Cellular and Molecular Biology I

BIO-205-4

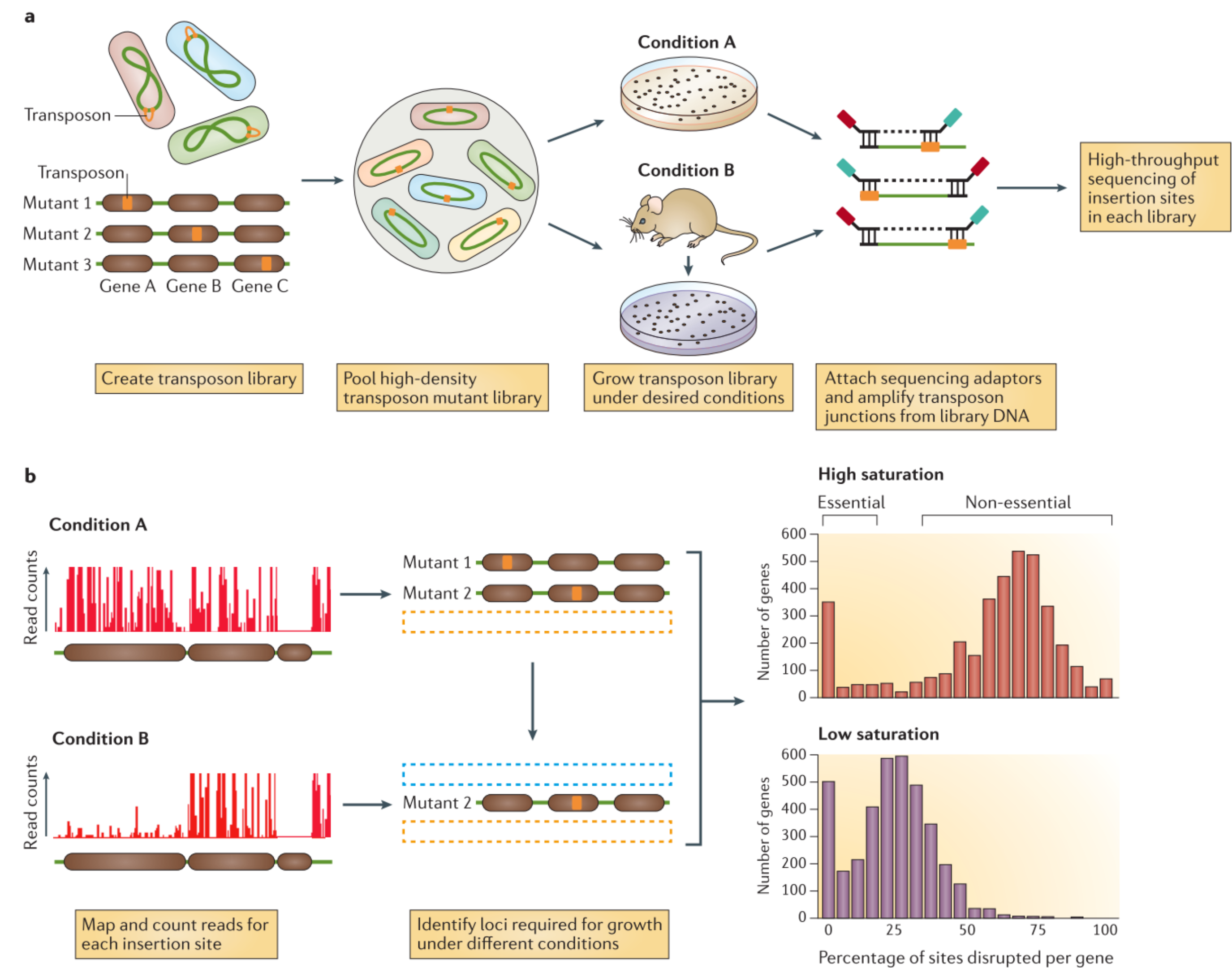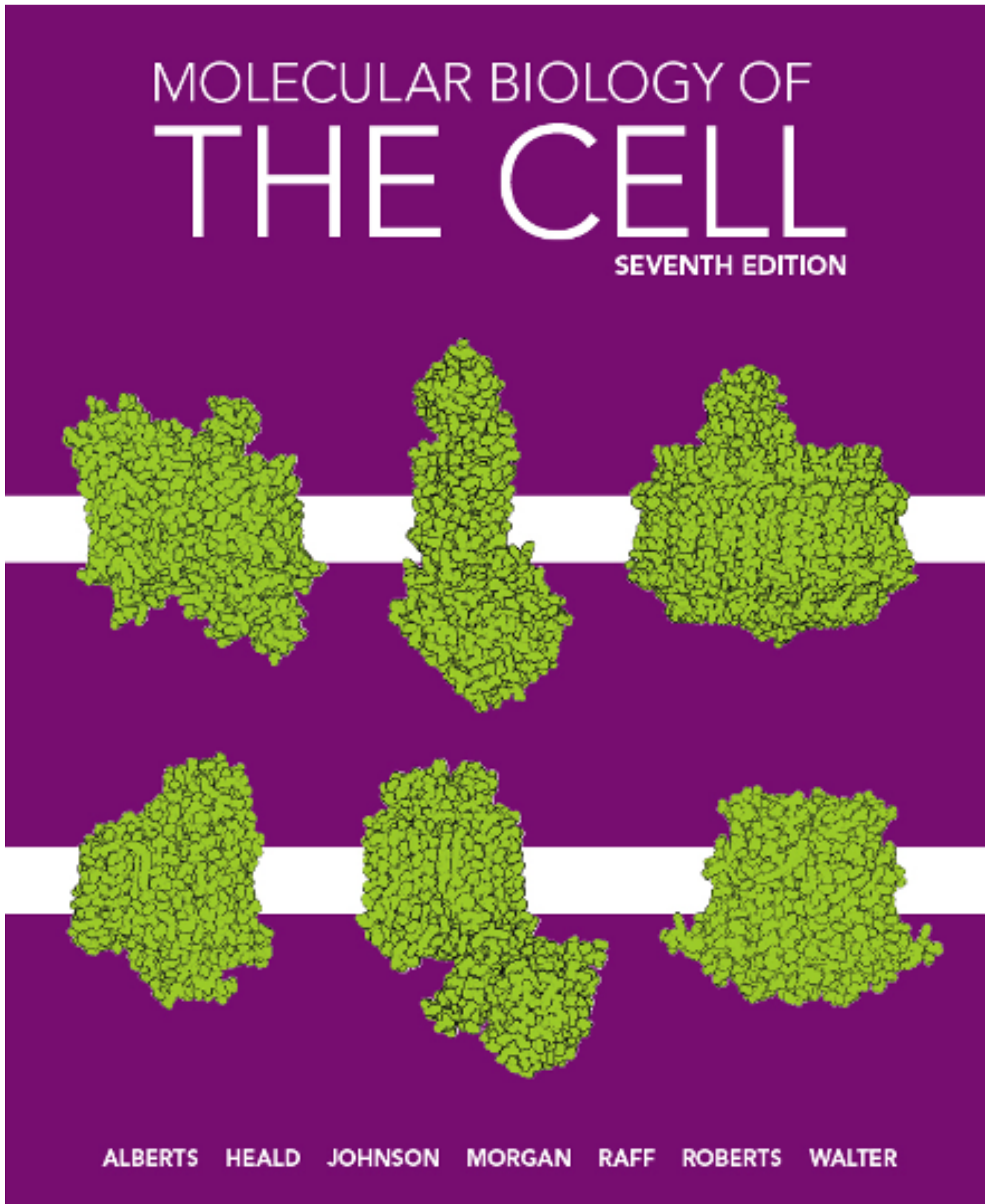**Camille Goemans - 2024**

Using transposons in research

Figure 1: Transposon insertion sequencing method (from Chao *et al.* 2016)

Chapter 5

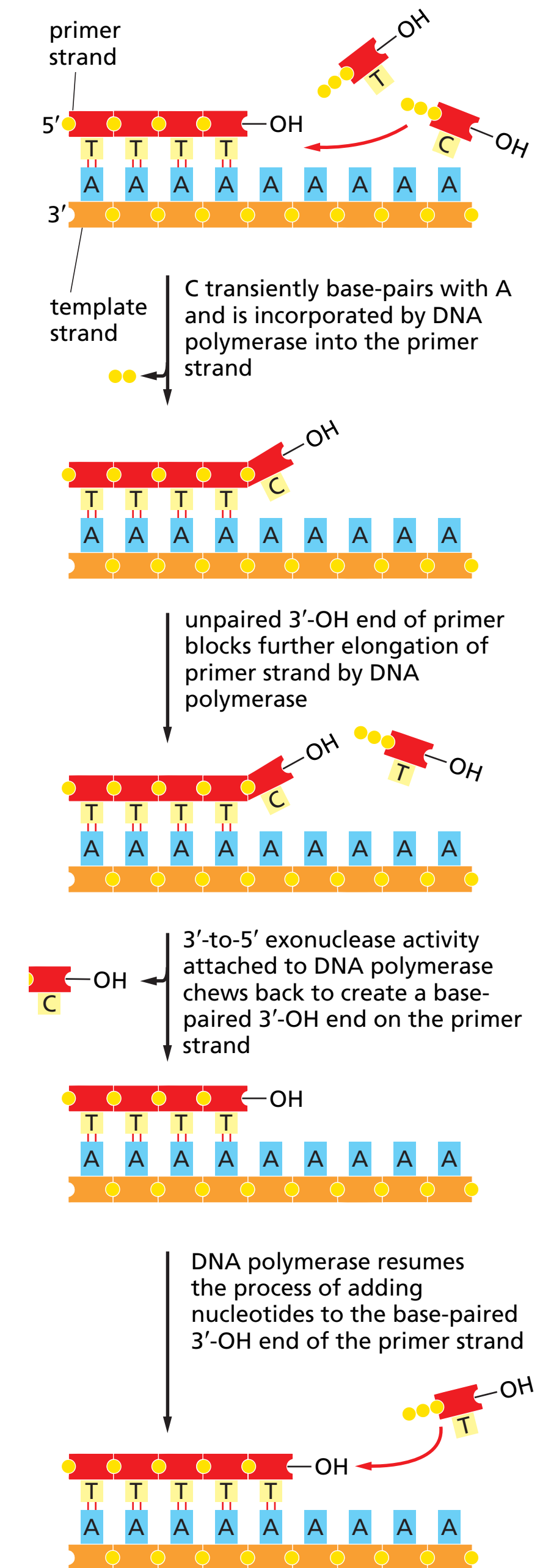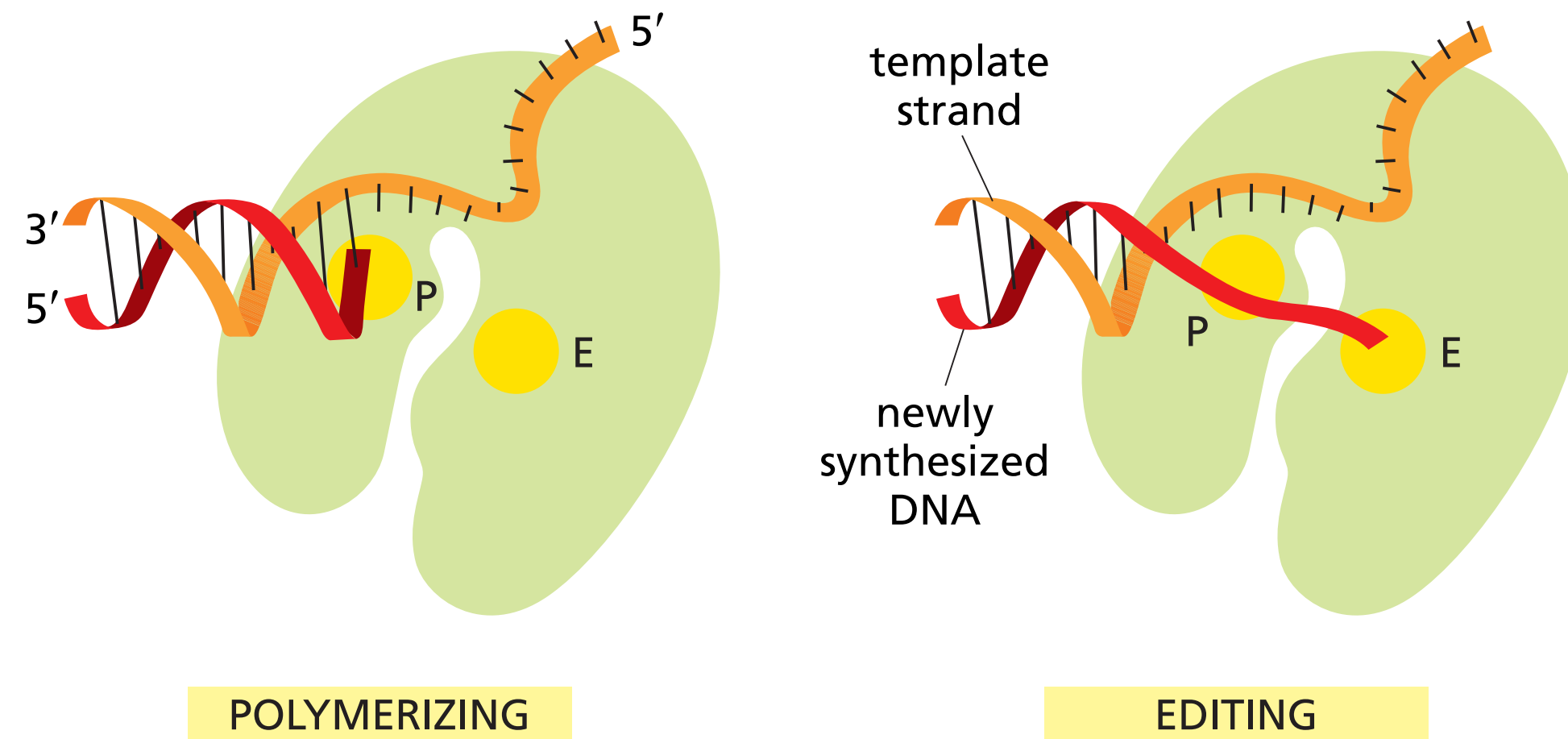## DNA Replication, Repair, and Recombination

Plan

- Quick recap

- <u>DNA replication mechanisms</u>

- The initiation and completion of DNA replication in chromosomes

- DNA repair

- Transposition and conservative site-specific recombination

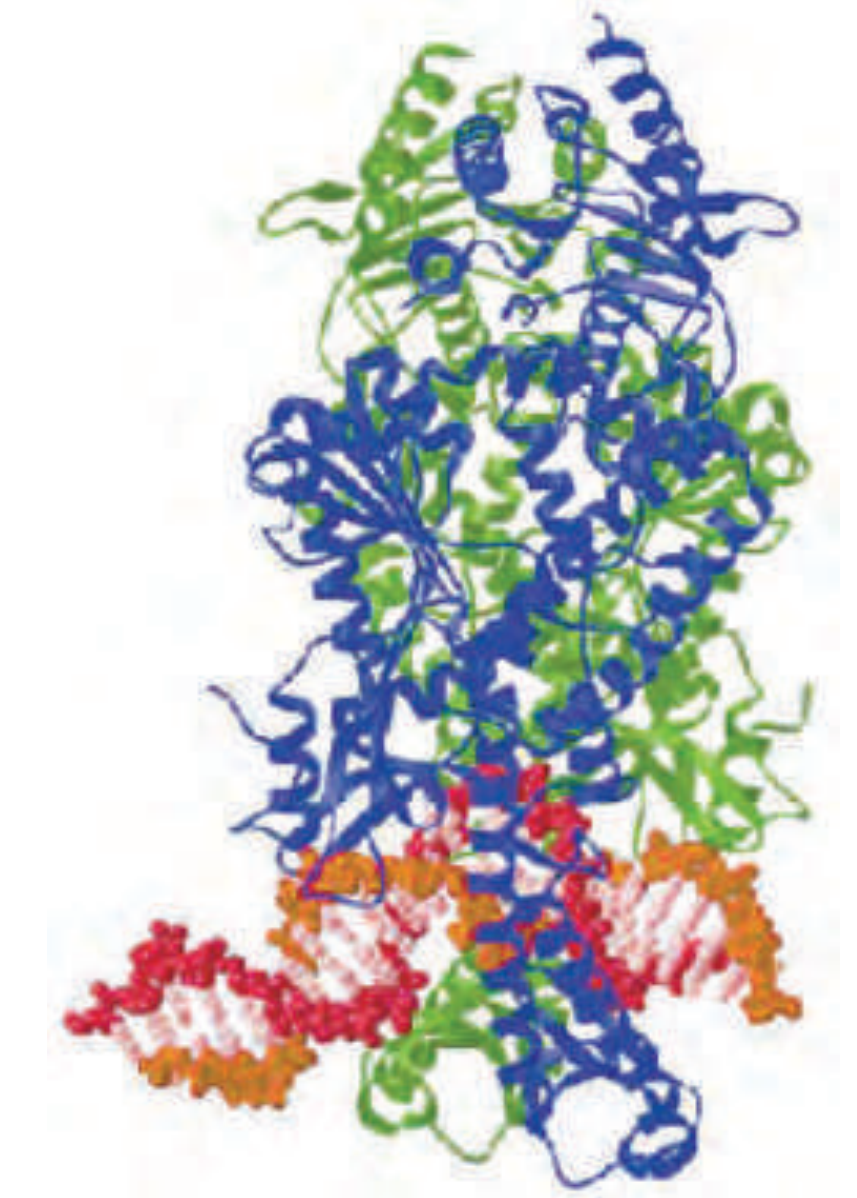# How does the DNA polymerase avoid errors?

Proofreading mechanisms

- Step 1 - **Proofreading by DNA polymerase**, incorrect nucleotides are harder to add to the chain (energetically less favorable) and are more likely to diffuse away

- Step 2- **Exonucleic proofreading by DNA polymerase** takes place when an incorrect nucleotide is inserted (only works when DNA is synthesised from 5' to 3')

POLYMERIZING

EDITING

C transiently base-pairs with A and is incorporated by DNA polymerase into the primer strand

unpaired 3'-OH end of primer blocks further elongation of primer strand by DNA polymerase

3'-to-5' exonuclease activity attached to DNA polymerase chews back to create a base-paired 3'-OH end on the primer strand

DNA polymerase resumes the process of adding nucleotides to the base-paired 3'-OH end of the primer strand

Proofreading mechanisms

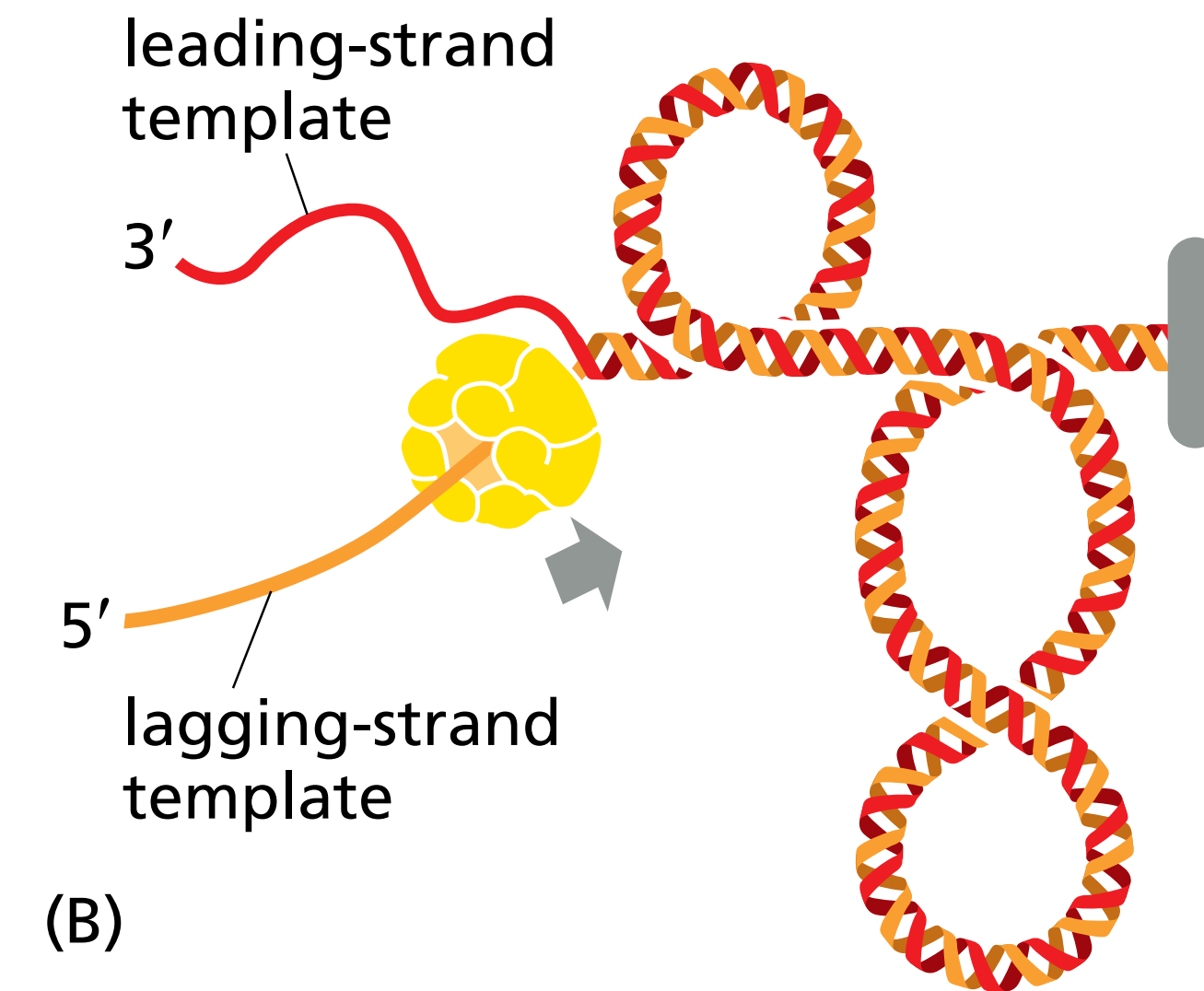- Step 3- **Strand-directed mismatch repair** detects distortion into mismatched DNA

- It needs to recognise the **new strand**

- In *E. coli*, this depends on the **methylation status** of all adenines on the DNA (which takes some time)

- Similar system in Eukaryotes, that recognises **knicks in the lagging strand** before the action of the DNA ligase - not clear how it works on the leading strand

error in newly made strand

BINDING OF MISMATCH PROOFREADING PROTEINS

MutS   MutL

DNA SCANNING DETECTS NICK IN NEW DNA STRAND

STRAND REMOVAL

REPAIR DNA SYNTHESIS

(A)

(B)

Preventing DNA tangling

- For replication to occur, the double helix needs to **unwind** ahead of the replication fork (which is energetically unfavourable)
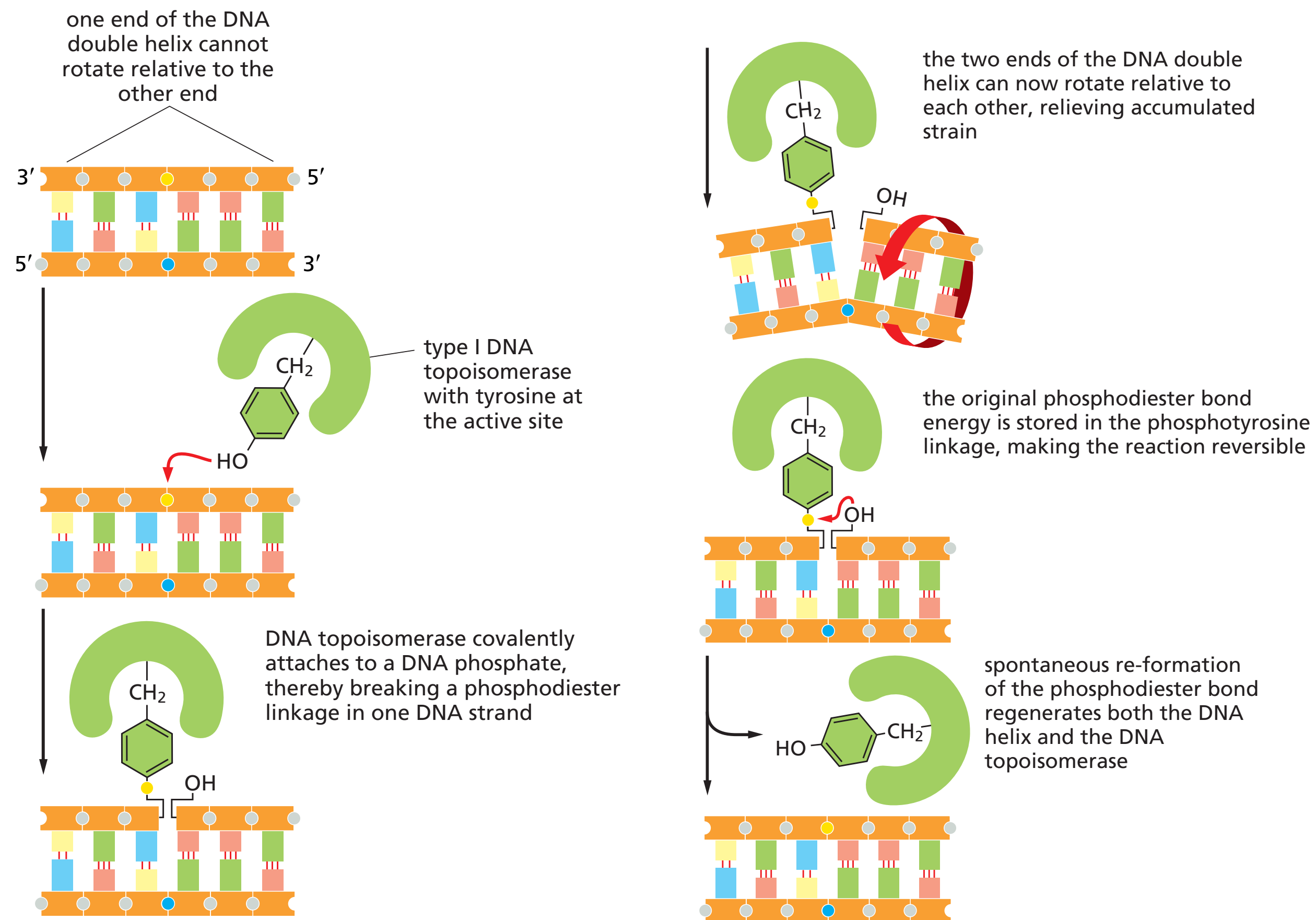

3′

5′

if the DNA cannot rapidly rotate, torsional stress will build up

(A)

leading-strand template

3′

5′

lagging-strand template

(B)

Preventing DNA tangling

- The overwinding is relieved by proteins known as **DNA topoisomerases**

- They have different **mechanisms of action** (one example here)

one end of the DNA double helix cannot rotate relative to the other end

type I DNA topoisomerase with tyrosine at the active site

DNA topoisomerase covalently attaches to a DNA phosphate, thereby breaking a phosphodiester linkage in one DNA strand

the two ends of the DNA double helix can now rotate relative to each other, relieving accumulated strain

the original phosphodiester bond energy is stored in the phosphotyrosine linkage, making the reaction reversible

spontaneous re-formation of the phosphodiester bond regenerates both the DNA helix and the DNA topoisomerase

Eukaryotes vs. Prokaryotes

- Most of what we know derives from **studies on bacteria and bacteriophages** (viruses that infect bacteria)

- Most mechanisms have been **conserved through evolution**

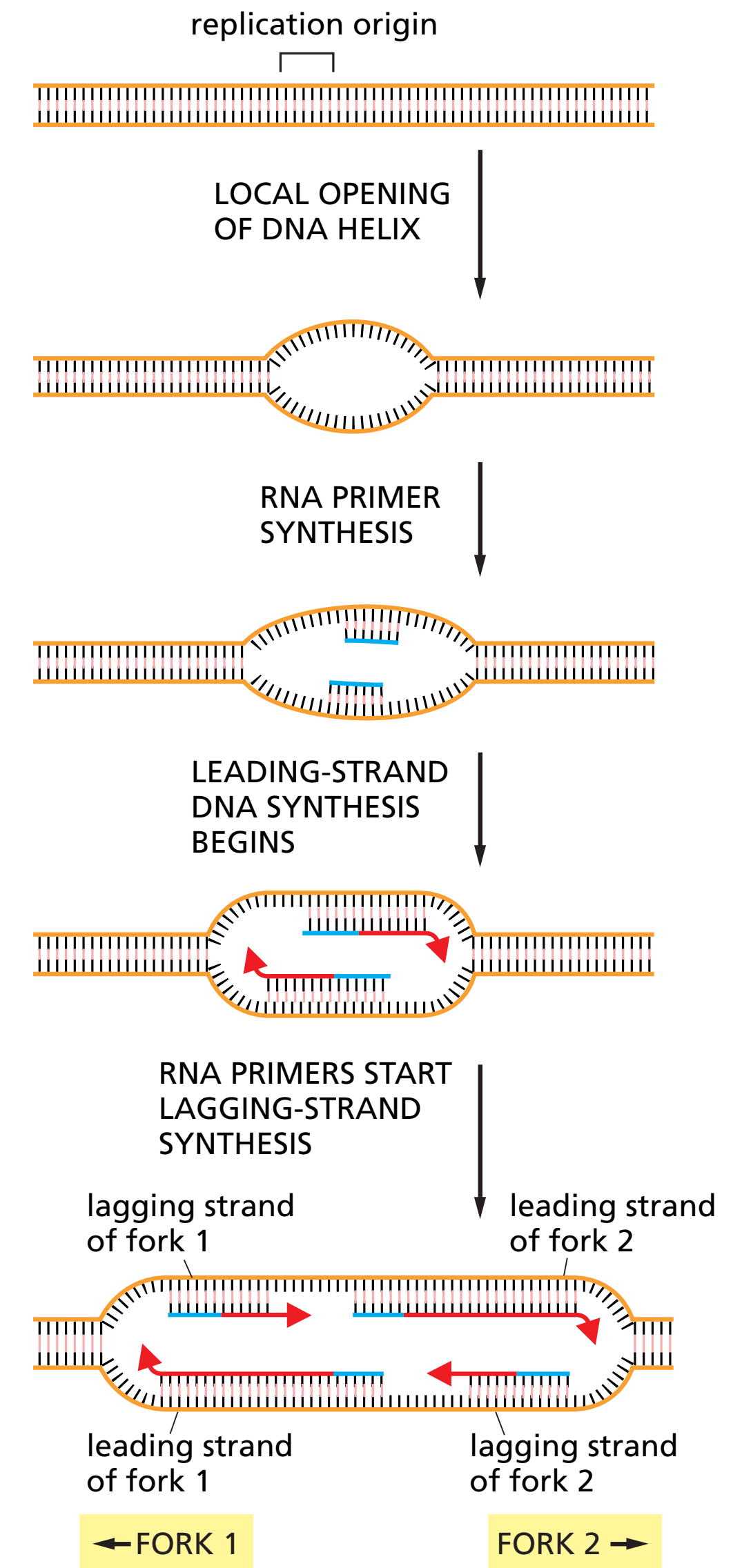- **More proteins** in Eukaryotes

Replication actors

- DNA polymerase

- Primase

- Ligase

- Helicase

- ssDNA binding proteins

- Sliding clamp

- DNA topoisomerase

Plan

- Quick recap

- DNA replication mechanisms

- <u>The initiation and completion of DNA replication in chromosomes</u>

- DNA repair

- Transposition and conservative site-specific recombination
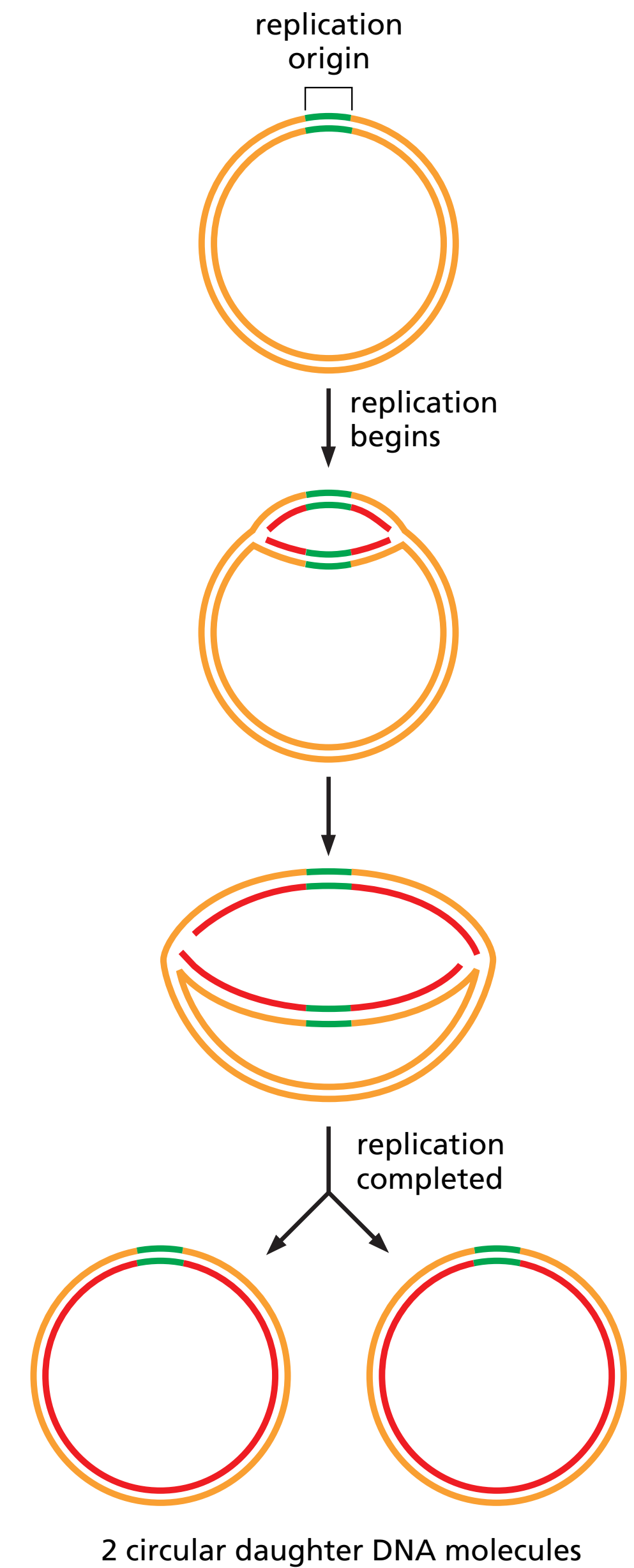
DNA synthesis begins at replication origins

- How does the whole process **start**?

- DNA double helix is **very stable**

- **Initiator proteins** bind to **replication origins** and open the two strands by breaking the hydrogen bonds

- In bacteria or yeasts, those are specific **DNA sequences** that attract initiator proteins (more AT-rich sequences that are held by less hydrogen bonds)

replication origin

LOCAL OPENING
OF DNA HELIX

RNA PRIMER
SYNTHESIS

LEADING-STRAND
DNA SYNTHESIS
BEGINS

RNA PRIMERS START
LAGGING-STRAND
SYNTHESIS

lagging strand
of fork 1

leading strand
of fork 2

leading strand
of fork 1

lagging strand
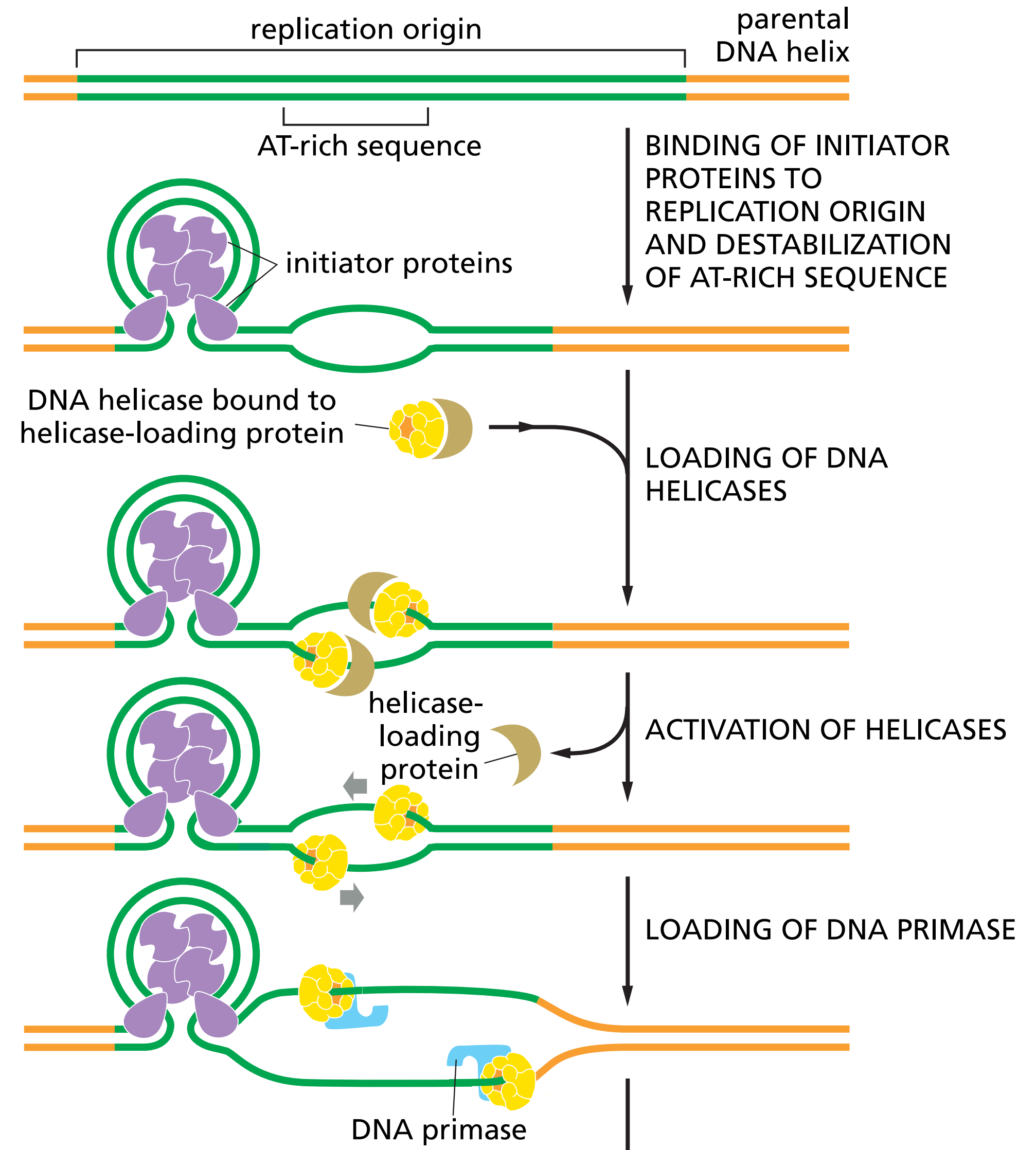of fork 2

◄— FORK 1

FORK 2 —►

In bacteria

- Reminder: bacteria typically have **one circular chromosome**

- Replication starts at **one origin of replication**

- Initiation of replication is **tightly regulated**

- Because they are growing **fast,** bacteria replicate their DNA **"continuously"**

replication
origin

replication
begins

replication
completed

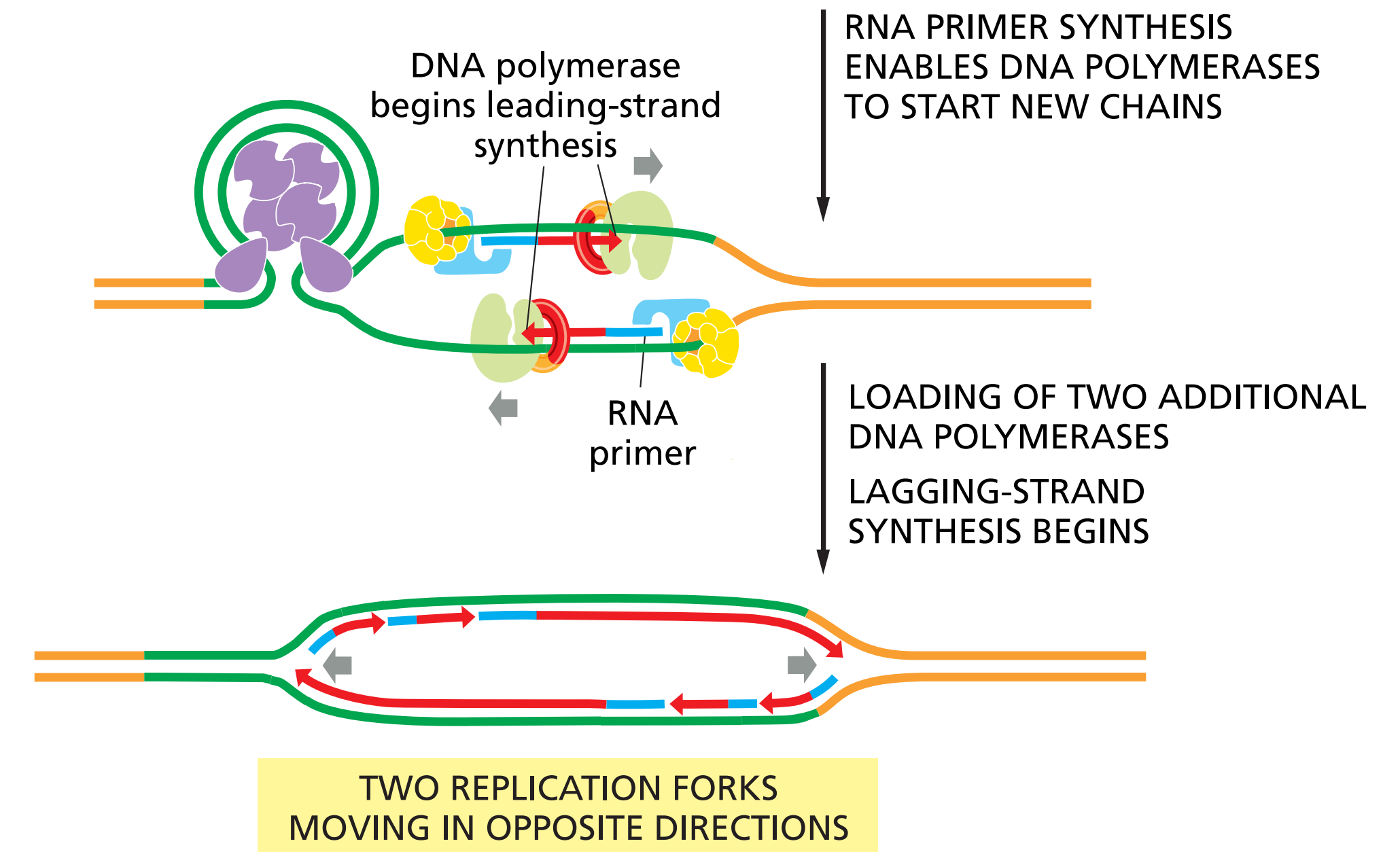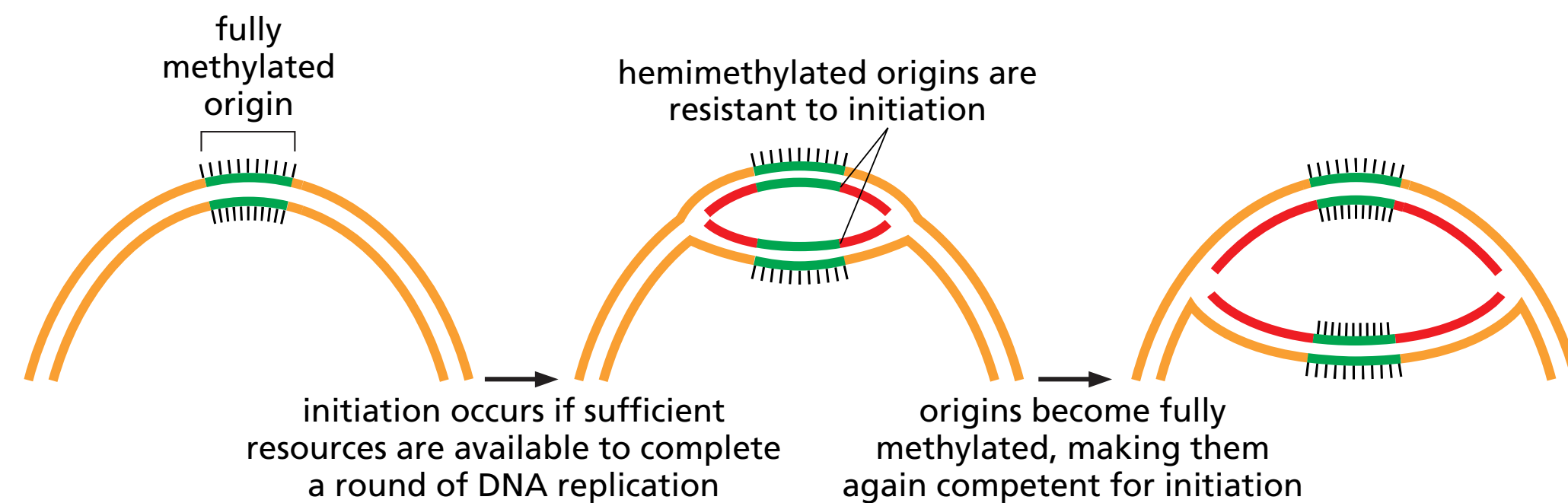2 circular daughter DNA molecules

In bacteria

1. **Initiator proteins** bind to the origin of replication

2. The helix is **destabilized**

3. **DNA helicases** are recruited

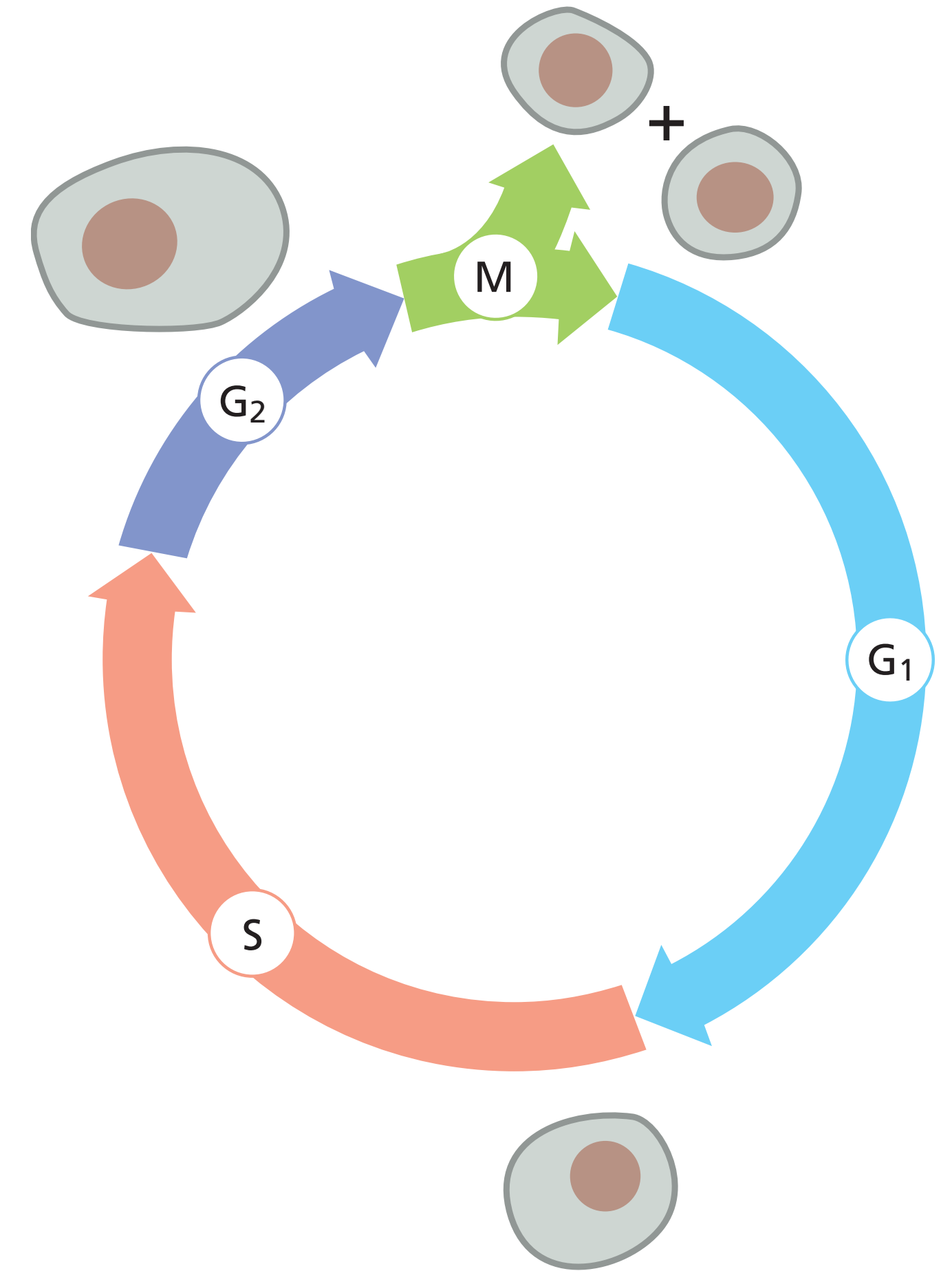4. **DNA primase** synthesizes the primers

In bacteria

5. Creation of **two replication forks** that move in opposite directions

6. **Inactivation** of the initiator proteins

7. **Refractory period** until origin is methylated

RNA PRIMER SYNTHESIS
ENABLES DNA POLYMERASES
TO START NEW CHAINS

DNA polymerase
begins leading-strand
synthesis

RNA
primer

LOADING OF TWO ADDITIONAL
DNA POLYMERASES

LAGGING-STRAND
SYNTHESIS BEGINS

TWO REPLICATION FORKS
MOVING IN OPPOSITE DIRECTIONS

fully
methylated
origin

hemimethylated origins are
resistant to initiation

initiation occurs if sufficient
resources are available to complete
a round of DNA replication

origins become fully
methylated, making them
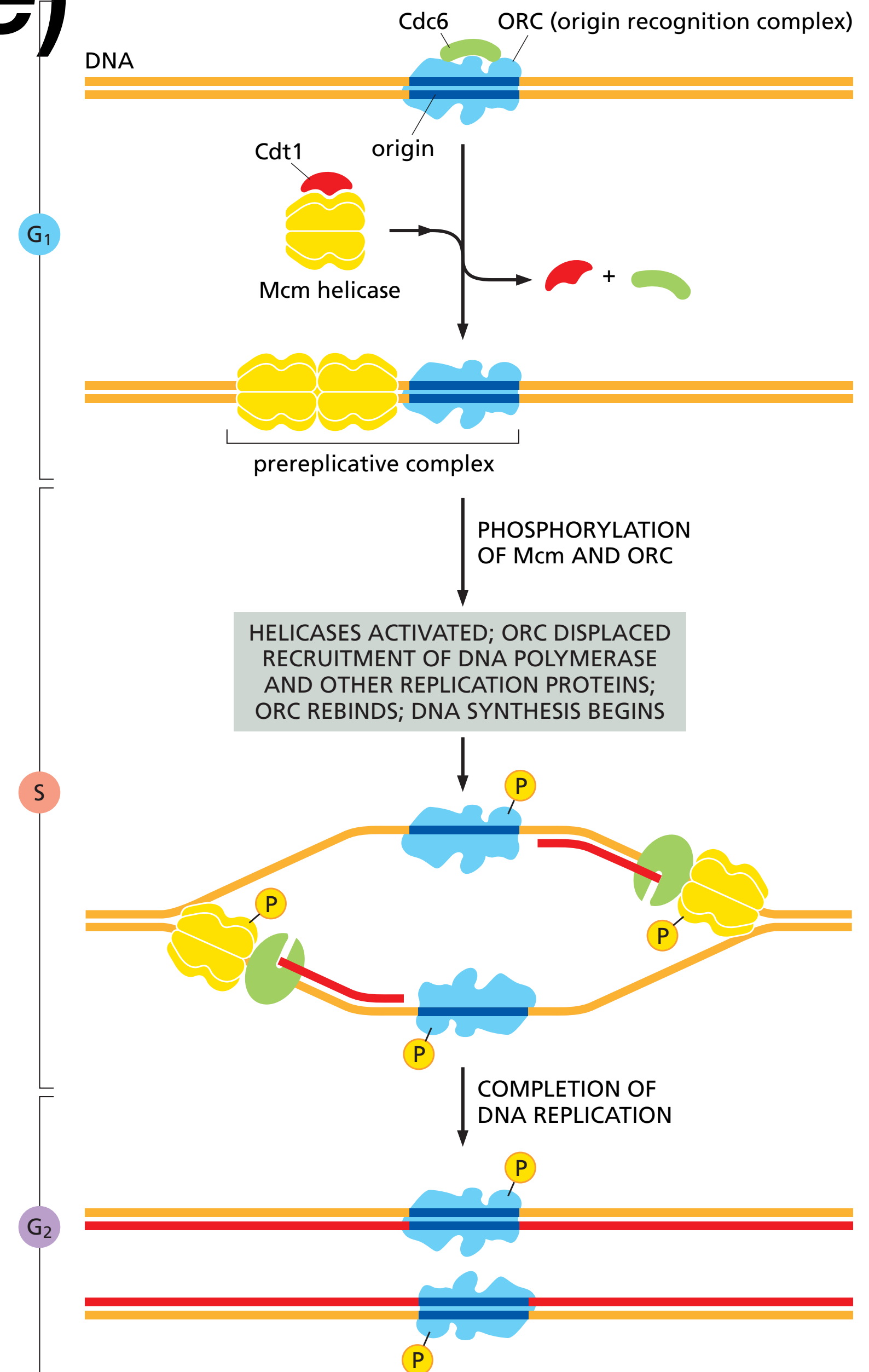again competent for initiation

In eukaryotes

MBoC6 m5.32/5.28

- **Multiple origins of replication** (~30 000 to ~50 000 for each cell division)

- The human genome has **more potential origins** than this; **different cells** use different origins

- **Two replication forks** are created at each origin (as in bacteria) and stop when they reach **another one** going in the opposite direction, or when they reach **the end of a chromosome**

- Replication only occurs during the **S-phase** of the cell cycle (8h in mammalian cells)
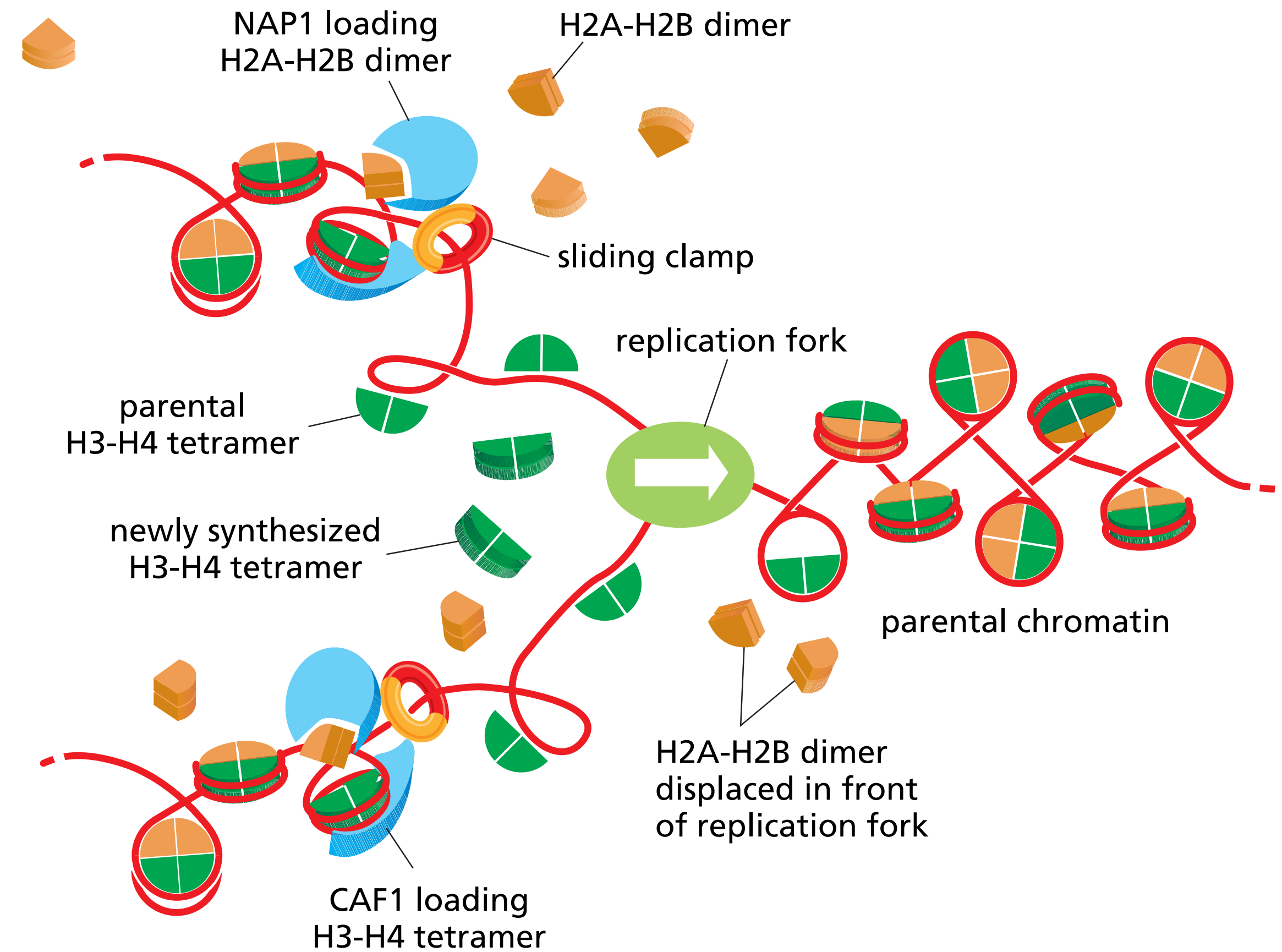
In eukaryotes (ex: *S. cerevisiae*)

- Binding of a large multi-subunit initiator protein called **ORC**

- **Phosphorylation of ORC by specific protein kinases** to prevent it from accepting other helicases

- This ensures only **one replication per cell cycle**

What about nucleosomes?

- **Chromatin-remodelling complexes destabilize** DNA-histones binding to allow the replication fork to move forward

- Broken into a H3-H4 tetramer and H2A or H2B dimers

- New histones are **synthesised** mostly in **S-phase**

- The assembly is promoted by **histone chaperones**

NAP1 loading H2A-H2B dimer

H2A-H2B dimer

sliding clamp

replication fork

parental H3-H4 tetramer

newly synthesized H3-H4 tetramer

parental chromatin

H2A-H2B dimer displaced in front of replication fork

CAF1 loading H3-H4 tetramer

The end of chromosomes
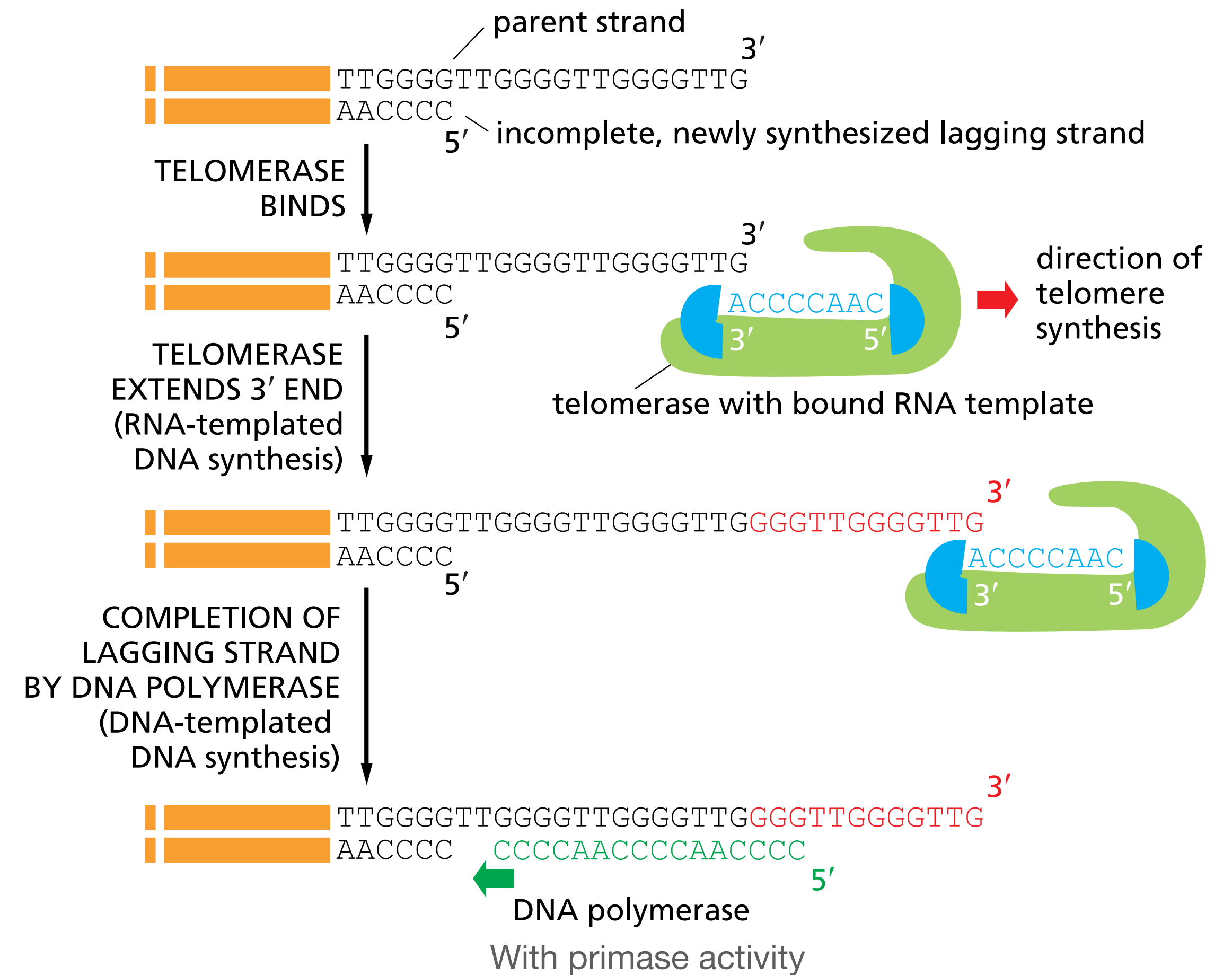
- At the lagging strand, the last RNA primer **cannot be replaced** by DNA (no primer for the DNA polymerase)

- Specific sequences at the end of the chromosomes = **telomeres**

- They contain **repeats** of a short sequence (~ 1000 times)

- In human, the sequence is GGGTTA

- Recognised by a **telomerase** that replenishes this sequence

- It uses an **RNA template** that it part of the telomerase itself

remainder of
telomerase RNA

telomerase protein
"fingers"

region of
telomerase RNA
used as template

3'

5'

"palm"—active site
of telomerase
protein

rest of
chromosome

newly
synthesized
telomere DNA

"thumb"

The end of chromosomes

- Similar to **reverse transcriptases** which make DNA using RNA templates

- Then, a **DNA primase/polymerase** finish strand synthesis

parent strand

3′

TTGGGGTTGGGGTTGGGGTTG
AACCCC

5′ incomplete, newly synthesized lagging strand

TELOMERASE
BINDS

3′

TTGGGGTTGGGGTTGGGGTTG
AACCCC

5′

ACCCCAAC
3′   5′

direction of
telomere
synthesis

telomerase with bound RNA template

TELOMERASE
EXTENDS 3′ END
(RNA-templated
DNA synthesis)

3′

TTGGGGTTGGGGTTGGGGTTGGGGTTGGGGTTG
AACCCC

5′

ACCCCAAC
3′   5′

COMPLETION OF
LAGGING STRAND
BY DNA POLYMERASE
(DNA-templated
DNA synthesis)

3′

TTGGGGTTGGGGTTGGGGTTGGGGTTGGGGTTG
AACCCC CCCCAACCCCAACCCC

5′

DNA polymerase

With primase activity

The end of chromosomes

- **Telomeres** must be distinguished in the cell from **accidental breaks** (otherwise the cells will try to repair them)

- Telomeres are protected by a protein cap called **sheltering**

- In somatic cells, **telomere gradually shorten -** this might be a regulation system for a given cell to **stop dividing**

Plan

- Quick recap

- DNA replication mechanisms

- The initiation and completion of DNA replication in chromosomes

- DNA repair

- Transposition and conservative site-specific recombination

DNA repair

- Tens of thousands **random changes in DNA** per human cell every day; less than 0.02% accumulate as permanent mutations

- The rest is repaired - very **important function**

- Initially studied in **bacteria**

- **Double helix** is ideally suited for repair (one strand remains intact)

- In general,

  1. Damage is **excised**

  2. Original **DNA sequence is restored** by a DNA polymerase
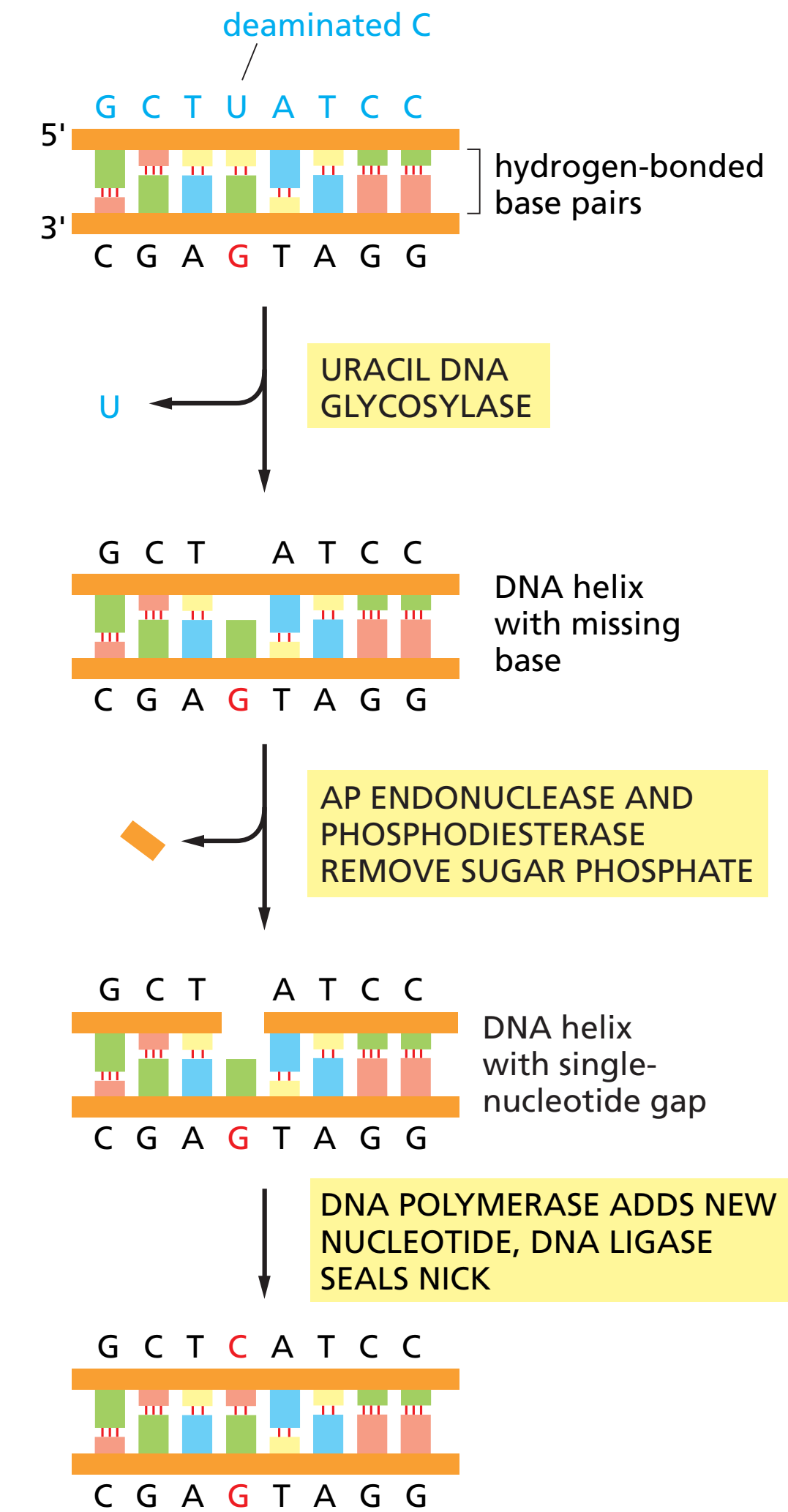
  3. **Sealed** by DNA ligase

DNA repair

- **Human diseases** associated with decreased repair

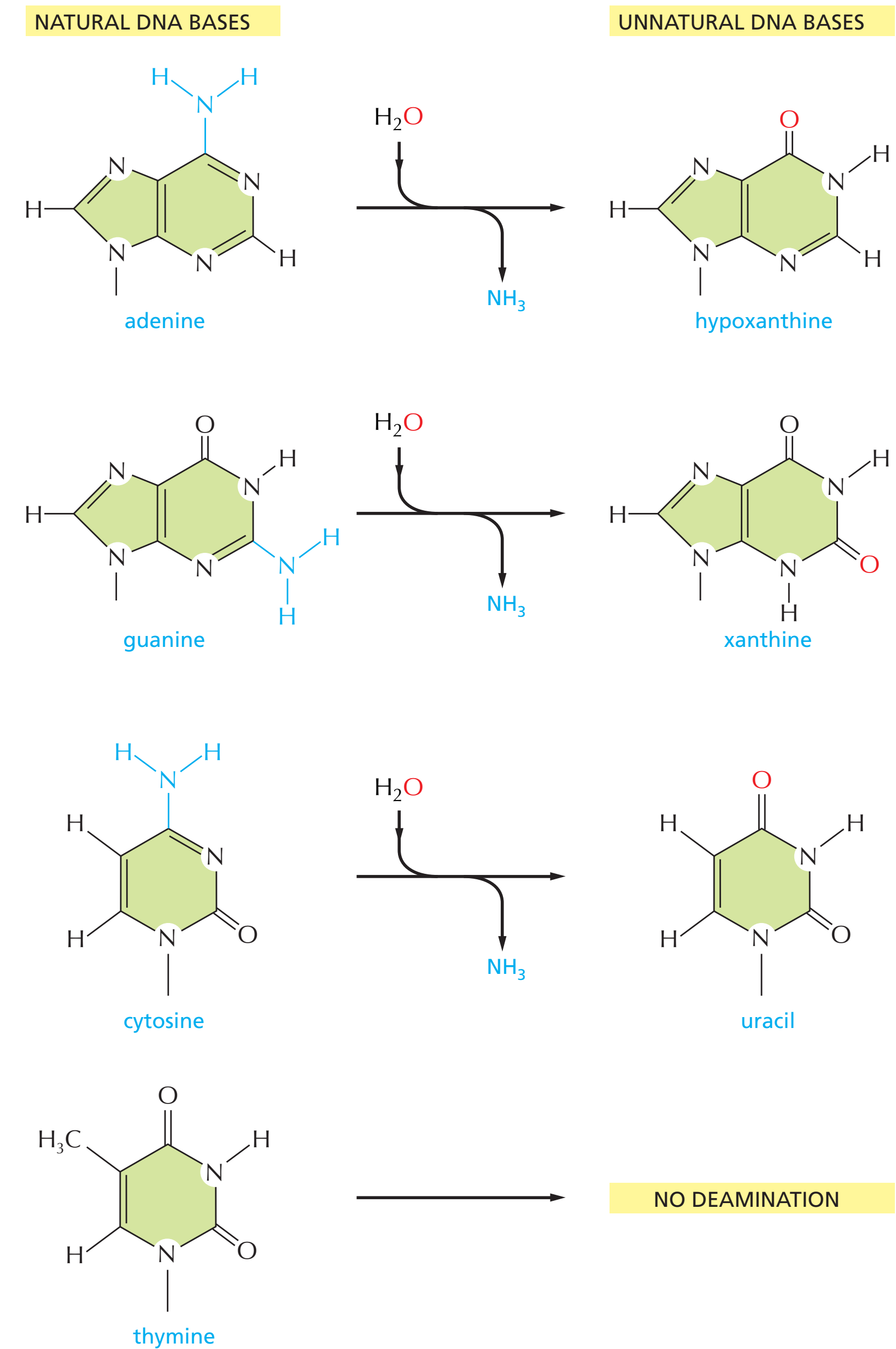| TABLE 5–2 Some Inherited Human Syndromes with Defects in DNA Repair		
Name	Phenotype	Enzyme or process affected
MSH2, 3, 6, MLH1, PMS2	Colon cancer	Mismatch repair
Xeroderma pigmentosum (XP) groups A–G	Skin cancer, UV sensitivity, neurological abnormalities	Nucleotide excision repair
Cockayne syndrome	UV sensitivity; developmental abnormalities	Coupling of nucleotide excision repair to transcription
XP variant	UV sensitivity, skin cancer	Translesion synthesis by DNA polymerase ν
Ataxia telangiectasia (AT)	Leukemia, lymphoma, γ-ray sensitivity, genome instability	ATM protein, a protein kinase activated by double-strand breaks
BRCA1	Breast and ovarian cancer	Repair by homologous recombination
BRCA2	Breast, ovarian, and prostate cancer	Repair by homologous recombination
Werner syndrome	Premature aging, cancer at several sites, genome instability	Accessory 3′-exonuclease and DNA helicase used in repair
Bloom syndrome	Cancer at several sites, stunted growth, genome instability	DNA helicase needed for recombination
Fanconi anemia groups A–G	Congenital abnormalities, leukemia, genome instability	DNA interstrand cross-link repair
46 BR patient	Hypersensitivity to DNA-damaging agents, genome instability	DNA ligase I

DNA repair mechanisms

(A) BASE EXCISION REPAIR

deaminated C

G C T U A T C C

5'

hydrogen-bonded base pairs

3'

C G A G T A G G

U ← URACIL DNA GLYCOSYLASE

G C T   A T C C

DNA helix with missing base

C G A G T A G G

AP ENDONUCLEASE AND PHOSPHODIESTERASE REMOVE SUGAR PHOSPHATE

G C T   A T C C

DNA helix with single-nucleotide gap

C G A G T A G G

DNA POLYMERASE ADDS NEW NUCLEOTIDE, DNA LIGASE SEALS NICK

G C T C A T C C

C G A G T A G G

- **Base-excision repair:** uses **DNA glycosylases** that recognise specific types

  of altered bases and remove them

DNA repair mechanisms

- The **chemistry of the bases** facilitates damage detection

- E.g. Every deamination leads to an **unnatural base**

adenine → hypoxanthine

guanine → xanthine

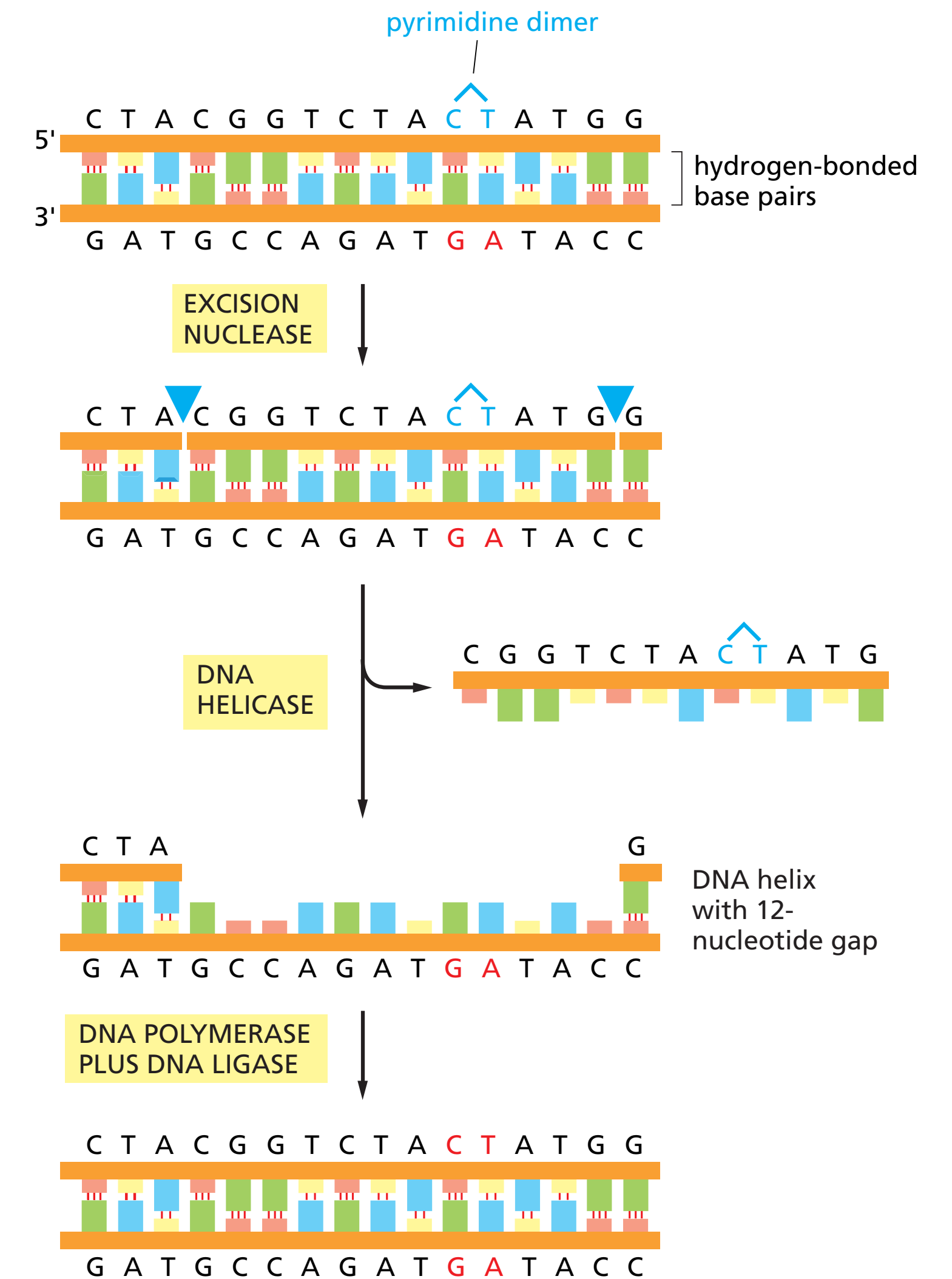cytosine → uracil

thymine → NO DEAMINATION

DNA repair mechanisms

- **Nucleotide-excision repair:**

  - Detects **distortions** of the double helix (rather than specific base-change)

  - Cleaves the phosphodiester backbone on both sides

  - A DNA helicase removes the damaged strand

  - DNA polymerase and ligase repair the DNA

(B)   NUCLEOTIDE EXCISION REPAIR

pyrimidine dimer

C T A C G G T C T A C T A T G G
5'
hydrogen-bonded base pairs
3'
G A T G C C A G A T G A T A C C

EXCISION NUCLEASE

C T A C G G T C T A C T A T G G
G A T G C C A G A T G A T A C C

DNA HELICASE

C G G T C T A C T A T G

C T A   G
DNA helix with 12-nucleotide gap
G A T G C C A G A T G A T A C C

DNA POLYMERASE PLUS DNA LIGASE

C T A C G G T C T A C T A T G G
G A T G C C A G A T G A T A C C

DNA repair mechanisms

- How to **direct DNA repair** machineries to **"important" DNA regions**?

  - Coupled with **RNA polymerase** (in charge of transcription)

  - When RNA polymerase stalls at DNA lesions, it **recruits DNA repair**

  - People that lack this coupled action (Cockayne Syndrome) have RNA polymerase permanently stalled on their DNA
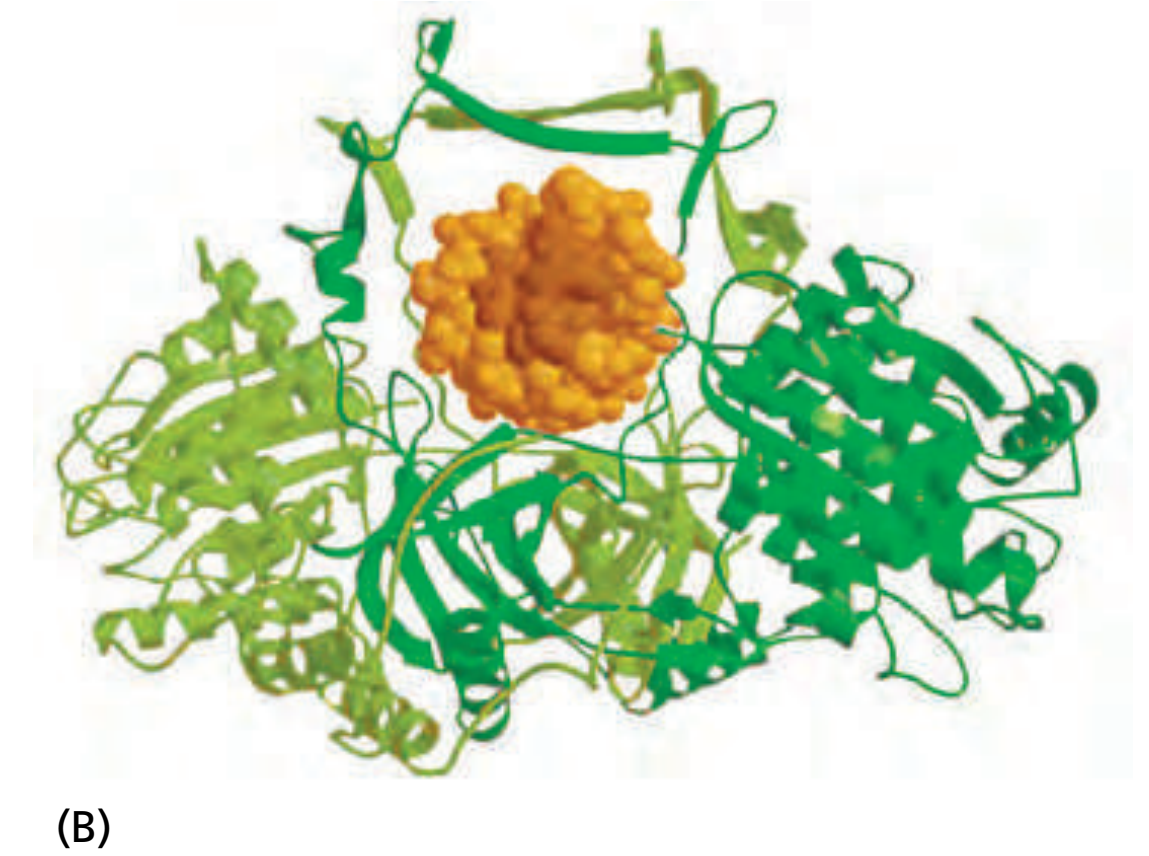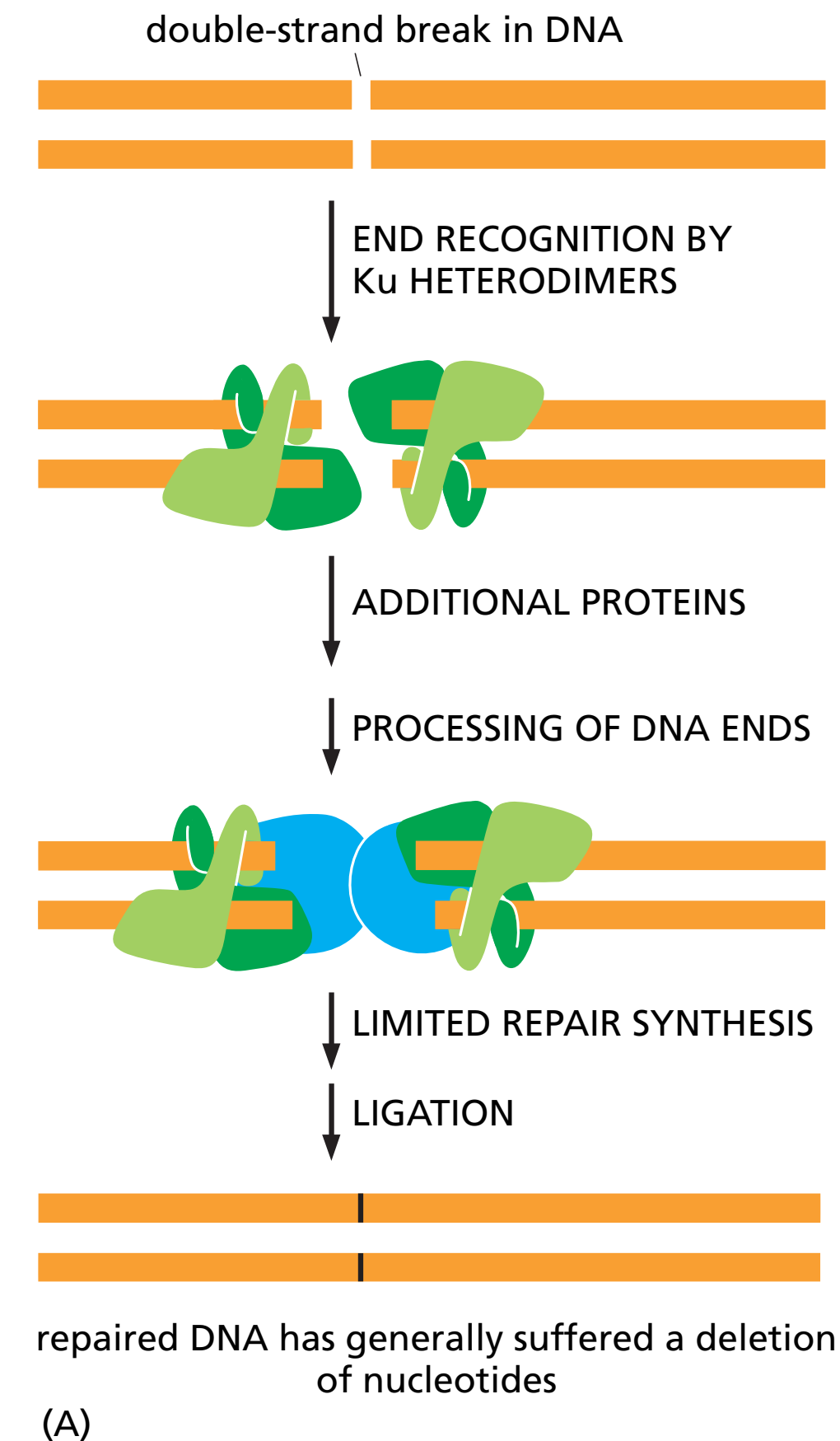
Cockayne syndrome	UV sensitivity; developmental abnormalities	Coupling of nucleotide excision repair to transcription

DNA repair mechanisms

- What about **double strand breaks**? No intact template DNA

  1. Non-homologous end joining

  - Broken ends are brought together

  - Ligated

  - Lost of a small DNA sequence

  - "Quick and dirty"

  - Danger: chromosome rearrangement as the system has no way to know if the sequences were together initially
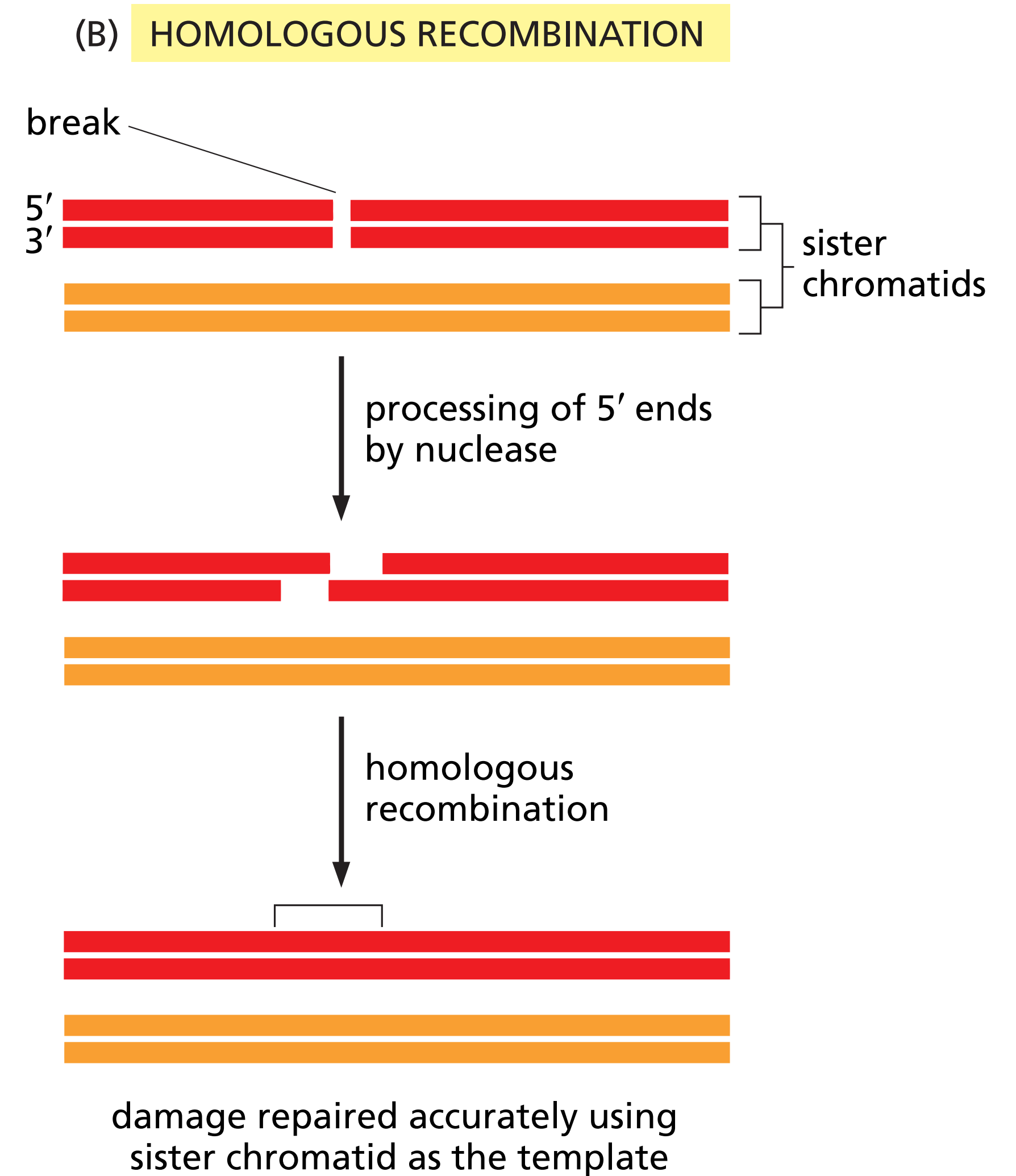
  - Telomeres are protected from this

double-strand break in DNA

END RECOGNITION BY Ku HETERODIMERS

ADDITIONAL PROTEINS

PROCESSING OF DNA ENDS

LIMITED REPAIR SYNTHESIS

LIGATION

repaired DNA has generally suffered a deletion of nucleotides

(A)

(B)

DNA repair mechanisms

- What about **double strand breaks**? No intact template DNA
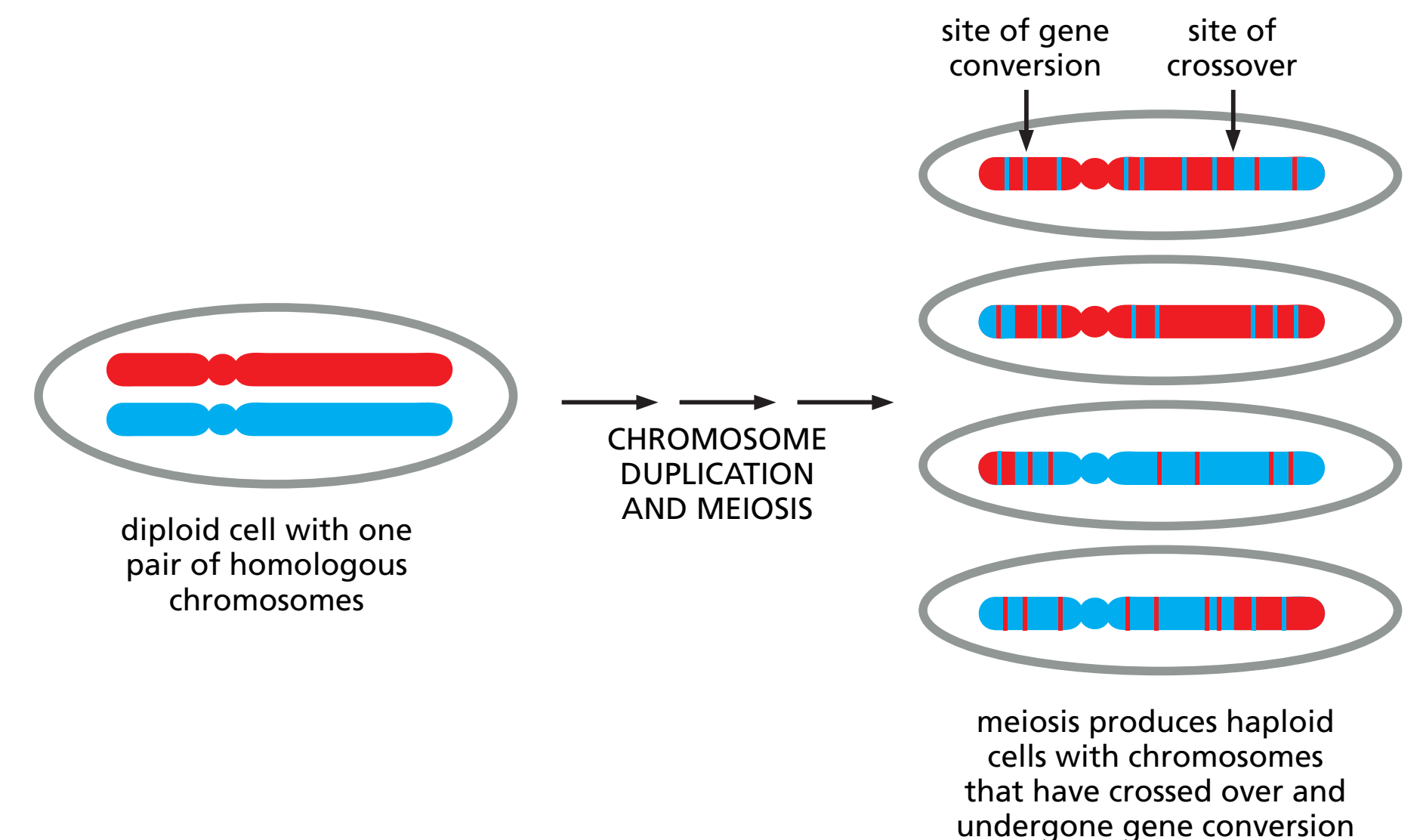
  2. Homologous recombination

  - More accurate

  - The sister chromatid is used as a template

  - In human, only occurs shortly after DNA replication when sister chromatids are available to serve as template

  - More info later

(B) HOMOLOGOUS RECOMBINATION

break

5′
3′

sister chromatids

processing of 5′ ends by nuclease

homologous recombination

damage repaired accurately using sister chromatid as the template

What is homologous recombination ?

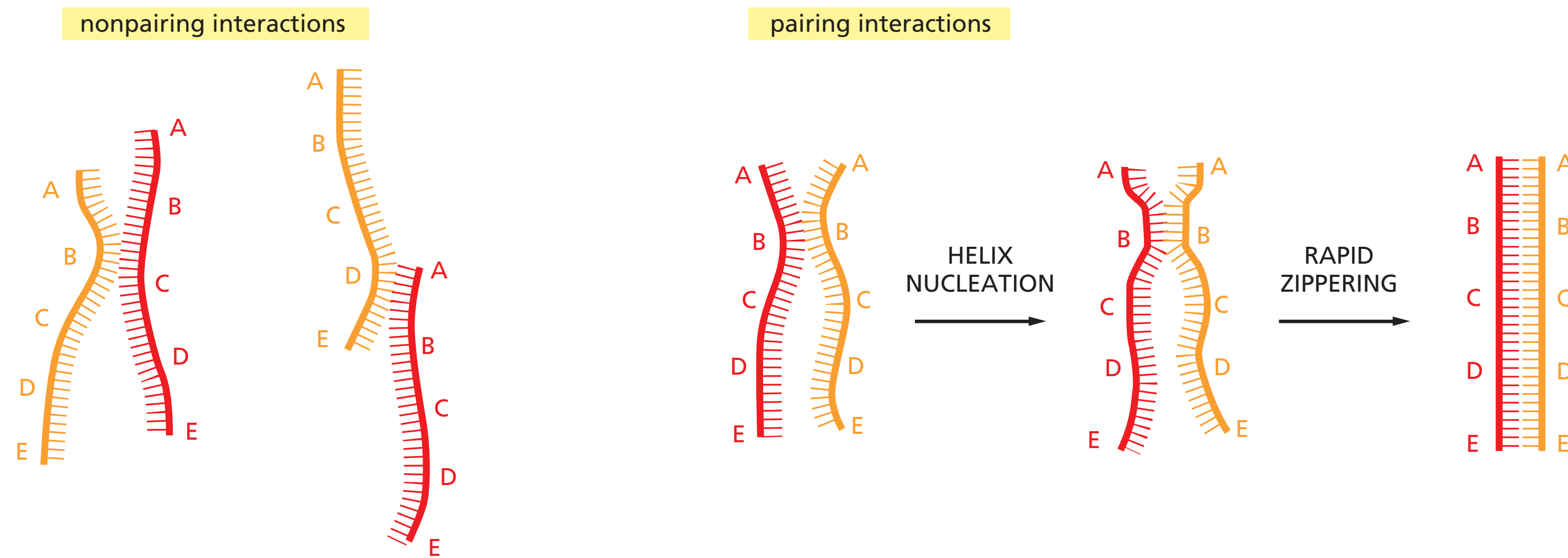**Homologous recombination** is the exchange of DNA strands between a pair of homologous duplex DNA sequences

- Can repair many types of **DNA damage** (e.g. double strand breaks)

- **Versatile** mechanism **conserved** in all cells

- During **meiosis,** plays a key role in gamete production (exchange of bits maternal and paternal genetic material) to create new combinations to be passed on to the offspring

- Most of what we know derives from **bacteria** and **viruses**

site of gene conversion   site of crossover

diploid cell with one pair of homologous chromosomes

CHROMOSOME DUPLICATION AND MEIOSIS

meiosis produces haploid cells with chromosomes that have crossed over and undergone gene conversion
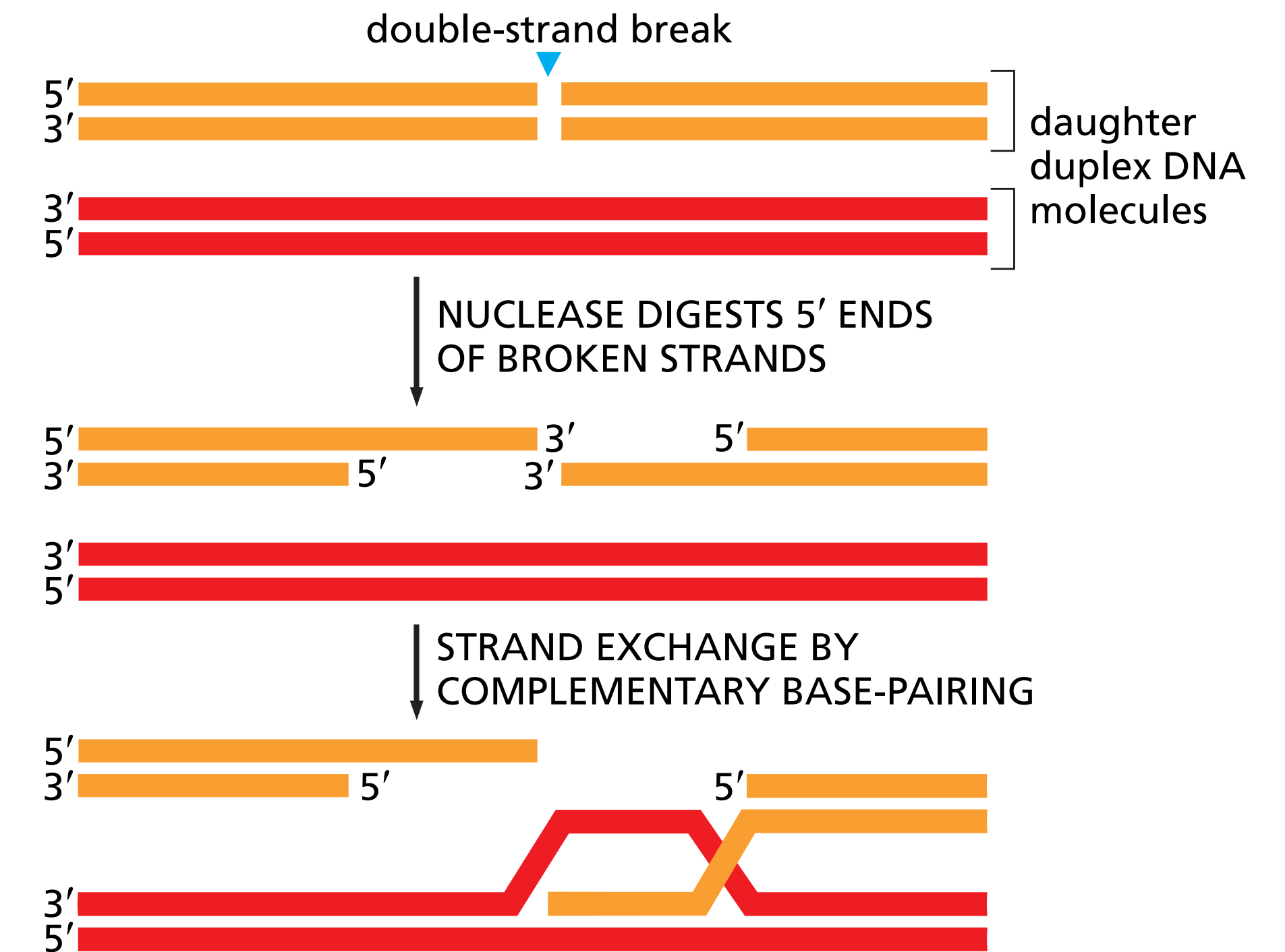
What is homologous recombination ?

Homologous recombination is the exchange of DNA strands between a pair of homologous duplex DNA sequences

- Only occurs between DNA duplexes that have extensive regions of sequence similarity (homology)

- Concept of **DNA hybridization:**
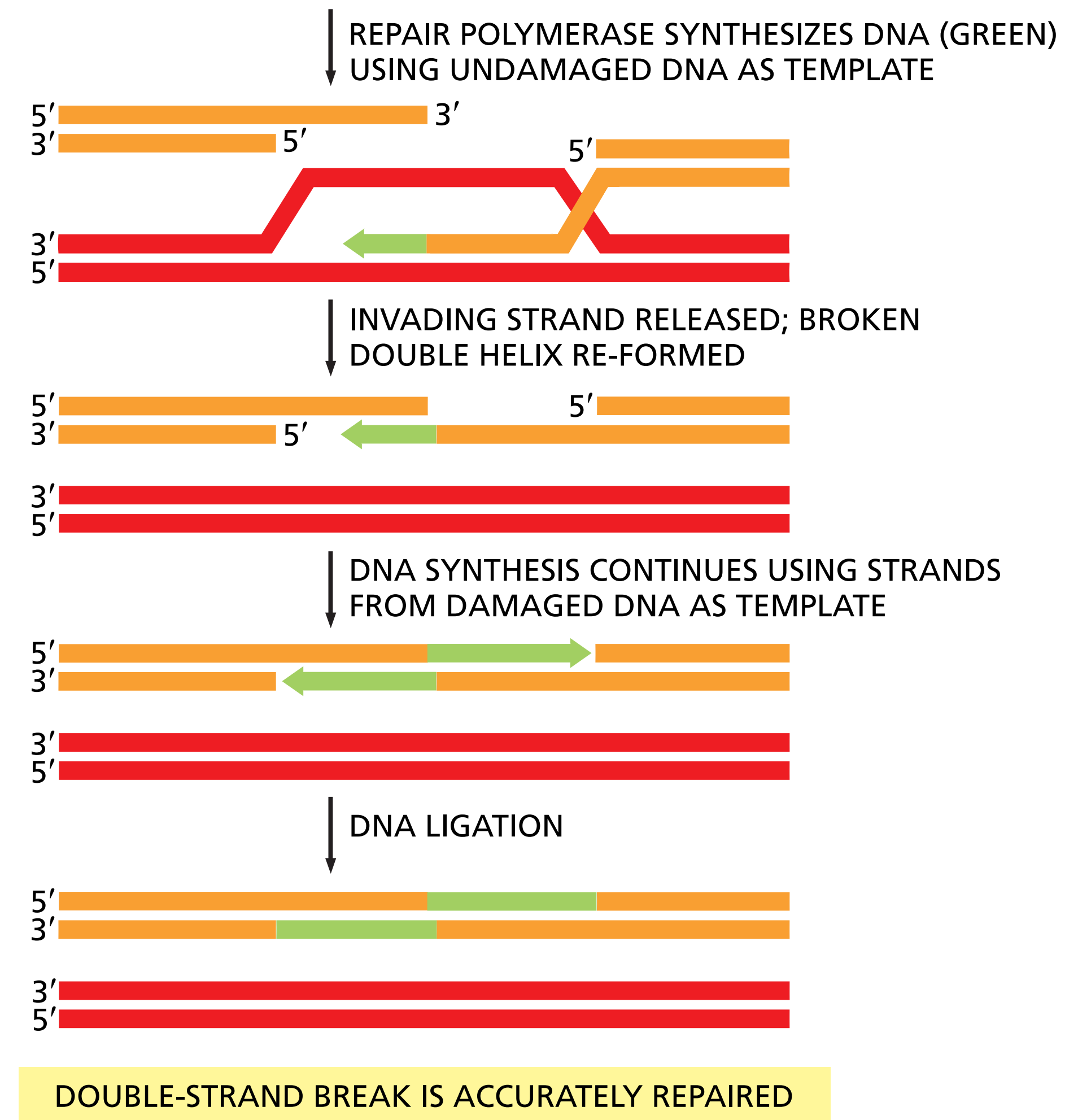
What is homologous recombination ?

1. The ends of the broken DNA are chewed back by **nucleases** to produce **overhanging single-strand 3' ends**

2. **Strand exchange** then happens: the single-strand 3' end searches for a homologous sequence through base-pairing
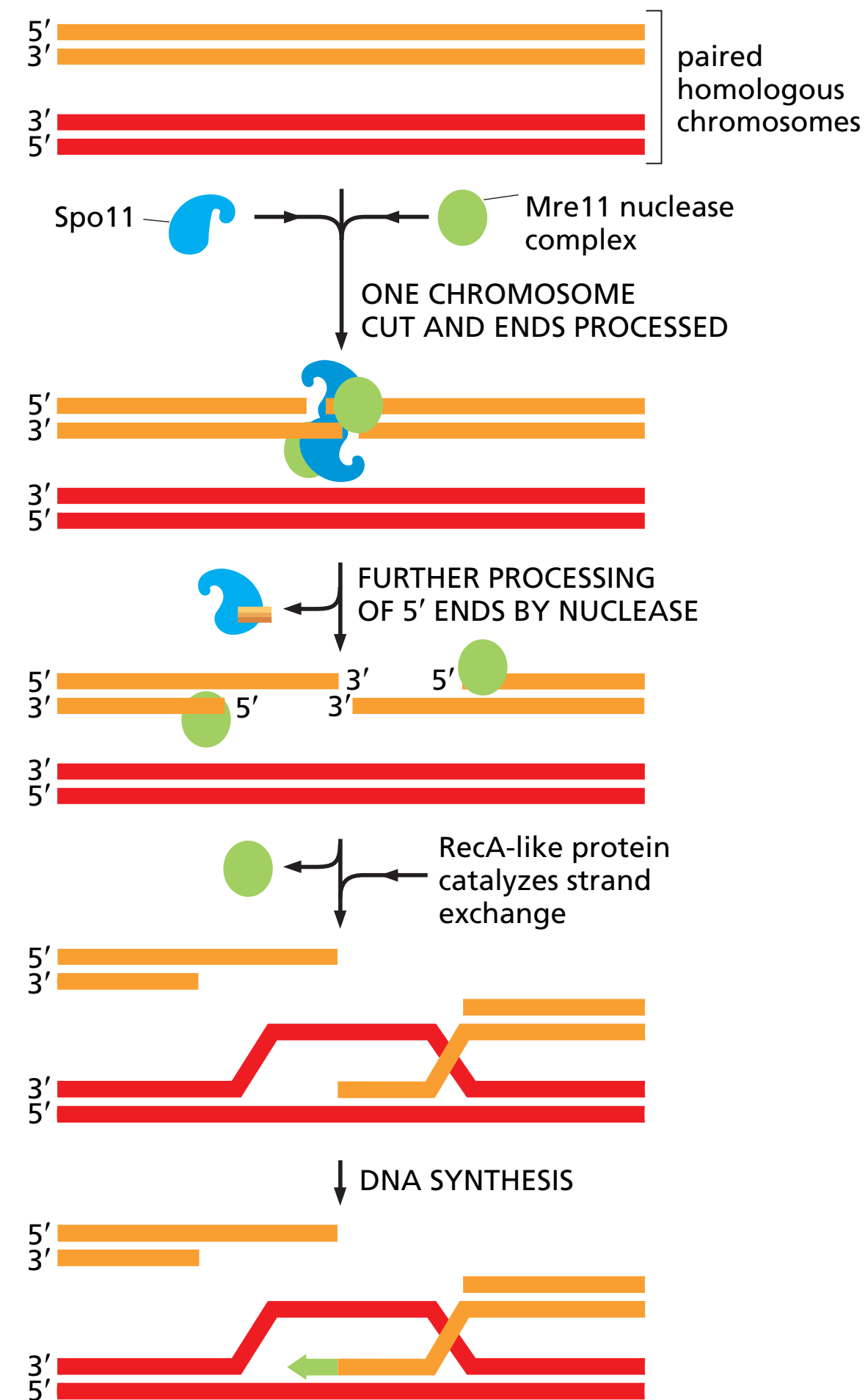
What is homologous recombination ?

3. **DNA polymerase** extends the invading strand

4. **Ligation** restore the two initial DNA duplexes

REPAIR POLYMERASE SYNTHESIZES DNA (GREEN)
USING UNDAMAGED DNA AS TEMPLATE

5′ — 3′
3′ — 5′   5′   5′

3′
5′

INVADING STRAND RELEASED; BROKEN
DOUBLE HELIX RE-FORMED

5′   5′
3′   5′

3′
5′

DNA SYNTHESIS CONTINUES USING STRANDS
FROM DAMAGED DNA AS TEMPLATE

5′
3′

3′
5′

DNA LIGATION

5′
3′

3′
5′

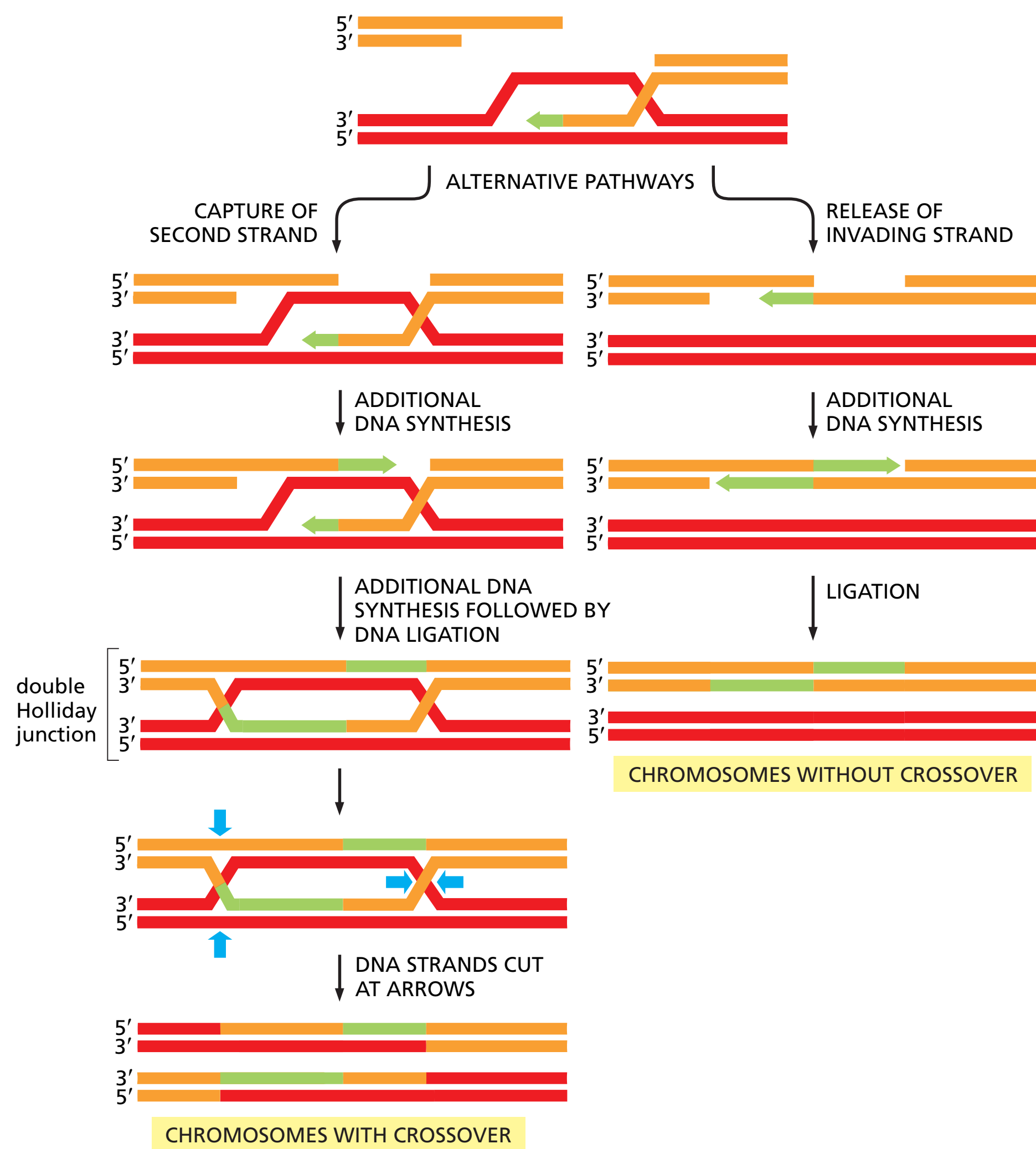DOUBLE-STRAND BREAK IS ACCURATELY REPAIRED

What is homologous recombination ?

Homologous recombination during meiosis can generate **chromosome crossovers**

What is homologous recombination ?

Homologous recombination during meiosis can generate **chromosome crossovers**
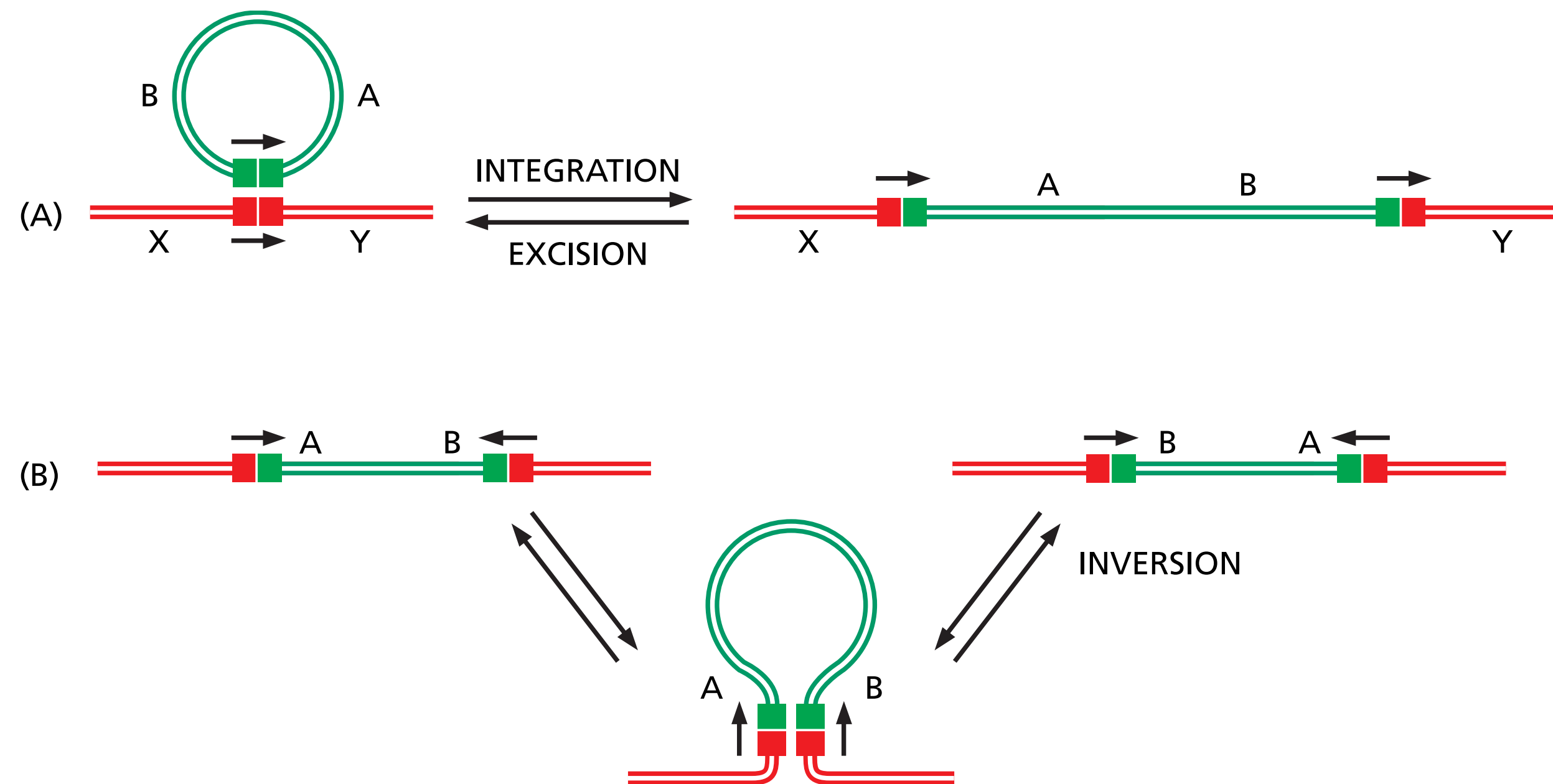
Conservative site-specific recombination

- For other types of mobile genetic elements

- Breakage and joining happens a specific sequences

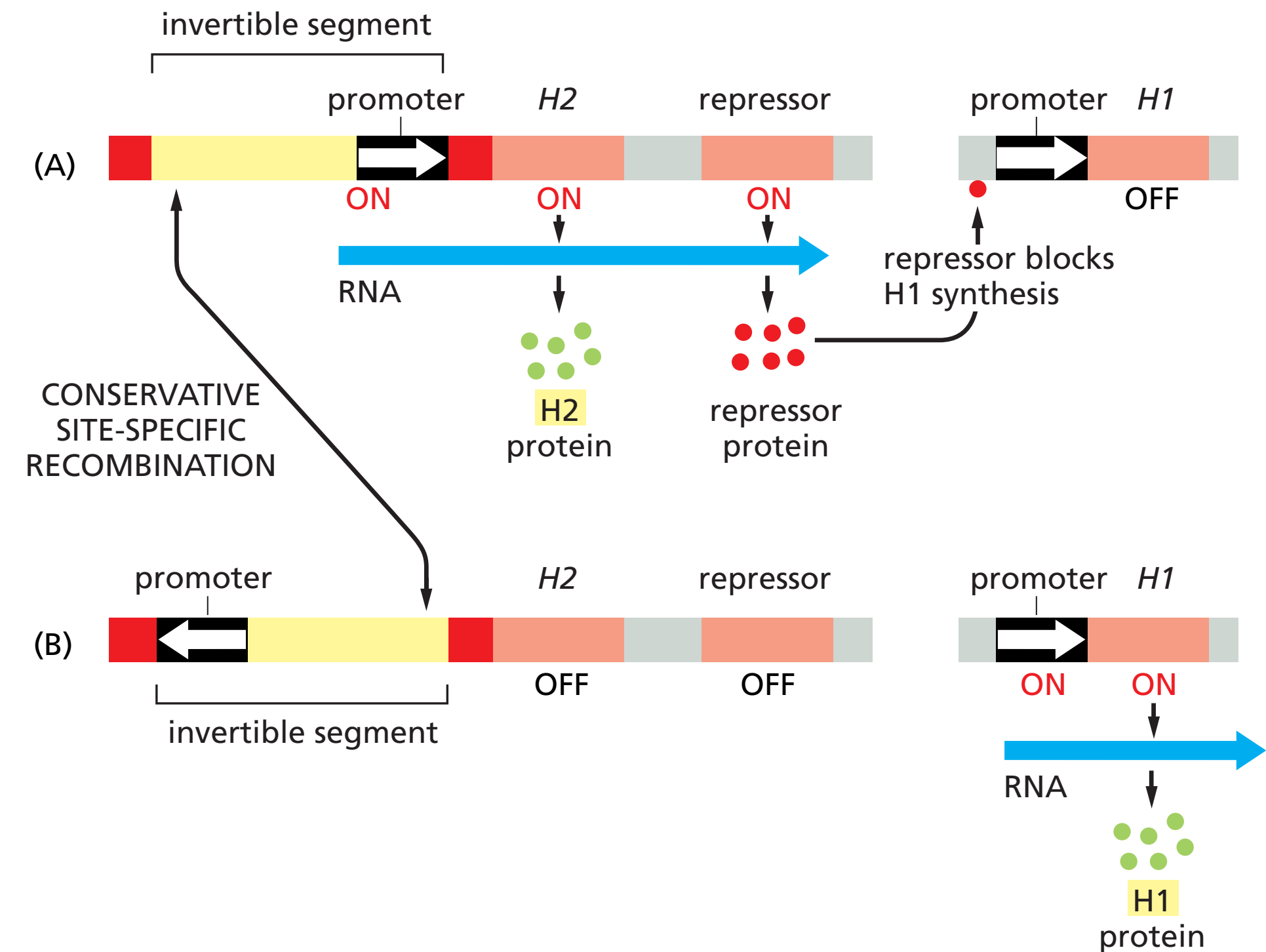- DNA integration, excision or inversion can occur

- Also used by some viruses

Differences with transposition

- Specific end-sequences on both donor and recipient DNA
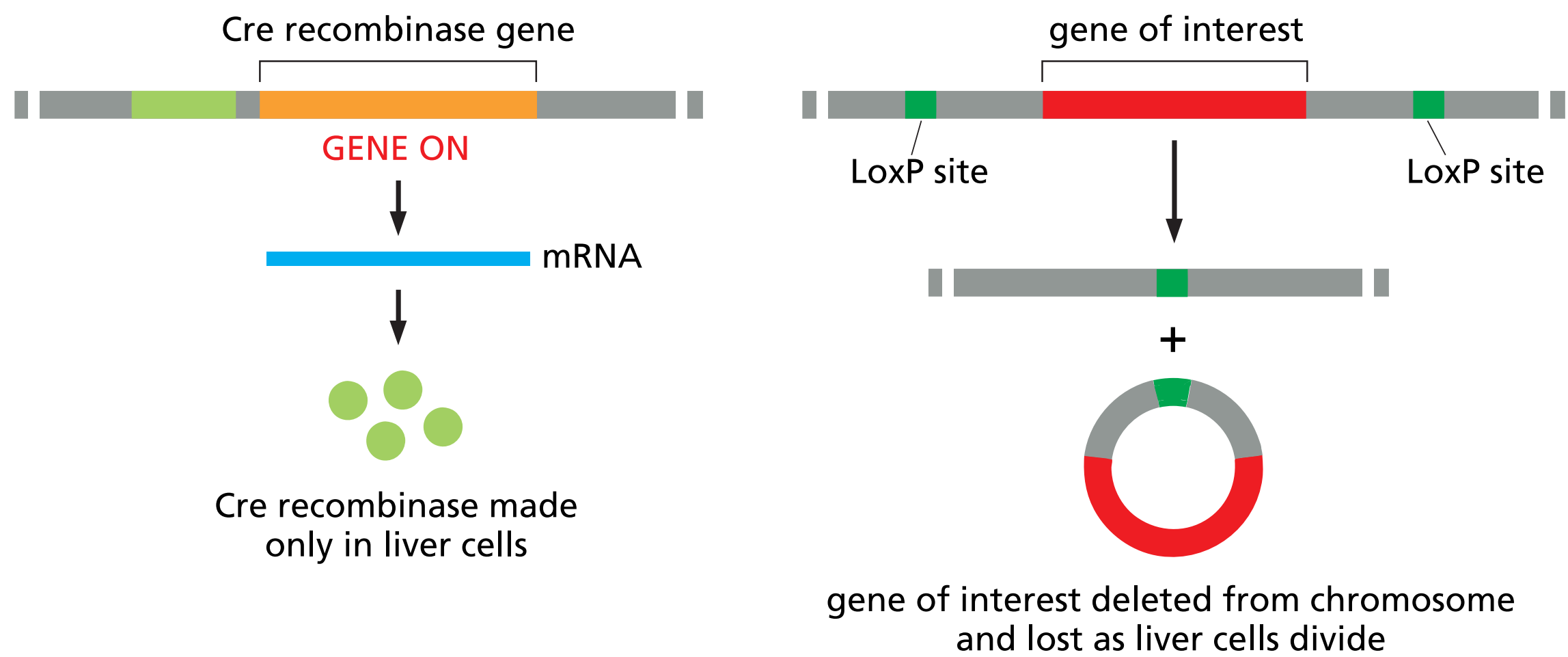
- Reaction mechanism are different

Conservative site-specific recombination

- Can be used to switch genes on and off in bacteria (phenomenon called **phase variation**)

- Inversion of piece of DNA that includes the promoter

- Alters the expression of gene (e.g. cell surface flagellin for which bacteria have two genes, H1 and H2)

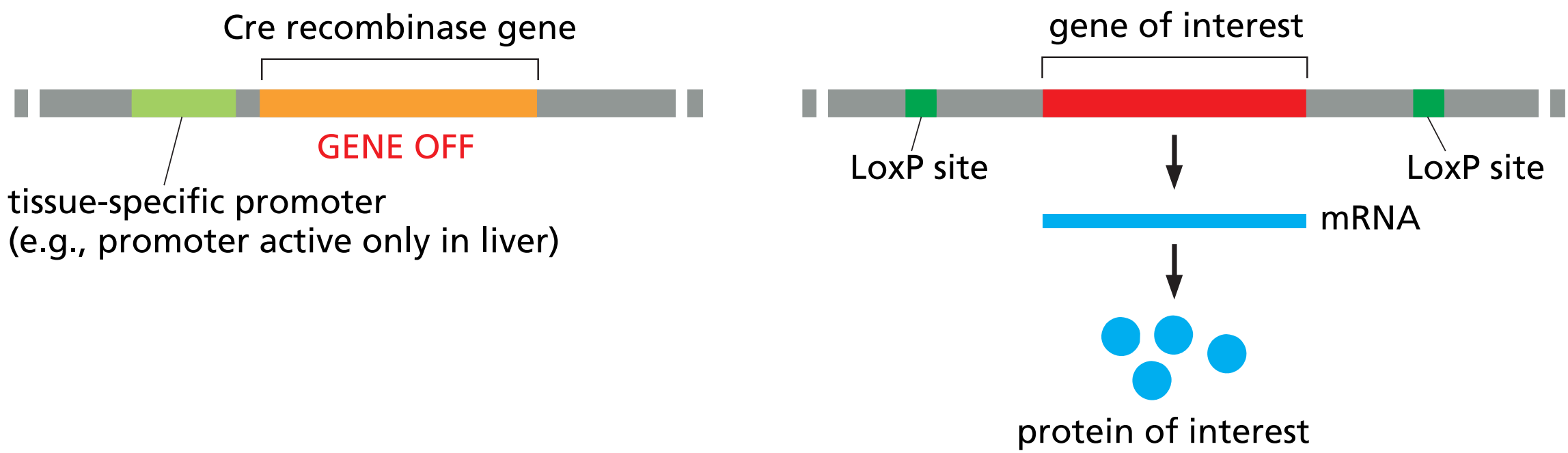- Protects bacteria from immune response

Conservative site-specific recombination

IN SPECIFIC TISSUE (e.g., LIVER)

Cre recombinase gene

GENE ON

mRNA

Cre recombinase made
only in liver cells

gene of interest

LoxP site   LoxP site

+

gene of interest deleted from chromosome
and lost as liver cells divide

IN OTHER TISSUES, THE GENE OF INTEREST IS EXPRESSED NORMALLY

Cre recombinase gene

GENE OFF

tissue-specific promoter
(e.g., promoter active only in liver)

gene of interest

LoxP site   LoxP site

mRNA

protein of interest

• Powerful tool for developmental biologists

Have a nice day!