DNA** **primase** binds the single stranded DNA and adds a short (~35 nt) RNA sequence on both strands of template DNA on either side of the replication origin (Figure 14-2 A & B). Then DNA polymerases can start synthesis of DNA molecules (Figure 14-2 B). Three major DNA polymerases are then involved: α, δ and ε. DNA pol α adds a short (20 to 30 nts) DNA fragment to the RNA primer on both strands, and then hands off synthesis to a second polymerase either DNA pol δ or DNA pol ε. A protein called **Proliferating Cell Nuclear Antigen** (PCNA) ‘clamps’ around the DNA and interacts with polymerase to help stabilize and maintain polymerase’s interaction with DNA. DNA replication will proceed in both directions, away from the origin.



**Figure 14-2: The start of DNA replication.** (A) After helicase (purple triangle) begins to unwind the DNA at the origin, DNA primase (pink box) can recognize the single-stranded DNA and begin to add a short RNA primer (shown in B). Note that it is estimated that DNA primase will add about 35 RNA nucleotides; this illustration is representing this addition using only 2 nucleotides (pink staple-each vertical line is one nucleotide). DNA polymerase (in blue) α will add nucleotides to the 3’-OH end of the primer and then ‘hand off’ the polymerization to either DNA polymerase δ or DNA polymerase ε. Polymerization will occur on the daughter strand in the 5’ to 3’ direction. On the top parental strand, this means polymerization will occur right to left and on the bottom parental strand it will occur left to right. PCNA (orange ring) is a protein that helps to keep DNA polymerase associated with the DNA template. (Illustration by Christopher G. Brown)

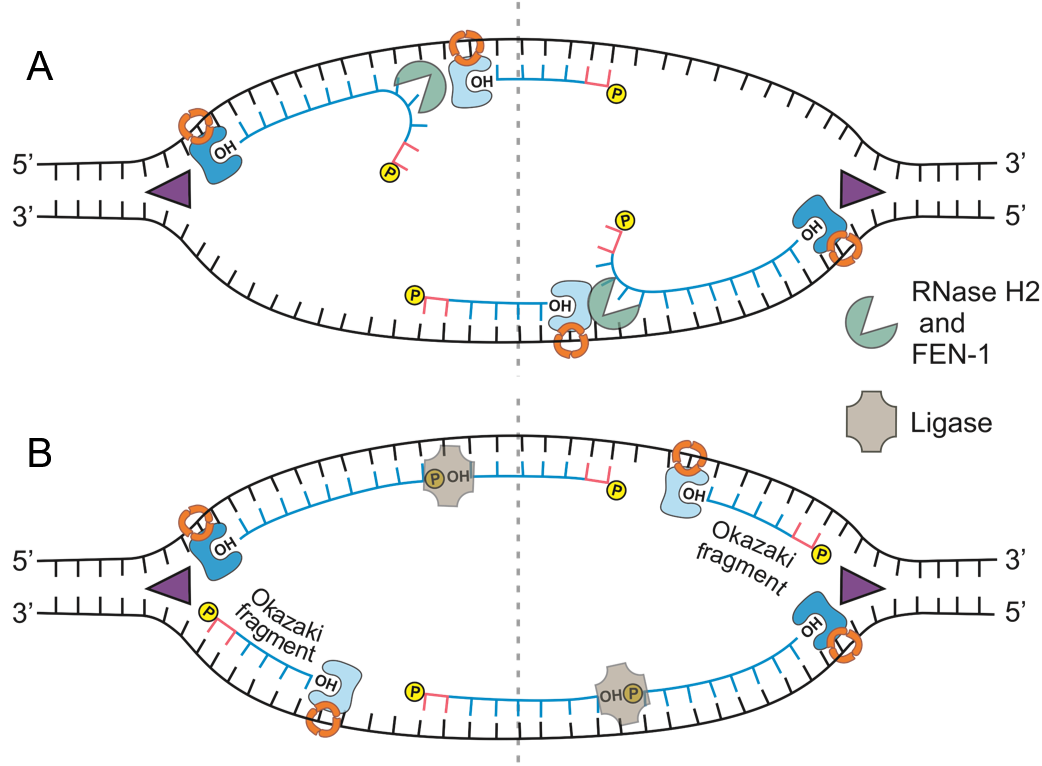
Notice in figure 14-2B that all of the DNA polymerases bind to the 3’-OH group on the RNA primer to initiation DNA replication.

As the DNA polymerases elongate from the RNA primer (Figure 14-3) you will notice that DNA polymerases that are darker blue can replicate the daughter strands without ‘running into’ another structure. As long as the helicase continues to unwind the DNA helix this polymerase complex can proceed in the 5’ to 3’ direction without stopping. Daughter strands that are synthesized in the same direction the replication fork is progressing are made in a continuous fashion and are termed **leading strands**. Alternatively, the DNA polymerases represented by the lighter blue color will ‘run into’ the previous RNA primer. These daughter strands are still being replicated 5’ to 3’ but they are not being replicated in the same direction as the replication fork that is nearest them AND they are replicated in shorter fragments. This daughter strand is called the **lagging strand**.



**Figure 14-3: Elongation of DNA replication part 2**. The DNA polymerases in dark blue (α and ε) replicate DNA in the same direction the replication fork is moving, called leading strand synthesis. The DNA polymerases in light blue (α and δ) replicate DNA away from their nearest replication fork in short fragments, called lagging strand synthesis. Lagging strand polymerases will ‘bump’ into neighboring RNA primers. Notice the return of DNA primase (pink box), which must return to the lagging strand to lay down a new primer so a new 3’OH will be available for replication. (Illustration by Christopher G. Brown)

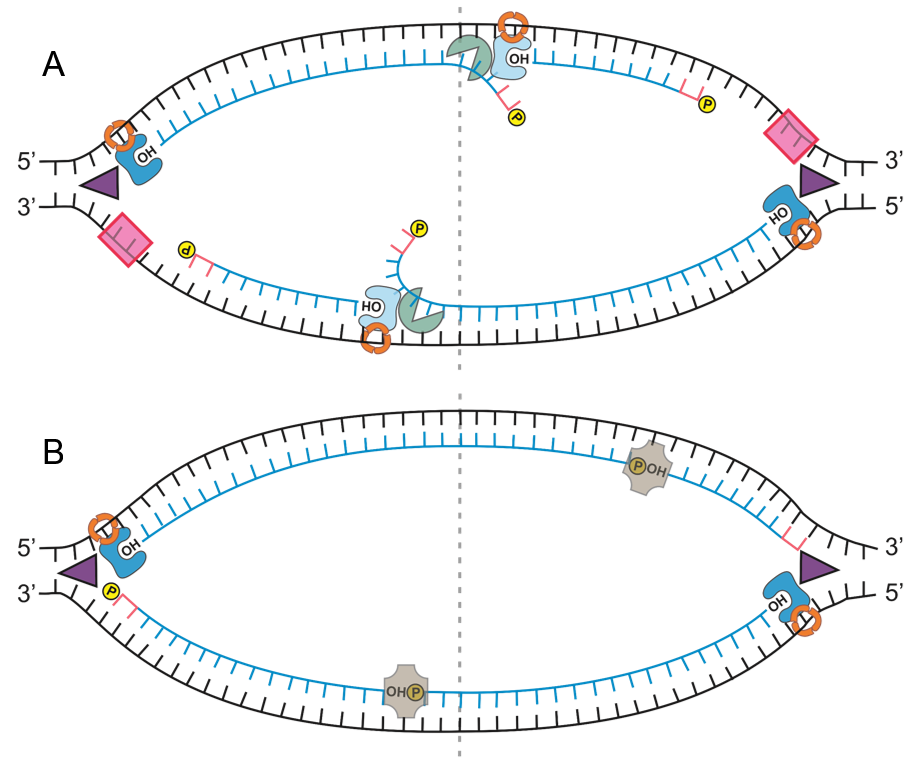
Once the lagging strand polymerase complex ‘bumps’ into the neighboring RNA molecule, it will begin to displace the neighbor by ‘peeling’ back the nucleotides and re-replicating this portion of the DNA (Figure 14-4 A). The peeling back causes a small ‘flap’ to form, special enzymes called flap endonucleases come in and cut the phosphodiester backbone near the RNA/DNA boundary. The two enzymes thought to be most responsible for this action are called **RNase H2** and **FEN-1** (flap endonuclease 1). Once the flap is removed, the lagging strand polymerases will fill in any remaining DNA sequence. It will then disassociate with the DNA leaving a ‘nick’ in the DNA backbone. An enzyme called DNA ligase will bind and chemically connect the free 3’-OH with the free 5’-phosphate group ‘sealing’ the backbone (Figure 14-4 B).



**Figure 14-4: Process of removing RNA primers and ligating DNA strands.** (A) Lagging strand DNA polymerases will run into neighboring RNA primers, displace them, and flap endonucleases (RNase H2 and FEN-1-teal PacMan™) will remove them. The DNA polymerases will replicate any remaining nucleotides and then disassociate with the DNA leaving a nicked backbone. (B) DNA Ligase (shown in tan) will connect the free 3’-hydroxyl and 5’-phosphate groups fixing the broken backbone. The short RNA/DNA fragments generated on the lagging strand are called Okazaki fragments. (Illustration by Christopher G. Brown)

By repeating this process on the lagging strand, the small RNA/DNA fragments (called **Okazaki fragments**) are pieced together to make one long daughter strand of DNA free of any RNA nucleotides (Figure 14-5).

By the end of S-phase, the cell will have completely and faithfully replicated the entire genome. Let’s consider a single chromosome for a moment. Prior to S-phase a single chromosome consists of two DNA strands, held together by hydrogen bonds in a single DNA helix. By the end of S-phase the chromosome consists of four DNA strands that are evenly split between two identical structures called **chromatids.** Because the chromatids are identical, they are called **sister chromatids.** How are these sisters actually held together? They are not hydrogen bound together like the individual helices, instead a special protein called **cohesin** literally “links” the two sister chromatids together. Cohesin remains bound to both sister chromatids until M-phase when a special enzyme called **separase** degrades it, allowing the sisters to separate. For more detailed information refer back to chapter 13.



**Figure 14-5: Repeated removal of RNA primers and action of ligase.** (A) The cycle from figure 14-4 continues until both parental strands of DNA have been replicated. (B) Notice as this process continues the daughter strand is becoming one long piece of DNA. The final RNA primers at the very ends of the lagging strand will be removed in a different process (not discussed here). (Illustration by Christopher G. Brown)

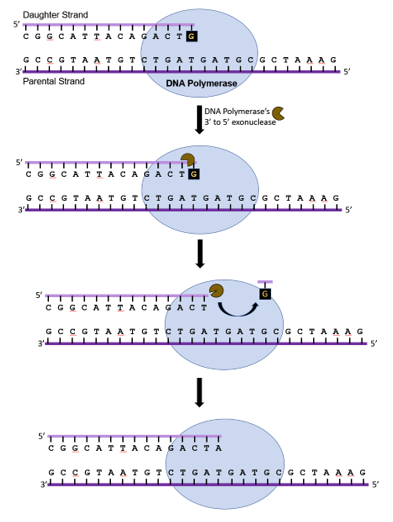
**Epigenetic patterning is retained after replication**

Methyl groups can be added to cytosine nucleotides as a mechanism of controlling what genes are transcribed in cells, one form of epigenetic patterning. This is often important in the identity and proper functioning of the cell. Therefore, it is important to ensure that any epigenetic patterning is retained after replication in the daughter strand. Because the cytosines in the parental strands will retain any methylation, DNA methyltransferases that recognize methylation of the parental strand will be able to methylate appropriate cytosines in daughter strand. Refer to figure 5-8 in chapter 5 to connect these ideas.

**Errors in DNA replication and proofreading**

DNA replication is a highly accurate process, but mistakes can occasionally occur, such as a DNA polymerase inserting a wrong base.

Most of the mistakes during DNA replication are promptly corrected by the proofreading ability of DNA polymerase itself (Figure 14-6). In proofreading, the DNA pol δ or ε reads the newly added base before adding the next one, so a correction can be made. The polymerase checks whether the newly added base has paired correctly with the base in the template strand. If it is the right base, the next nucleotide is added. If an incorrect base has been added, the enzyme makes a cut at the phosphodiester bond and releases the wrong nucleotide on the daughter strand. This is performed by the 3' exonuclease action of DNA pol δ or ε. Once the incorrect nucleotide has been removed, it can be replaced by the correct one. Note that DNA pol α does not have proofreading ability because it does not contain the 3’ exonuclease activity.



**Figure 14-6:  Proofreading by DNA polymerase.** DNA polymerases delta and epsilon contain a 3’ to 5’ exonuclease enzyme that allows the removal of improperly paired nucleotides (nt). The enzyme recognizes the mismatch, slows to allow the 3’ to 5’ exonuclease to interact with the improperly paired nt. It cleaves the backbone of the daughter strand to remove the wrong base and then the 5’ to 3’ replication machinery adds the correct nucleotide. (Illustration by Jennell Talley.)

## Telomere replication

Unlike prokaryotic chromosomes, eukaryotic chromosomes are linear. As you’ve learned, the enzyme DNA polymerase can add nucleotides only in the 5' to 3' direction. Leading strand synthesis continues until the end of the chromosome is reached. On the lagging strand, DNA is synthesized in short stretches, each of which is initiated by a separate primer. When the replication fork reaches the end of the linear chromosome, there is no way to replace the primer on the 5’ end of the lagging strand leaving a structure called a 3’-overhang or a 5’-gap (Fig 14-7). Interestingly, this overhang is critically important in forming a ‘cap-like’ structure at the end of the chromosome called the telomere. The telomere has a highly repetitive sequence that is recognized by proteins that bind the DNA sequence and ‘hide’ the linear end of the chromosome by forming a loop. This structure is important because it is telling the cell “Hey, I’m the normal end of the chromosome-leave me alone; don’t try and fix me!” Because the cell needs this 3’-overhang to create the correct end structure, the leading strand is resected (or chewed up) by an exonuclease to create a 3’-overhang. After each round of replication, the leading strands are continually shortened (the lagging strands also gradually shorten due to removal of the last RNA primer, but this shortening is actually less detrimental compared to the degradation on the leading strands). If the shortening at the ends of the chromosomes was not fixed, eventually the highly repetitive sequence of the telomere would be lost and the cap structure at the ends of the chromosomes would no longer form, causing the cell to signal that the DNA was damaged.

To combat the loss of DNA sequence in the telomere and prevent a DNA damage response, an enzyme called telomerase ([Figure](https://openstax.org/books/biology-2e/pages/14-5-dna-replication-in-eukaryotes#fig-ch14_05_02) 14-7), adds nucleotides to the ends of chromosomes. Telomerase is highly expressed in early development and gametic cells; it is moderately expressed in many adult stem cells, but not expressed in most adult somatic cells. Telomerase is an enzyme that contains a catalytic part and a built-in RNA template. Because it uses an RNA template to build DNA (transcription in reverse), it is called a reverse transcriptase. Telomerase attaches to the 3’-overhang (location of a 3’OH group) of the chromosome, and adds DNA nucleotides that are complementary to the RNA template. Once the 3' end of the template is sufficiently elongated, DNA polymerase can add a new primer and fill in the nucleotides complementary to the ends of the chromosomes (Figure 14-7). The final RNA primer will be removed, re-establishing the 3’-overhang and allowing the cap structure to be formed once again.

Telomeres shorten as cells age due to the routine loss of sequence during each round of cell division. However in many cancer cell types, the expression of the telomerase activity has been reactivated, allowing cancer cells to lengthen telomeres and become immortal. Because of this action, there is great interest in better understanding how telomerase and the telomere are regulated in cells and if there are ways to exogenously control the enzyme.



**Figure 14-7: Telomerase.**  The ends of linear chromosomes are maintained by the action of telomerase. Briefly, Telomerase attaches to the 3’-overhang of the chromosome, and adds complementary DNA nucleotides to the RNA template. Once the 3' end of the template is sufficiently elongated, DNA polymerase adds a new primer and fills in the nucleotides complementary to the ends of the chromosomes.

# DNA Damage and Repair

We generally accept the notion that replication duplicates genetic material faithfully. At the same time, evolution would not be possible without mutations, and mutations are not possible without at least some adverse consequences.

Since the complex chemistry of replication is subject to errors, cells have evolved systems of DNA repair to survive and fix mutations. As we saw, DNA polymerases have proofreading ability so that incorrectly inserted nucleotides in the daughter strand can be quickly removed by exonuclease activity and repaired using the parent strand by the polymerase itself. Beyond this, multiple mechanisms have evolved to repair mismatched base pairs and other kinds of damaged DNA that escape early detection. For small mutations (single point mutations) the strand without the mutation will be used to help repair the DNA. For larger mutations (large deletions or breaks in the backbone), a homologous strand of DNA is often used to help repair the damage. How often and where DNA damage occurs is random, as is which damage will be repaired and which will escape detection to become a mutation.

## Causes of DNA Damage

The simplest damage to DNA during replication is called a **point mutation**, which is the result of the accidental insertion of a ‘wrong’ nucleotide into a growing DNA strand. Other mutations include DNA deletions, duplications, and inversions, any of which might escape repair. The causes of DNA damage can be chemical or physical and can be the result of spontaneous intracellular events (e.g., oxidative reactions) or environmental factors (radiation, exogenous chemicals, etc.). Based on studies of different kinds of DNA damage, it is estimated that DNA damaging events might be occurring at the rate of 10,000 per day; therefore there must be important DNA repair pathways at work to protect cells against such a high rate of DNA damage. Environmental factors that can damage DNA include UV light, X-rays and other radiation, as well as chemicals (e.g., toxins, carcinogens, and drugs.). Both germline (sex cells) and somatic (non-sex) cells can be affected. While mutations can and do cause debilitating diseases, most mutations are likely *silent*; they do not cause disease, or we do not detect any effect associated with the mutation. Besides, most DNA damage is repaired. Cells correct more than 99.9% of mistaken base changes before they have a chance to become mutations. That is why we think of replication as a “faithful” process. Let’s look at some common types of DNA damage that are usually repaired:

* **Depurination**- the hydrolytic removal of guanine or adenine base from the sugar (deoxyribose) in a DNA strand. This base removal leaves the corresponding sugar and phosphate of the backbone present in the strand.
* **Pyrimidine dimers (thymine dimers)**- formed when neighboring pyrimidines (in the same DNA strand) become covalently linked due to UV exposure. This is more commonly found between two thymine bases.
* **Deamination**- the hydrolytic removal of amino (-NH2) groups from guanine (most common), cytosine or adenine. This mutation changes the chemical structure, and therefore changes the base attached to the backbone.
* Oxidative damage- oxidation of any base, but most commonly purines. This damage is most prevalent in mitochondria where the majority of the highly oxidative pathways of cell respiration occur.
* Inappropriate methylation- addition of a methyl group to any bases, but most commonly purines.
* Break in the phosphodiester backbone- can occur during replication or from radiation or chemical exposure, and can be either a single-stranded or double-stranded break

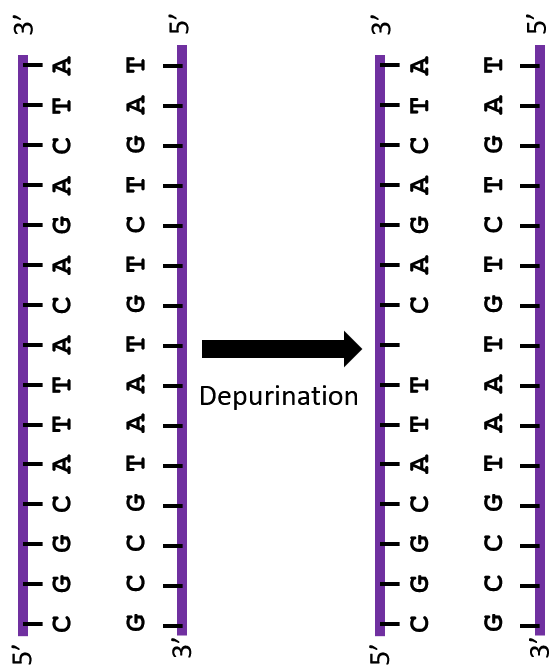
## Consequences of Uncorrected DNA Damage

While bacteria suffer DNA damage, the discussion here will focus here on eukaryotes since they have evolved the most sophisticated DNA repair mechanisms. Remember that unrepaired DNA damage will be passed onto daughter cells in mitosis, or might be passed on to the next generation if the mutation occurs in a germline cell.

Next, let us consider some*molecular consequences* of uncorrected DNA damage.

**Depurination**

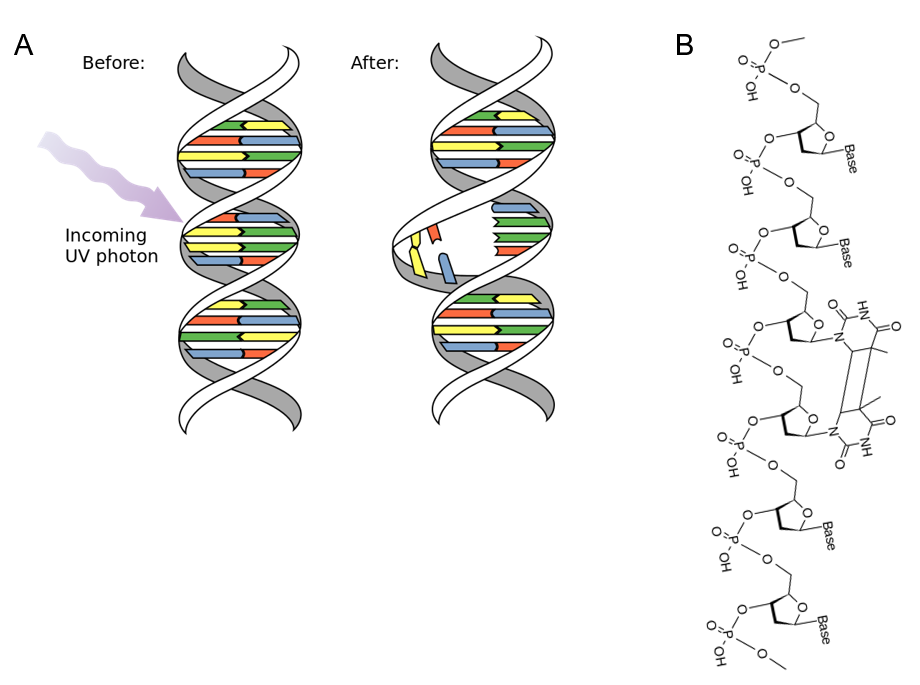
This is the spontaneous, hydrolytic removal of guanine or adenine base from the sugar in a DNA strand (Figure 14-8). If not repaired, depurination typically results in a single base-pair mutation after the following round of replication. One strand will be replicated normally because its base remains, but when the daughter strand of the DNA without a base (due to depurination) is replicated, any deoxyribonucleotide can be added because there is no base to pair properly with (often the nucleotide contains an adenine base). If the added base is not the correct base then there will be a point mutation after replication of the affected strand.



**Figure 14-8: Depurination**. The spontaneous breaking of the glycosidic bond between the purine base and the deoxyribose sugar forms an abasic site. Note the backbone phosphodiester bond is unharmed in this type of damage. (Illustration by Jennell Talley).

**Pyrimidine Dimers**

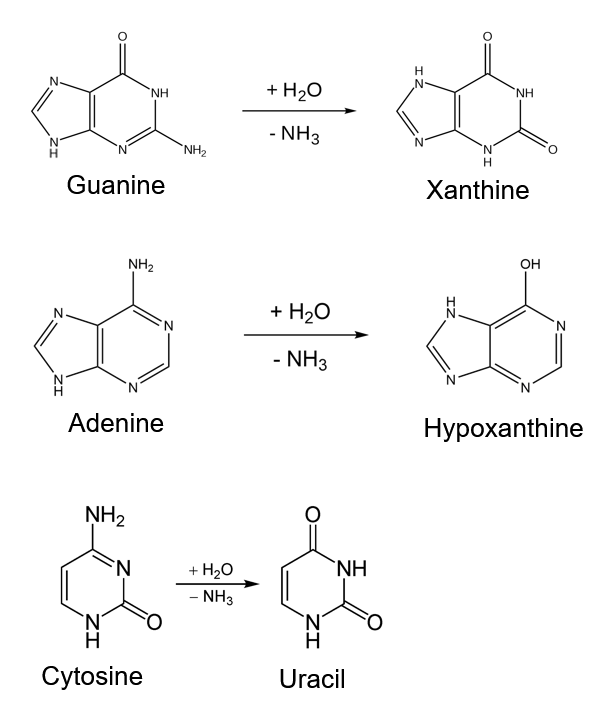
UV light exposure of DNA can cause adjacent pyrimidines (commonly thymines; less often, cytosines) on a single DNA strand to dimerize (Figure 14-9). If these go undetected, they can inhibit proper functioning of DNA polymerase, either causing the polymerase to misread the base or cause replication to stop due to distortion of the DNA backbone.



**Figure 14-9: Pyrimidine Dimers.** Pyrimidine dimers are formed when UV light excites electrons and they cause covalent bonds to form between neighboring pyrimidine structures. (A) Pyrimidine nucleotides are shown in yellow. Notice that the bonding happens between pyrimidines on the same strand of DNA, causing a bulging of the DNA backbone. (B) The chemical structure of an intra-strand pyrimidine dimer illustrating the covalent connections between two thymine nucleotides.

**Deamination**

Deamination (Figure 14-10) is the hydrolytic removal of amine (-NH2) groups from guanine (most common), cytosine or adenine. Deamination does not affect thymine (because it has no –amine groups). Uncorrected deamination results in a base substitution on one chromosome and no change on the other. Deamination of adenine or guanine results in unnatural bases (hypoxanthine and xanthine, respectively). These are easily recognized and corrected by DNA repair systems.



**Figure 14-10: Deamination.** Hydrolysis of an amine group from guanine forms xanthine, from adenine forms hypoxanthine, and from cytosine forms uracil.

## DNA Repair Mechanisms

Many enzymes and proteins are involved in DNA repair. Some of these function in normal replication, mitosis, and meiosis, but were co-opted for DNA repair activities. Among different DNA repair pathways that have been identified, we will look at **Base Excision Repair (BER)**, **Nucleotide Excision Repair (NER)**, **Non-homologous End-Joining (NHEJ)**, and **Homologous Recombination Repair (HRR)**.

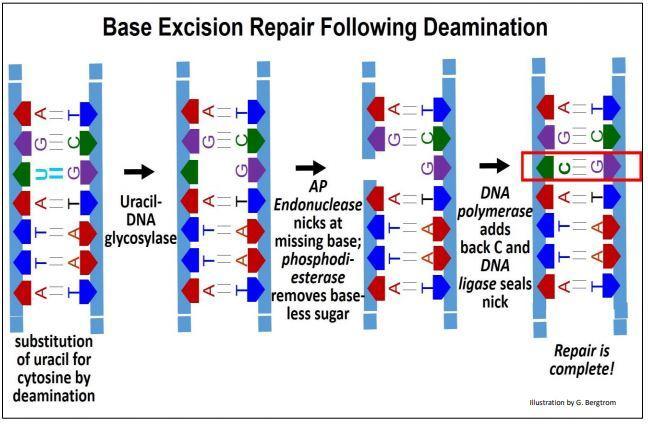
DNA repair typically has three major steps:

* Detecting the mutation (typically repair proteins detect a structural change in the DNA, such as distortion of the DNA backbone)
* Removing the mutation (typically by cutting the sequence out: endo- or exonuclease activity)
* Replacing the mutation (usually re-replicating the sequence and ligating the backbone)

The proteins involved in repair will often work together to accomplish the steps above. As you learn, some of these repair pathways try to spot how the pathways are similar.

**Base Excision Repair**

Upon detection and recognition of a damaged base, specific DNA glycosylases catalyze hydrolysis of the offending base (ex. Oxidized base, methylated base, or abasic site) from affected deoxyribose in the DNA. This type of repair is to remove and replace a single nucleotide. Events of base excision repairare summarized in Figure 14-11.



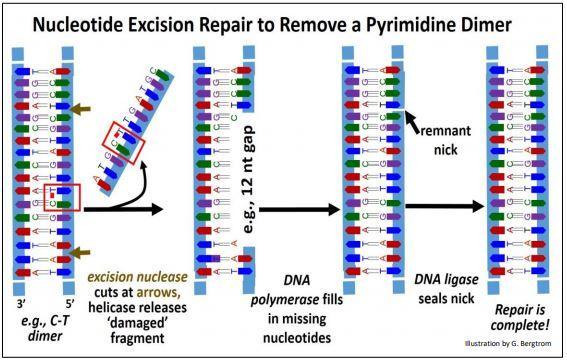
**Figure 14-11: Base Excision Repair (BER).** A damaged base is recognized by repair proteins. AP endonuclease cuts the phosphodiester backbone and then phosphodiesterase removes the rest of the nucleotide (sugar and phosphate). DNA polymerase re-replicates the single nucleotide. Ligase seals the backbone back together.

After a DNA glycosylase removes an offending base, an AP (Apurinic/Apyrimidinic) endonuclease recognizes the deoxyribose with the missing base, and nicks the DNA at that nucleotide. Phosphodiesterase next hydrolyzes the remaining phosphate-ester bond of ‘base-less’ sugar-phosphate, removing it from the DNA strand. DNA polymerase (beta or lambda) then adds the correct nucleotide to the 3’ end of the nick. Finally, DNA ligase III seals the remaining nick in the strand.

**Nucleotide Excision Repair**

The cell uses NER to repair larger (bulkier) mutations, for example, thymidine dimers. The events of nucleotide excision repair are shown in Figure 14-12 for a pyrimidine dimer.

In this example, an Excision Nuclease recognizes a pyrimidine dimer and hydrolyzes phosphodiester bonds between nucleotides several bases away from the mutation. A DNA helicasethen unwinds and separates the DNA fragment containing the dimerized bases from the damaged DNA strand. Finally, DNA polymerase acts 5’-3’ to fill in the gap and DNA ligase seals the remaining nick to complete the repair. This repair pathway is complex and it is thought that three different DNA polymerases (delta, epsilon, and kappa) may be working together to perform NER.

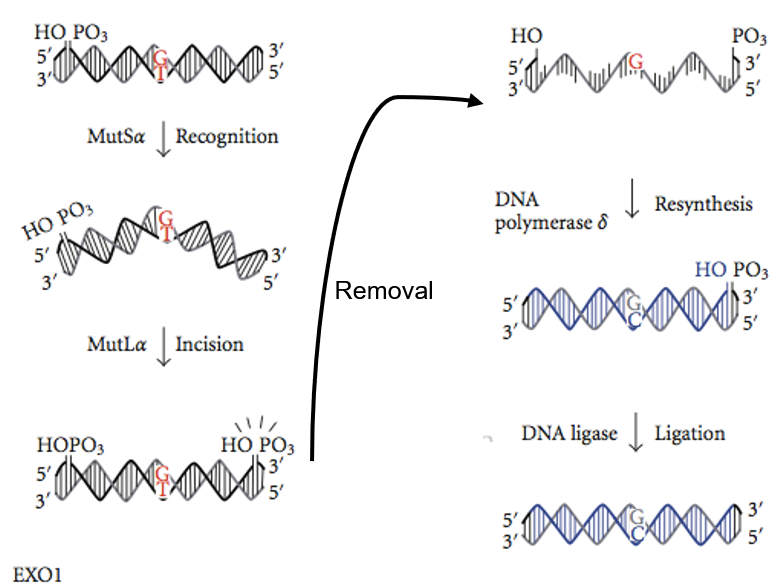


**Figure 14-12: Nucleotide Excision Repair (NER).** Repair proteins recognize a lesion (mutation) in the DNA. Excision nuclease cuts the phosphodiester backbone on either side of the mutation. Helicase unwinds the DNA allowing the damaged DNA to diffuse away. DNA polymerase re-replicates the gap. DNA ligase seals the backbone back together.

**Mismatch Repair**

**DNA Mismatch Repair**occurs when DNA polymerase proofreading misses an incorrect base insertion into a new DNA strand. This repair mechanism relies on the fact that double-stranded DNA shows a specific pattern of methylation. These methylation patterns are related to epigenetic patterns of gene activity and chromosome structure that are expected to be inherited by daughter cells. When DNA replicates, the methyl groups on the template DNA strands remain, but the newly synthesized DNA is unmethylated. In fact, it will take some time for methylation enzymes to locate and methylate the appropriate nucleotides in the new DNA. In the intervening time, several proteins and enzymes can detect inappropriate base pairing (the mismatches) and initiate mismatch repair.

MutSα detects the mismatch, while MutLα uses endonuclease activity to break the phosphodiester backbone on the 5’ side of the mutation. Exo1 may be involved in breaking the backbone on the 3’ side of the mutation. This then allows this section, containing the mutation to be removed by the action of a helicase, which unwinds the DNA strands releasing the mutated strand of DNA. Next, DNA polymerase 𝜹 replicates the DNA off of the parental strand and ligase fixes the nick in the backbone. The process is illustrated in Figure 14-13.

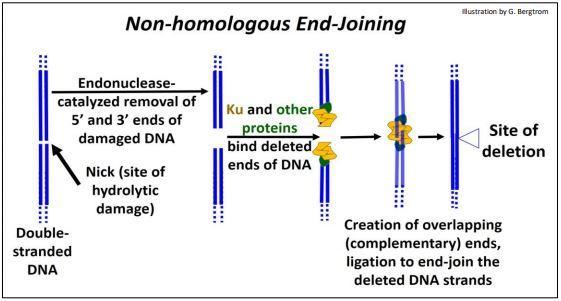


**Figure 14-13: Mismatch Repair.** The repair proteins MutSα detects the mismatch, MutLα uses endonuclease activity to break the phosphodiester backbone on the 5’ side of the mutation, it is thought Exo1 may be involved in breaking the backbone on the 3’ side of the mutation and then a helicase unwinds the DNA strands releasing the mutated strand of DNA. DNA polymerase δ replicates the DNA using the parental strand as a template and ligase fixes the nick in the backbone.

**Non-homologous End-Joining**

DNA replication errors can cause double-stranded breaks, as can environmental factors (ionizing radiation, oxidation, etc.). Repair by non-homologous end-joining (NHEJ) fixes damage due to breaks in the phosphodiester backbone, by deleting damaged and adjacent DNA and rejoining the ‘cut’ ends as shown in Figure 14-14.

Once the site of a double-stranded break is recognized, nucleotides are removed from the ends of both strands at the break-site leaving ‘blunt ends’. Next, several proteins bring DNA strands together and further cleave single DNA strands to create staggered (overlapping, or complementary) ends. The overlapping ends of these DNA strands form hydrogen bonds. Finally, DNA ligase seals the hydrogen-bonded overlapping ends of DNA strands, leaving the repaired bases. In older people, there is evidence of more than 2000 ‘footprints’ of this kind of repair per cell. How is this possible? This*quick-fix* repair often works with no ill effects because most of the eukaryotic genome does not encode genes or even regulatory DNA (whose damage would otherwise be more serious). If the bases that are deleted come from regions of the DNA that code for genes or are in regulatory regions of genes, there can be deleterious consequences.



**Figure 14-14: Non-homologous End-joining (NHEJ).** A double-strand break is recognized and the ends are resected (chewed back) and made to be blunt-ended. A complex of proteins recognizes the break and helps to bring them together (note any deleted nucleotides are now permanently lost from the DNA sequence). If any replication occurs, it is typically done by polymerase δ. Ligase will join the two ends together.

**Homologous Recombination**

**Homologous recombination (HRR)** is a complex but normal and frequent part of meiosis in eukaryotes. You may recall that homologous recombination occurs in the first cell division of meiosis (Meiosis I). During prophase I homologous chromosomes align. The DNA backbone is broken to encourage an exchange of alleles, and then ligation occurs to reseal the now recombinant DNA molecules. Novel recombinations of variant alleles in the chromosomes of sperm and eggs ensure genetic diversity in species. The key point is that DNA breakage of DNA is required to exchange alleles between homologous chromosomes.

Cells use the same proteins to repair DNA damaged by double-stranded breaks that operates on sister chromatids in meiosis. In both cases, the process accurately repairs damaged DNA without any deletions, resulting in accurate repair of the DNA and less damage compared with NHEJ. These mechanisms are conserved in the cells of all species. This indicates an evolutionary imperative of accurate repair to the survival of species, no less than the imperative to maintain genetic diversity of species.

**Repair of a Double-Stranded Break using Synthesis-Dependent Strand Annealing**

Homologous recombination is carried out by a number of proteins whose activities are coordinated. Figure 14-15 shows this process. First, some proteins recognize the double-strand break and recruit exonucleases that resect, or remove nucleotides from the broken ends. It might seem strange that the cell is deleting some nucleotides, but this process ensures repair of the break without losing DNA sequence (unlike in the results of Non-Homologous End Joining). Temporarily deleting bases helps to create a DNA overhang, the 3’-DNA overhang is single-stranded, meaning the nucleotide bases are unpaired. This allows other proteins to bind to the overhang and they use the sequence of the overhang to “search” for a complementary sequence to use to repair itself. The complementary sequence is often found on the neighboring sister chromatid, which would only be present after S phase), or perhaps from the homologous chromosome. Once the complementary sequence is found, proteins will “invade” the homologous sequence to allow the overhang to base pair with the complementary sequence. This causes the displacement of the non-complementary strand and the formation of a structure called a “d-loop” (d for displacement). The 3’-end is now aligned with a complementary sequence and DNA polymerase will add complementary bases to the 3’-overhang, replacing the bases that were deleted earlier. Hundreds of bases can be added to the 3’-overhang, eventually providing enough DNA sequence that the overhang can base pair with its original complementary DNA strand. The DNA missing from the complementary sequence can now be filled in by DNA polymerase, and ligase can connect the free 3’ hydroxyl and 5’ phosphate groups. The DNA is now completely repaired, without deleting any genetic information!

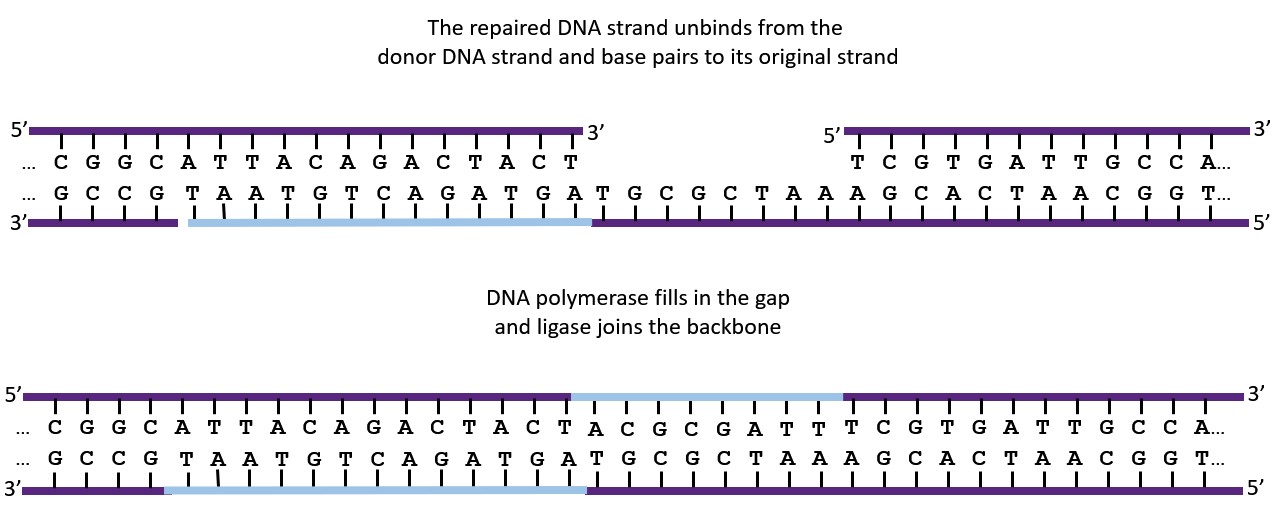
This version of homologous recombination repair is known as synthesis-dependent strand annealing (SDSA) and is the simplest; there are a few other ways the cell can use homologous recombination to repair double strand breaks, but we will not discuss them here.

Interestingly, there are two proteins involved in homologous recombination repair that you may have heard of before. These are products of the BRCA1 and BRCA2 genes (the same ones that when mutated, increase the likelihood of developing breast cancer). BRCA1 and BRCA2 are expressed mainly in breast tissue, their wild-type (normal) gene products participate in homologous recombination repair of double-stranded DNA breaks. Therefore, when they are mutated DNA repair is decreased, which causes an increase in DNA damage in breast cells.

This is the likely basis of the increased chance of developing breast cancer associated with loss-of-function mutations in BRCA1 and BRCA2. Because these genes are important in maintaining genome stability and therefore prevent tumor formation, these genes are examples of tumor suppressors (refer to chapter 13 for more information about tumor suppressors).

Figure 14-15: Repair of a double-strand break by HRR. 
This is known as resection and allows for the formation of a 3’-single-stranded overhang. Other proteins (yellow triangles) recognize and bind the overhang sequence. These proteins “guide” the overhang to a region of homologous DNA sequence on a different DNA strand, usually a sister chromatid (after S phase) or a homologous chromosome.Figure 14-15: Repair of a double-strand break by HRR. 
The overhang “invades” the donor DNA helix and displaces a strand of DNA, forming a structure called a d-loop. Figure 14-15: Repair of a double-strand break by HRR. 
The thick black arrow identifies where a double-strand (ds) break has occurred. First proteins (blue spheres) recognize the break in the backbone of the DNA, these help to recruit other proteins known as exonucleases (brown Pac Man) to remove some nucleotides from the 5’ ends.

Figure 14-15: Repair of a double-strand break by HRR. 
DNA polymerase (blue kidney bean structure) recognizes the free 3’-hydroxyl group and uses the donor DNA template to replace the deleted nucleotides.



**Figure 14-15: Repair of a double-strand break by HRR.** The thick black arrow identifies where a double-strand (ds) break has occurred. First proteins (blue spheres) recognize the break in the backbone of the DNA, these help to recruit other proteins known as exonucleases (brown Pac Man) to remove some nucleotides from the 5’ ends. This is known as resection and allows for the formation of a 3’-single-stranded overhang. Other proteins (yellow triangles) recognize and bind the overhang sequence. These proteins “guide” the overhang to a region of homologous DNA sequence on a different DNA strand, usually a sister chromatid (after S phase) or a homologous chromosome. The overhang “invades” the donor DNA helix and displaces a strand of DNA, forming a structure called a d-loop. DNA polymerase (blue kidney bean structure) recognizes the free 3’-hydroxyl group and uses the donor DNA template to replace the deleted nucleotides. At some point, the repaired DNA strand becomes free of the donor DNA and hydrogen bonds back to its original DNA sequence. Polymerase finishes replicating any areas it can and then ligase comes in to connect the DNA fragments together. Please note that this figure is illustrating one way a ds break is repaired using homologous sequences; this version is called synthesis-dependent strand annealing. There are other ways that are more complex the cell can use to repair ds breaks. (Illustration by Jennell Talley)