DNA proofreading and repair

Mechanisms to correct errors during DNA replication and to repair DNA damage over the cell's lifetime.

Key points:

- Cells have a variety of mechanisms to prevent **mutations**, or permanent changes in DNA sequence.
- During DNA synthesis, most DNA polymerases "check their work," fixing the majority of mispaired bases in a process called **proofreading**.
- Immediately after DNA synthesis, any remaining mispaired bases can be detected and replaced in a process called **mismatch repair**.
- If DNA gets damaged, it can be repaired by various mechanisms, including **chemical reversal**, **excision repair**, and **double-stranded break repair**.

Introduction

What does DNA have to do with cancer? Cancer occurs when cells divide in an uncontrolled way, ignoring normal "stop" signals and producing a tumor. This bad behavior is caused by accumulated **mutations**, or permanent sequence changes in the cells' DNA.

Replication errors and DNA damage are actually happening in the cells of our bodies all the time. In most cases, however, they don’t cause cancer, or even mutations. That’s because they are usually detected and fixed by DNA proofreading and repair mechanisms. Or, if the damage cannot be fixed, the cell will undergo programmed cell death (apoptosis) to avoid passing on the faulty DNA.

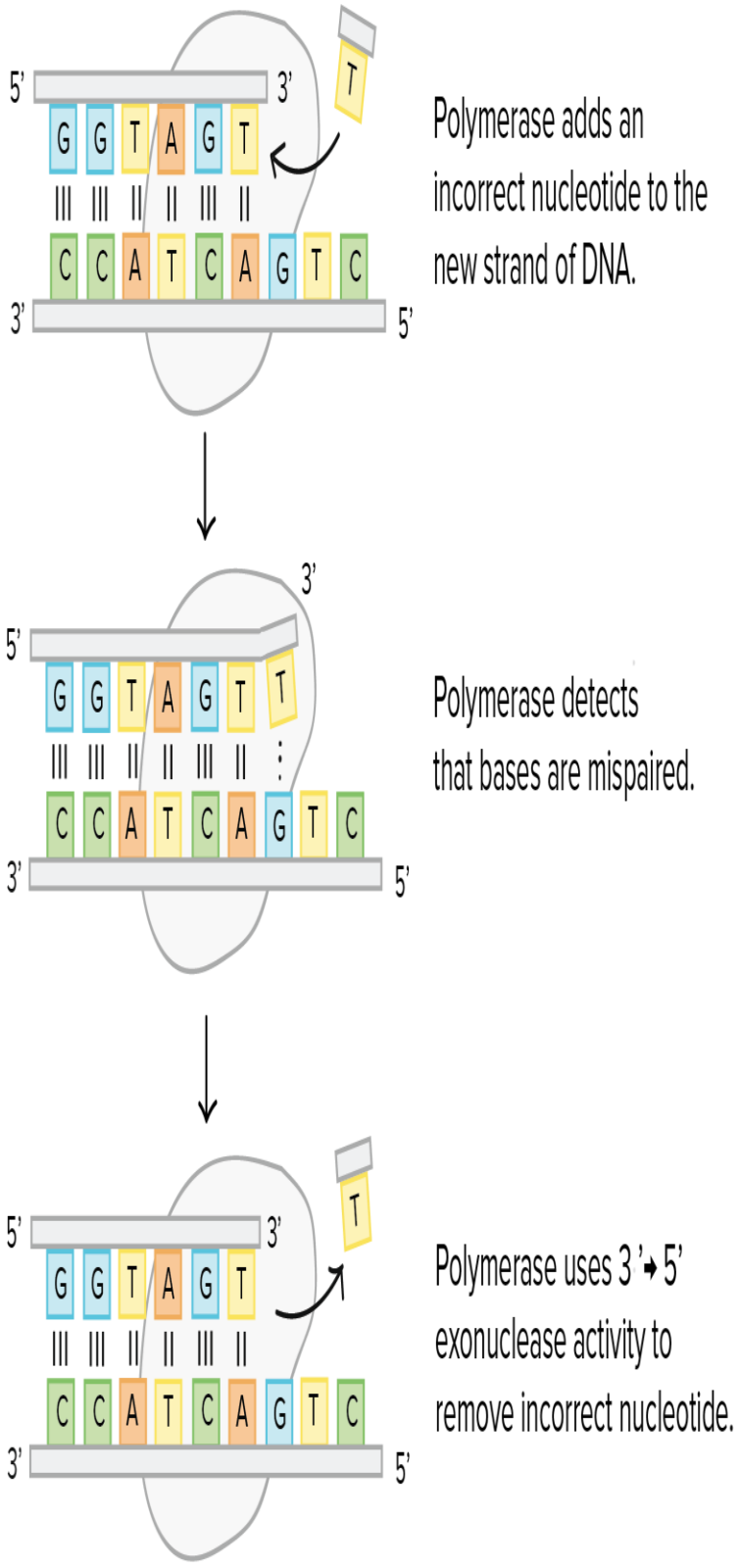
Mutations happen, and get passed on to daughter cells, only when these mechanisms fail. Cancer, in turn, develops only when multiple mutations in division-related genes accumulate in the same cell.

In this article, we'll take a closer look at the mechanisms used by cells to correct replication errors and fix DNA damage, including:

- Proofreading, which corrects errors during DNA replication
- Mismatch repair, which fixes mispaired bases right after DNA replication
- DNA damage repair pathways, which detect and correct damage throughout the cell cycle

Proofreading

DNA polymerases are the enzymes that build DNA in cells. During DNA replication (copying), most DNA polymerases can "check their work" with each base that they add. This process is called **proofreading**. If the polymerase detects that a wrong (incorrectly paired) nucleotide has been added, it will remove and replace the nucleotide right away, before continuing with DNA synthesis^1ıstart superscript, 1, end superscript.



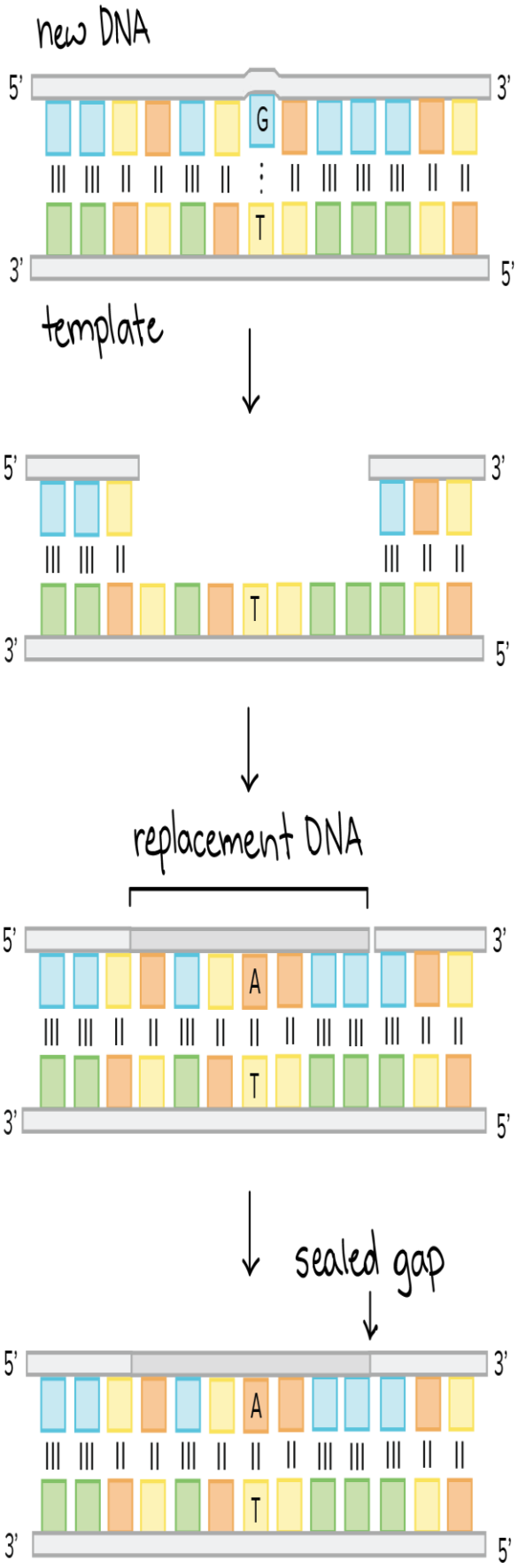
Proofreading:

1. DNA polymerase adds a new base to the 3' end of the growing, new strand. (The template has a G, and the polymerase incorrectly adds a T rather than a C to the new strand.)
2. Polymerase detects that the bases are mispaired.
3. Polymerase uses 3' to 5' exonuclease activity to remove the incorrect T from the 3' end of the new strand.

Mismatch repair

Many errors are corrected by proofreading, but a few slip through. **Mismatch repair** happens right after new DNA has been made, and its job is to remove and replace mis-paired bases (ones that were not fixed during proofreading). Mismatch repair can also detect and correct small insertions and deletions that happen when the polymerases "slips," losing its footing on the template^22squared.

How does mismatch repair work? First, a protein complex (group of proteins) recognizes and binds to the mispaired base. A second complex cuts the DNA near the mismatch, and more enzymes chop out the incorrect nucleotide and a surrounding patch of DNA. A DNA polymerase then replaces the missing section with correct nucleotides, and an enzyme called a DNA ligase seals the gap^22squared.



A mismatch is detected in newly synthesized DNA.

The new DNA strand is cut, and the mispaired nucleotide and its neighbors are removed.

The missing patch is replaced with correct nucleotides by a DNA polymerase.

A DNA ligase seals the gap in the DNA backbone.

Mismatch repair.

1. A mismatch is detected in newly synthesized DNA. There is a G in the new strand paired with a T in the template (old) strand.
2. The new DNA strand is cut, and a patch of DNA containing the mispaired nucleotide and its neighbors is removed.
3. The missing patch is replaced with correct nucleotides by a DNA polymerase.
4. A DNA ligase seals the remaining gap in the DNA backbone.

One thing you may wonder is how the proteins involved in DNA repair can tell "who's right" during mismatch repair. That is, when two bases are mispaired (like the G and T in the drawing above), which of the two should be removed and replaced?

In bacteria, original and newly made strands of DNA can be told apart by a feature called *methylation state*. An old DNA strand will have methyl (-\text {CH}_3−CH3minus, start text, C, H, end text, start subscript, 3, end subscript) groups attached to some of its bases, while a newly made DNA strand will not yet have gotten its methyl group^3cubed.

In eukaryotes, the processes that allow the original strand to be identified in mismatch repair involve recognition of nicks (single-stranded breaks) that are found only in the newly synthesized DNA^3cubed.

DNA damage repair mechanisms

Bad things can happen to DNA at almost any point in a cell's lifetime, not just during replication. In fact, your DNA is getting damaged all the time by outside factors like UV light, chemicals, and X-rays—not to mention spontaneous chemical reactions that happen even without environmental insults!^4start superscript, 4, end superscript

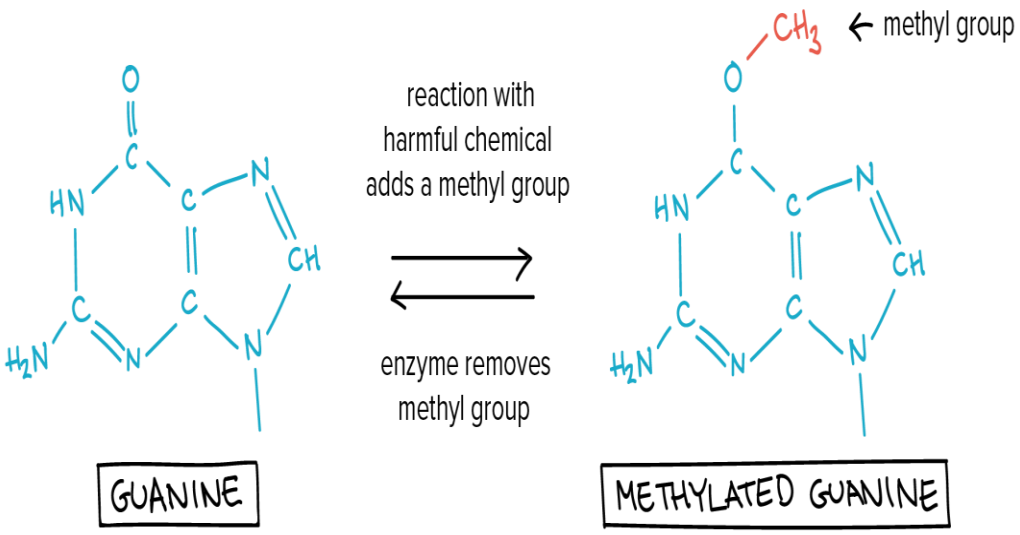
Fortunately, your cells have repair mechanisms to detect and correct many types of DNA damage. Repair processes that help fix damaged DNA include:

- **Direct reversal:** Some DNA-damaging chemical reactions can be directly "undone" by enzymes in the cell.
- **Excision repair:** Damage to one or a few bases of DNA is often fixed by removal (excision) and replacement of the damaged region. In **base excision repair,** just the damaged base is removed. In **nucleotide excision repair**, as in the mismatch repair we saw above, a patch of nucleotides is removed.
- **Double-stranded break repair:** Two major pathways, non-homologous end joining and homologous recombination, are used to repair double-stranded breaks in DNA (that is, when an entire chromosome splits into two pieces).

Reversal of damage

In some cases, a cell can fix DNA damage simply by reversing the chemical reaction that caused it. To understand this, we need to realize that "DNA damage" often just involves an extra group of atoms getting attached to DNA through a chemical reaction.

For example, guanine (G) can undergo a reaction that attaches a methyl (-\text{CH}_3−CH3minus, start text, C, H, end text, start subscript, 3, end subscript) group to an oxygen atom in the base. The methyl-bearing guanine, if not fixed, will pair with thymine (T) rather than cytosine (C) during DNA replication. Luckily, humans and many other organisms have an enzyme that can remove the methyl group, reversing the reaction and returning the base to normal^5start superscript, 5, end superscript.



Methylation of guanine

A normal guanine base undergoes a reaction with a harmful chemical, causing a methyl (-\text{CH}_3−CH3minus, start text, C, H, end text, start subscript, 3, end subscript) group to be added to the carbonyl O found on one of the rings of the base.

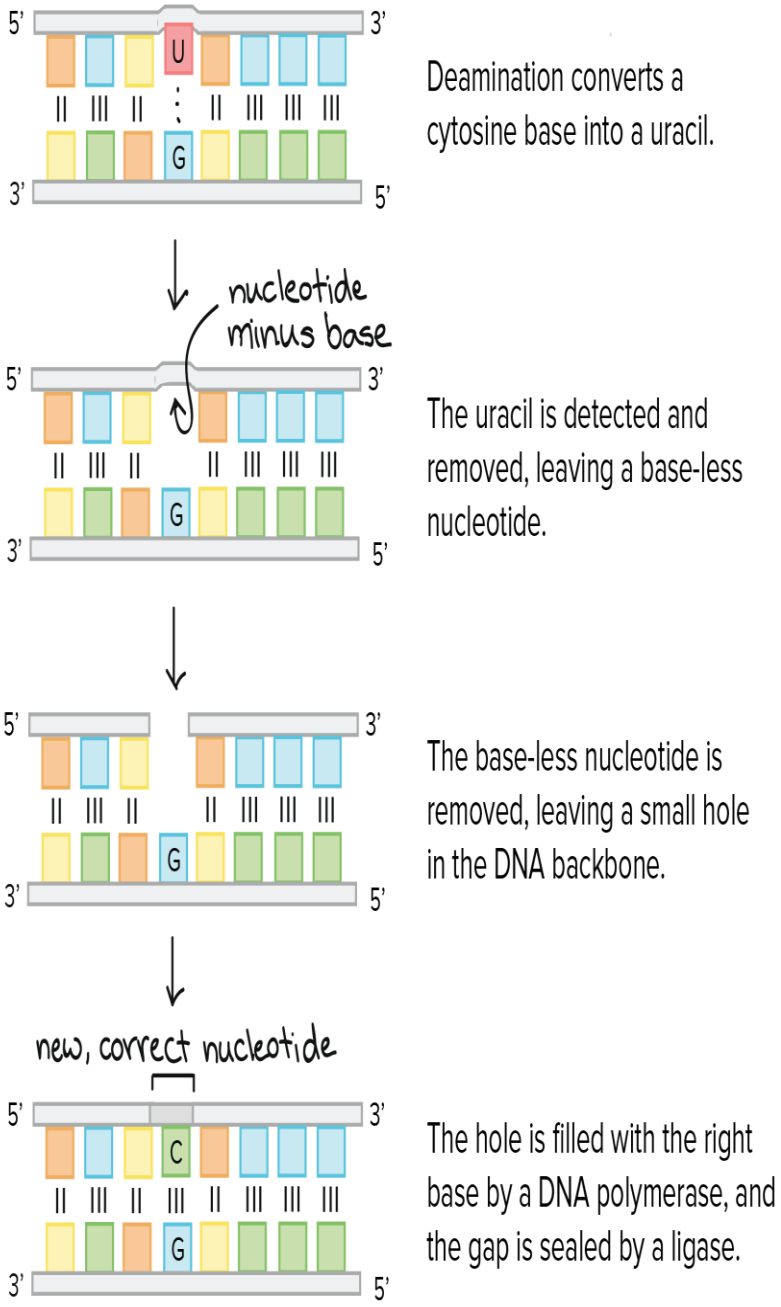
The methyl group can be removed from the damaged, methylated base by an enzyme found in the cell.

Base excision repair

Base excision repair is a mechanism used to detect and remove certain types of damaged bases. A group of enzymes called glycosylases play a key role in base excision repair. Each glycosylase detects and removes a specific kind of damaged base.

For example, a chemical reaction called deamination can convert a cytosine base into uracil, a base typically found only in RNA. During DNA replication, uracil will pair with adenine rather than guanine (as it would if the base was still cytosine), so an uncorrected cytosine-to-uracil change can lead to a mutation^5[5]start superscript, 5, end superscript.

To prevent such mutations, a glycosylase from the base excision repair pathway detects and removes deaminated cytosines. Once the base has been removed, the "empty" piece of DNA backbone is also removed, and the gap is filled and sealed by other enzymes^6start superscript, 6, end superscript.



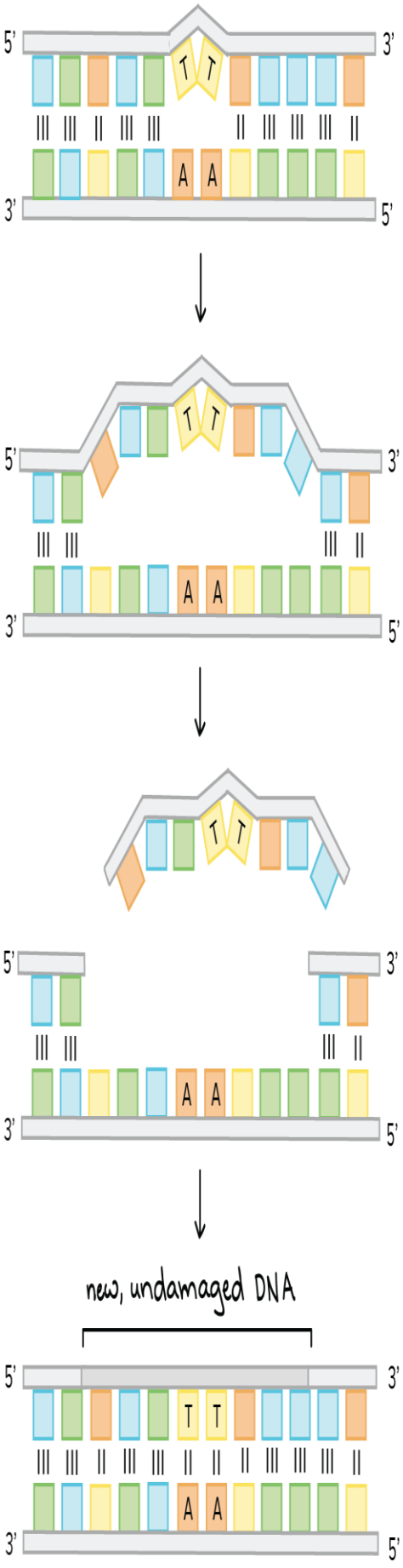
Base excision repair of a deaminated cytosine.

1. Deamination converts a cytosine base into a uracil. This results in a double helix in which a G in one strand is paired with a U in the other. The U was formerly a C, but was converted to U via deamination.
2. The uracil is detected and removed, leaving a base-less nucleotide.
3. The base-less nucleotide is removed, leaving a 1-nucleotide hole in the DNA backbone.
4. The hole is filled with the right base by a DNA polymerase, and the gap is sealed by a ligase.

Nucleotide excision repair

Nucleotide excision repair is another pathway used to remove and replace damaged bases. Nucleotide excision repair detects and corrects types of damage that distort the DNA double helix. For instance, this pathway detects bases that have been modified with bulky chemical groups, like the ones that get attached to your DNA when it's exposed to chemicals in cigarette smoke^77start superscript, 7, end superscript.

Nucleotide excision repair is also used to fix some types of damage caused by UV radiation, for instance, when you get a sunburn. UV radiation can make cytosine and thymine bases react with neighboring bases that are also Cs or Ts, forming bonds that distort the double helix and cause errors in DNA replication. The most common type of linkage, a **thymine dimer**, consists of two thymine bases that react with each other and become chemically linked^88start superscript, 8, end superscript.



UV radiation produces a thymine dimer.

Once the dimer has been detected, the surrounding DNA is opened to form a bubble.

Enzymes cut the damaged region out of the bubble.

A DNA polymerase replaces the excised (cut-out) DNA, and a ligase seals the backbone.

Nucleotide excision repair of a thymine dimer.

1. UV radiation produces a thymine dimer. In a thymine dimer, two Ts that are next to each other in the same strand link up via a chemical reaction between the bases. This creates a distortion in the shape of the double helix.
2. Once the dimer has been detected, the surrounding DNA is opened to form a bubble.
3. Enzymes cut the damaged region (thymine dimer plus neighboring regions of same strand) out of the bubble.
4. A DNA polymerase replaces the excised (cut-out) DNA, and a ligase seals the backbone.

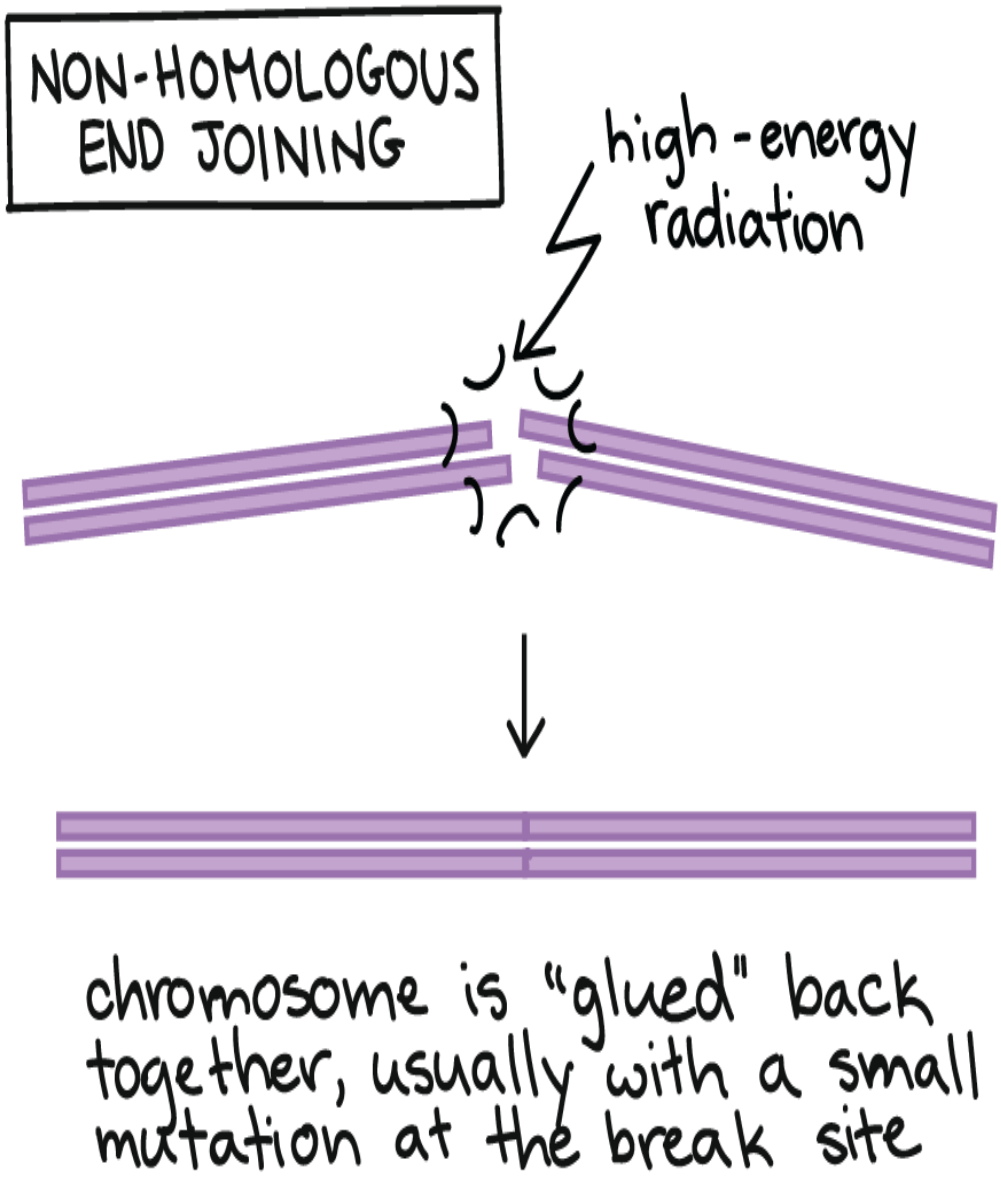
   In nucleotide excision repair, the damaged nucleotide(s) are removed along with a surrounding patch of DNA. In this process, a helicase (DNA-opening enzyme) cranks open the DNA to form a bubble, and DNA-cutting enzymes chop out the damaged part of the bubble. A DNA polymerase replaces the missing DNA, and a DNA ligase seals the gap in the backbone of the strand^9start superscript, 9, end superscript.

Double-stranded break repair

Some types of environmental factors, such as high-energy radiation, can cause double-stranded breaks in DNA (splitting a chromosome in two). This is the kind of DNA damage linked with superhero origin stories in comic books, and with disasters like Chernobyl in real life.

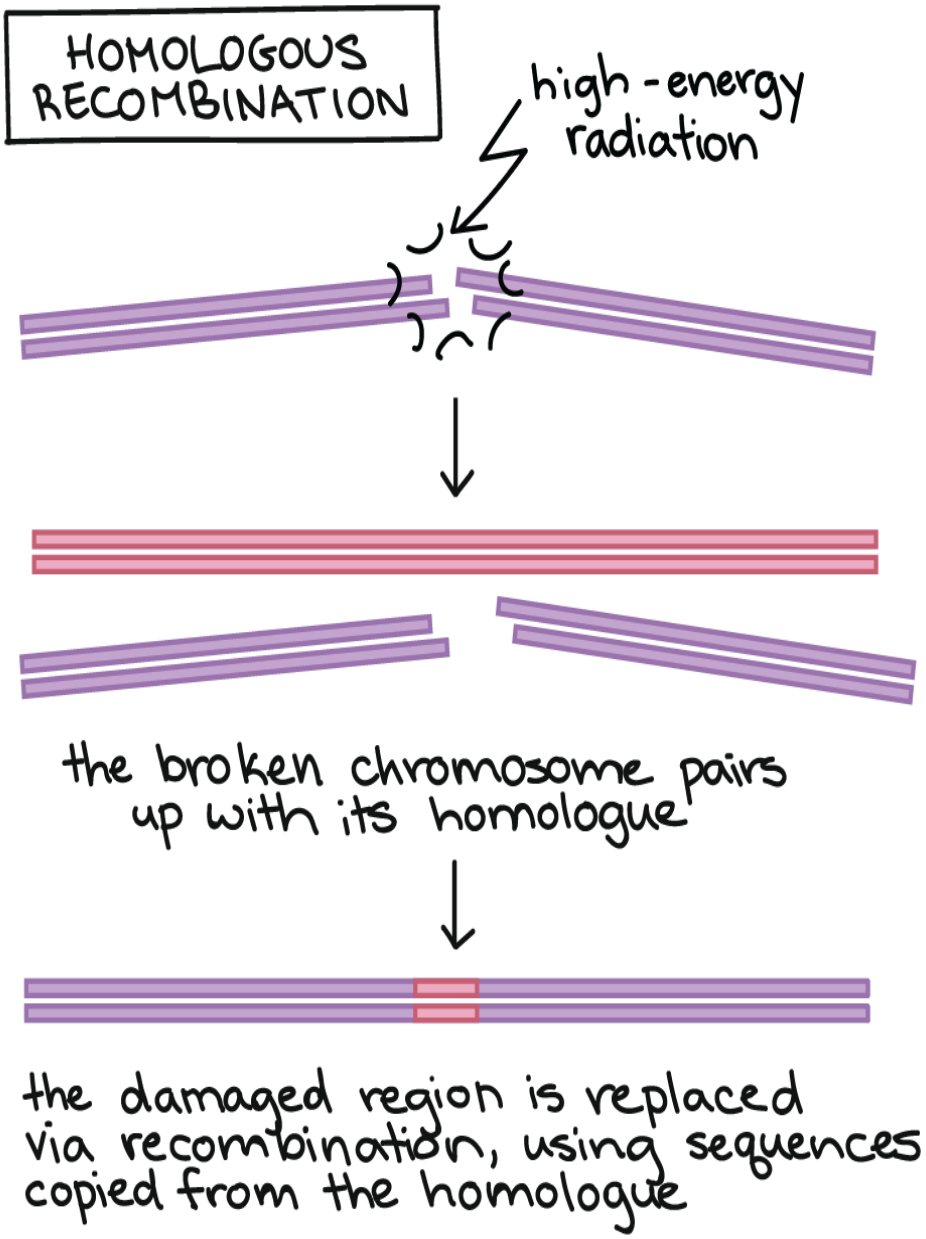
Double-stranded breaks are dangerous because large segments of chromosomes, and the hundreds of genes they contain, may be lost if the break is not repaired. Two pathways involved in the repair of double-stranded DNA breaks are the non-homologous end joining and homologous recombination pathways.

In **non-homologous end joining**, the two broken ends of the chromosome are simply glued back together. This repair mechanism is “messy” and typically involves the loss, or sometimes addition, of a few nucleotides at the cut site. So, non-homologous end joining tends to produce a mutation, but this is better than the alternative (loss of an entire chromosome arm)^{10}10start superscript, 10, end superscript.



A double-stranded break may be repaired by non-homologous end joining. The chromosome is "glued" back together, usually with a small mutation at the break site.

In **homologous recombination**, information from the homologous chromosome that matches the damaged one (or from a sister chromatid, if the DNA has been copied) is used to repair the break. In this process, the two homologous chromosomes come together, and the undamaged region of the homologue or chromatid is used as a template to replace the damaged region of the broken chromosome. Homologous recombination is "cleaner" than non-homologous end joining and does not usually cause mutations^{11}11start superscript, 11, end superscript.



The double-stranded break may be repaired by homologous recombination. The broken chromosome pairs up with its homologue. The damaged region is replaced via recombination, using sequences copied from the homologue.

DNA proofreading and repair in human disease

Evidence for the importance of proofreading and repair mechanisms comes from human genetic disorders. In many cases, mutations in genes that encode proofreading and repair proteins are associated with heredity cancers (cancers that run in families). For example:

- **Hereditary nonpolyposis colorectal cancer** (also called **Lynch syndrome**) is caused by mutations in genes encoding certain mismatch repair proteins^{12,13}12,13start superscript, 12, comma, 13, end superscript. Since mismatched bases are not repaired in the cells of people with this syndrome, mutations accumulate much more rapidly than in the cells of an unaffected person. This can lead to the development of tumors in the colon.
- People with **xeroderma pigmentosum** are extremely sensitive to UV light. This condition is caused by mutations affecting the nucleotide excision repair pathway. When this pathway doesn't work, thymine dimers and other forms of UV damage can't be repaired. People with xeroderma pigmentosum develop severe sunburns from just a few minutes in the sun, and about half will get skin cancer by the age of 101010 unless they avoid the sun.

DNA: Damage Types and Repair Mechanisms

Introduction to DNA Damage and Repair:

DNA is a highly stable and versatile molecule. Though sometimes the damage is caused to it, it is able to maintain the integrity of information contained in it. The perpetuation of genetic material from generation to generation depends upon keeping the rates of mutation at low level. DNA has many elaborate mechanisms to repair any damage or distortion.

The most frequent sources of damage to DNA are the inaccuracy in DNA replication and chemical changes in DNA. Malfunction of the process of replication can lead to incorporation of wrong bases, which are mismatched with the complementary strand. The damage causing chemicals break the backbone of the strand and chemically alter the bases. Alkylation, oxidation and methylation cause damage to bases. X-rays and gamma radiations cause single or double stranded breaks in DNA.

A change in the sequence of bases if replicated and passed on to the next generation becomes permanent and leads to mutation. At the same time mutations are also necessary which provide raw material for evolution. Without evolution, the new species, even human beings would not have arisen. Therefore a balance between mutation and repair is necessary.

Types of Damage:

Damage to DNA includes any deviation from the usual double helix structure.

1. Simple Mutations:

Simplest mutations are switching of one base for another base. In transition one pyrimidine is substituted by another pyramidine and purine with another purine. Trans-version involves substitution of a

pyramidine by a purine and purine by a pyrimidine such as T by G or A and A by C or T. Other simple mutations are detection, insertion of a single nucleotide or a small number of nucleotides. Mutations which change a single nucleotide are called point mutations.

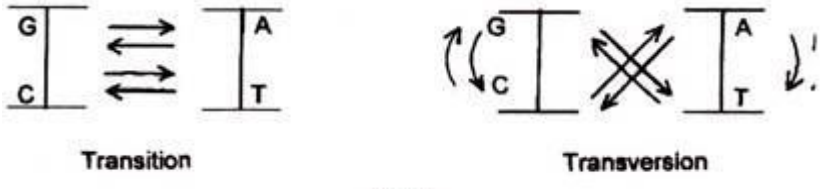


Fig. 5.1.

2. Deamination:

The common alteration of form or damage includes deamination of cytosine (C) to form uracil (u) which base pairs with adenine (A) in next replication instead of guanine (G) with which the original cytosine would have paired.

As uracil is not present in DNA, adenine base pairs with thymine (T). Therefore C-G pair is replaced by T-A in next replication cycle. Similarly, hypoxanthine results from adenine deamination.

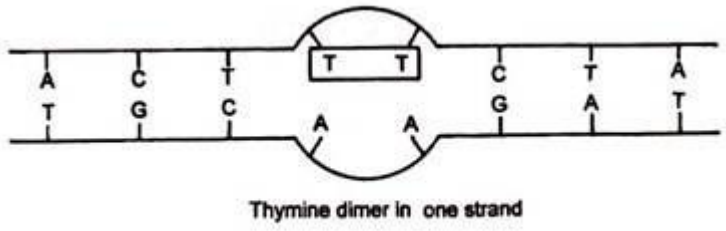
3. Missing Bases:

Cleavage of N-glycosidic bond between purine and sugar causes loss of purine base from DNA. This is called depurination. This apurinic site becomes non-coding lesion.

4. Chemical Modification of Bases:

Chemical modification of any of the four bases of DNA leads to modified bases. Methyl groups are added to various bases. Guanine forms 7- methylguanine, 3-methylguanine. Adenine forms 3-methyladenine. Cytosine forms 5- Methylcytosine.

Replacement of amino group by a keto group converts 5-methylcytosine to thymine.



Thymine dimer in one strand

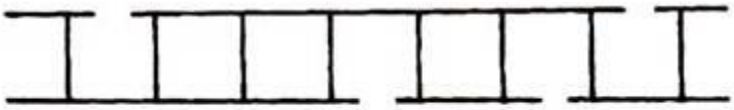
Fig. 5.2.

5. Formation of Pyrimidine Dimers (Thymine Dimers):

Formation of thymine dimers is very common in which a covalent bond (cyclobutyl ring) is formed between adjacent thymine bases. This leads to loss of base pairing with opposite stand. A bacteria may have thousands of dimers immediately after exposure to ultraviolet radiations.

6. Strand Breaks:

Sometimes phosphodiester bonds break in one strand of DNA helix. This is caused by various chemicals like peroxides, radiations and by enzymes like DNases. This leads to breaks in DNA backbone. Single strand breaks are more common than double strand breaks.



Double strand breaks

Fig. 5.3.

Sometimes X-rays, electronic beams and other radiations may cause phosphodiester bonds breaks in both strands which may not be directly opposite to each other. This leads to double strand breaks.

Some sites on DNA are more susceptible to damage. These are called hot-stops.

Repair Mechanisms:

Most kinds of damage create impediments to replication or transcription. Altered bases cause mispairing and can cause permanent alteration to DNA sequence after replication.

In order to maintain the integrity of information contained in it, the DNA has various repair mechanisms.

1. Direct Repair:
The damage is reversed by a repair enzyme which is called photoreactivation. This mechanism involves a light dependant enzyme called DNA photolyase. The enzyme is present in almost all cells from bacteria to animals. It uses energy from the absorbed light to cleave the C-C bond of cyclobutyl ring of the thymine dimers. In this way thymine dimers are monomerized.

2. Excision Repair:
It includes base excision repair and nucleotide excision repair. Base excision repair system involves an enzyme called N-glycosylase which recognizes the abnormal base and hydrolyes glycosidic bond between it and sugar.

Another enzyme, an endonuclease cleaves the DNA backbone on the 5′-side of the abnormal base. Then the DNA polymerase by its exonuclease activity removes the abnormal base. DNA polymerase then replaces it with normal base and DNA ligase seals the region.

Nucleotide repair system includes three steps, incision, excision and synthesis. Incision is done by endonuclease enzyme precisely on either side of the damaged patch of the strand. In this way damaged portion of the strand is cleaved.

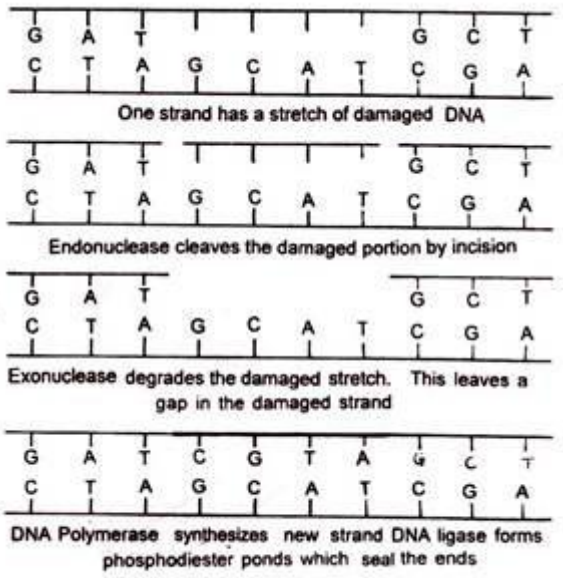


Fig. 5.4.

Endonuclease enzymes involved are UvrA, UvrB which recognize the damaged stretch of the strand. UvrC makes two cuts (incision)

on either side. Exonuclease removes the damaged strand. Enzyme involved is UvrD.

Later, DNA polymerase synthesizes the new strand by using complementary strand as a template. DNA ligase forms phosphodiester bonds which seal the ends on newly synthesized strand.

3. Mismatch Base Repair:
Sometimes wrong bases are incorporated during replication process, G-T or C-A pairs are formed. The wrong base is always incorporated in the daughter strand only. Therefore in order to distinguish the two strands for the purpose of repair, the adenine bases of the template strand are labelled or tagged by methyl groups. In this way the newly replication DNA helix is hemimethylated. The excision of wrong bases occur in the non-methylated or daughter strand.

4. Recombination Repair or Retrieval System:
In thymine dimer or other type of damage, DNA replication cannot proceed properly. A gap opposite to thymine dimer is left in the newly synthesized daughter strand. The gap is repaired by recombination mechanism or retrieval mechanism called also sister strand exchange.

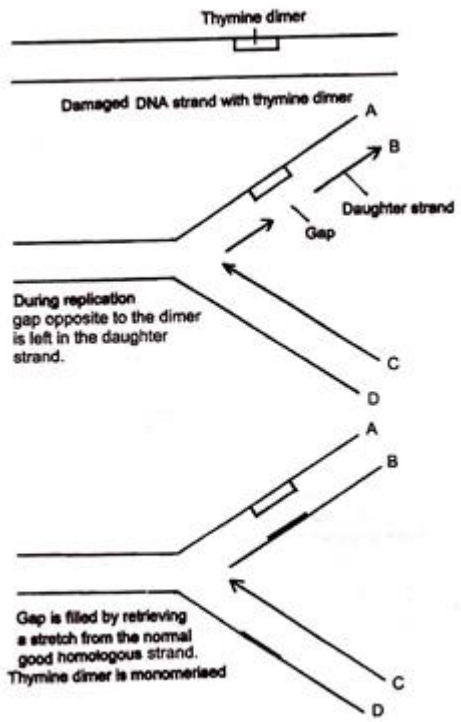


Fig. 5.5

During replication of DNA two identical copies are produced. Replicating DNA molecule has four strands A, B, C and D. Strands A and C have same DNA , sequence. Strands B and D also have same sequence as they are identical. A thymine dimer is present in strand A. The replication fork passes the dimmer as it cannot form hydrogen bonds with incoming adenine bases, thus creating a gap in the newly synthesized strand B.

In recombination repair system a short identical segment of DNA is retrieved from strand D and is inserted into the gap of strand B. But this creates a gap in strand D which is easily filled up by DNA polymerase using normal strand C as a template. This event is dependent on the activity of a special protein Rec A. The Rec A protein plays its role in retrieving a portion of the complementary strand from other side of the replication fork to fill the gap. Rec A is a strand exchange protein.

After filling both gaps, thymine is monomerised. So in this repair mechanism a portion of DNA strand is retrieved from the normal homologous DNA segment. This is also known as daughter strand gap repair mechanism.

5. SOS Repair Mechanism:

Sometimes the replicating machinery is unable to repair the damaged portion and bypasses the damaged site. This is known as translesion synthesis also called bypass system and is emergency repair system. This mechanism is catalyzed by a special class of DNA polymerases called Y-family of DNA polymerases which synthesized DNA directly across the damaged portion.

Websites

https://www.khanacademy.org/science/high-school-biology/hs-molecular-genetics/hs-discovery-and-structure-of-dna/a/dna-proofreading-and-repair

http://www.biologydiscussion.com/dna/dna-damage-types-and-repair-mechanisms-with-diagram/16332