

DNA REPAIR

Mechanisms of DNA Damage Recognition and Correction, with Prokaryotic and Eukaryotic Comparisons

1. Introduction: Why DNA Repair Matters

DNA is under constant chemical assault, both from normal cellular metabolism and from the external environment. It has been estimated that a single human cell suffers tens of thousands of DNA-damaging events every day. Left unrepaired, this damage causes mutations, blocks replication and transcription, and can drive cancer, premature ageing, and cell death. Cells therefore maintain a battery of **DNA repair pathways**, each specialised for a different class of damage.

(a) Common Types of DNA Damage

- **Spontaneous depurination/depyrimidination** — hydrolysis of the N-glycosidic bond removes a base, leaving an abasic (AP) site.
- **Deamination** — cytosine converts to uracil (and 5-methylcytosine to thymine), a common and highly mutagenic spontaneous change if uncorrected before replication.
- **Oxidative damage** — reactive oxygen species (ROS) generated by normal metabolism produce lesions such as 8-oxoguanine, which mispairs with adenine.
- **UV-induced damage** — ultraviolet light causes adjacent pyrimidines to fuse into **cyclobutane pyrimidine dimers** or 6-4 photoproducts, distorting the double helix.
- **Alkylation and bulky adducts** — chemical mutagens (e.g. from tobacco smoke) add alkyl or bulky chemical groups to bases.
- **Replication errors** — DNA polymerases occasionally misincorporate a nucleotide, creating a base-base mismatch.
- **Double-strand breaks (DSBs)** — both strands of the helix are severed at the same point, caused by ionising radiation, replication fork collapse, or reactive chemicals; these are the most dangerous form of damage.

The general logic of repair

Nearly every repair pathway follows the same basic logic: (1) **recognise** the damage or distortion, (2) **remove** the damaged section (a single base, a short patch, or process broken ends), (3) **resynthesise** the correct sequence using the intact complementary or homologous strand as a template, and (4) **seal** the remaining nick with DNA ligase.

2. Direct Reversal Repair

A small number of lesions can be corrected in a single enzymatic step, without removing any nucleotides from the DNA backbone at all.

- **Photoreactivation** — the enzyme **photolyase** uses energy absorbed from visible (blue) light to directly break the covalent bonds of a UV-induced pyrimidine dimer, restoring the original bases in place. Photolyase is found in bacteria, fungi, plants, and many animals, but is **absent in placental mammals**, including humans, who must rely on nucleotide excision repair for these lesions instead.

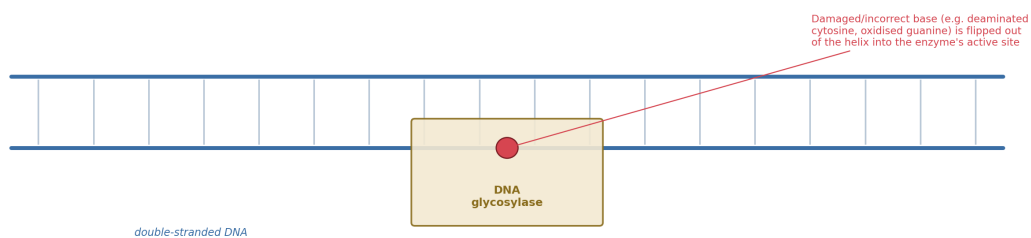
- **Alkyltransferase repair** — the enzyme **O⁶-methylguanine methyltransferase (MGMT)** directly removes an alkyl (e.g. methyl) group from damaged O⁶-methylguanine, transferring it onto one of its own cysteine residues. This permanently inactivates the enzyme after a single use, making MGMT a "suicide enzyme" that is not catalytically recycled.

3. Base Excision Repair (BER)

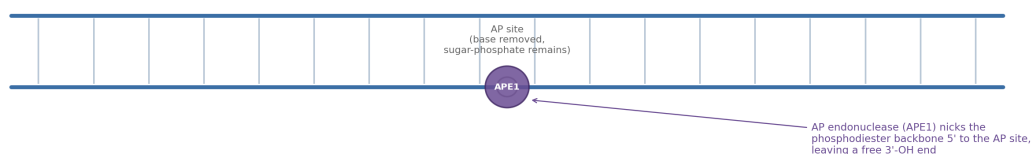
Base excision repair corrects small, non-helix-distorting base lesions — such as deaminated, oxidised, or alkylated bases — that arise mainly from spontaneous chemistry and normal metabolism. It is highly conserved, working the same basic way in bacteria and eukaryotes.

Base Excision Repair (BER) — Step by Step

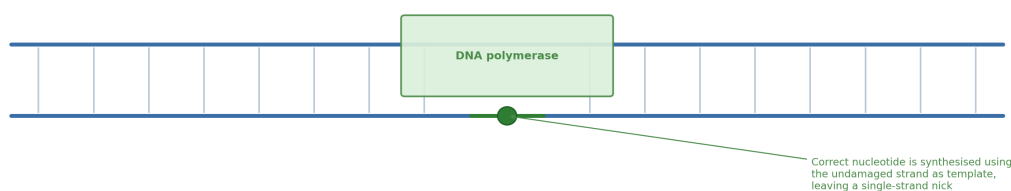
1. **Damage recognition** — a DNA glycosylase scans DNA and flips out the damaged base



2. **Base removal and incision** — an AP (apurinic/apyrimidinic) site is cut by AP endonuclease



3. **Gap filling** — DNA polymerase (Pol I in bacteria; Pol β in eukaryotes) inserts the correct base



4. **Sealing** — DNA ligase (LigI/LigIII) seals the remaining nick, restoring an intact duplex

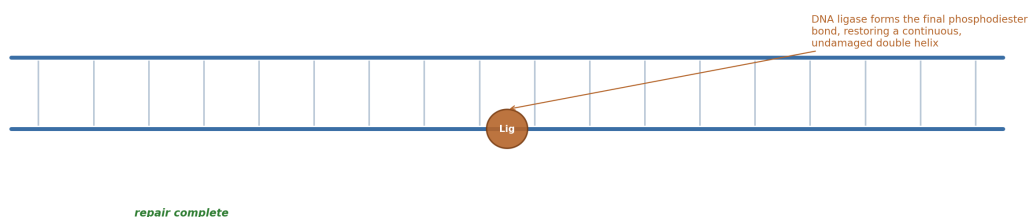


Fig. 1 — Base excision repair: a damaged base is flipped out and removed by a DNA glycosylase, the resulting AP site is incised, the gap is filled, and the nick is sealed.

- A **DNA glycosylase** — there are several, each recognising a specific type of damaged base (e.g. uracil-DNA glycosylase for deaminated cytosine, OGG1 for 8-oxoguanine) — flips the damaged base out of the helix and cleaves the N-glycosidic bond, leaving an **AP (apurinic/apyrimidinic) site**.
- **AP endonuclease** (APE1 in humans) nicks the phosphodiester backbone immediately 5' of the AP site.
- **DNA polymerase** (Pol I in bacteria, which also removes the remaining abasic sugar-phosphate with its 5'→3' exonuclease activity; Pol β in eukaryotes) fills the single-nucleotide gap using the intact complementary strand as template.
- **DNA ligase** (LigA in bacteria; LigIII, with the scaffold protein XRCC1, in eukaryotes for short-patch BER) seals the remaining nick, restoring an intact duplex.
- A **long-patch BER** variant exists in eukaryotes, in which Pol δ/ϵ (with PCNA) displaces and resynthesises a stretch of 2-10 nucleotides, and the displaced flap is removed by FEN1 before ligation by LigI.

4. Nucleotide Excision Repair (NER)

Nucleotide excision repair removes **bulky, helix-distorting lesions** that BER glycosylases cannot recognise — most importantly UV-induced pyrimidine dimers and bulky chemical adducts. Rather than excising a single base, NER cuts out a short single-stranded oligonucleotide patch containing the lesion.

Nucleotide Excision Repair: Prokaryotic vs Eukaryotic

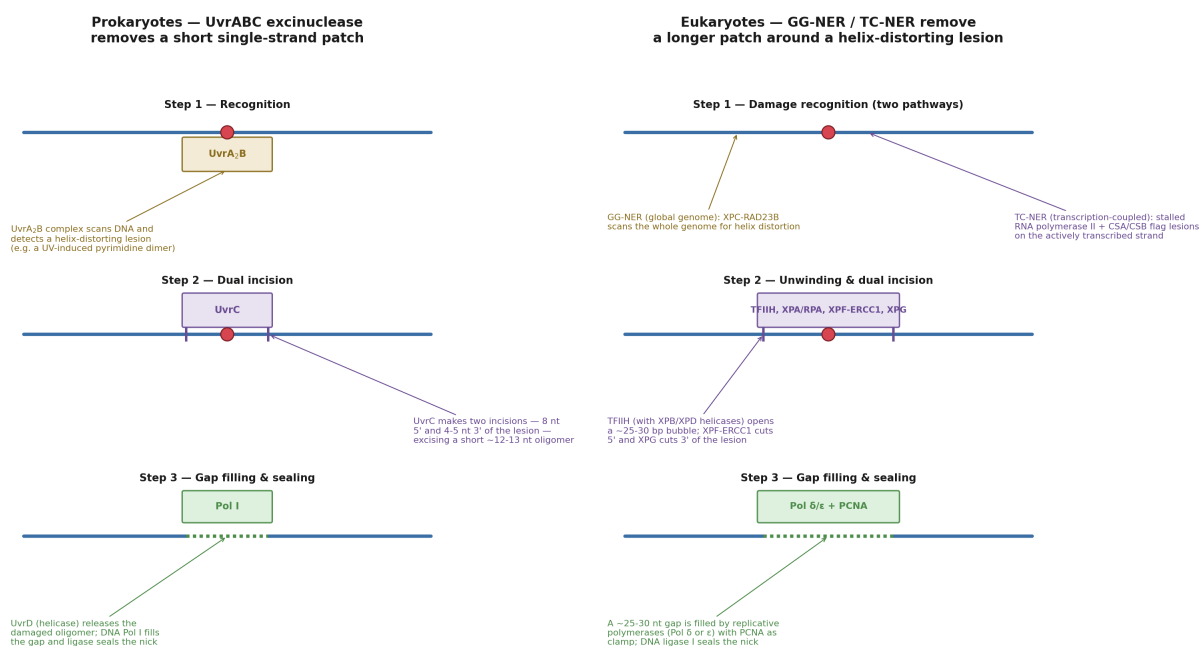


Fig. 2 — Nucleotide excision repair in bacteria (UvrABC excinuclease) versus eukaryotes (GG-NER and TC-NER), each ending in dual incision, gap filling, and ligation.

(a) Prokaryotic NER — the UvrABC System

- A **UvrA₂B** complex scans the genome and detects helix distortion; UvrA then dissociates, leaving UvrB stably bound at the lesion.

- **UvrC** is recruited and makes two incisions flanking the lesion — about 8 nucleotides 5' and 4-5 nucleotides 3' of it — excising a short ~12-13 nucleotide single-stranded fragment (the complex is therefore called an **excinuclease**).
- The helicase **UvrD** releases the excised oligonucleotide; **DNA polymerase I** fills the gap and **DNA ligase** seals the remaining nick.

(b) Eukaryotic NER — Two Sub-Pathways

- **Global genome NER (GG-NER)** continuously surveys the entire genome; the damage sensor **XPC-RAD23B** detects helix distortion anywhere, including in transcriptionally silent DNA.
- **Transcription-coupled NER (TC-NER)** is triggered when RNA polymerase II stalls at a lesion on the template strand of an actively transcribed gene; the factors **CSA** and **CSB** recruit the downstream repair machinery. TC-NER is faster and prioritises repair of the strand being transcribed.
- Both sub-pathways converge on the same downstream steps: the transcription factor **TFIIH** (using its **XPB** and **XPD** helicase subunits) unwinds roughly 25-30 base pairs around the lesion, and **XPA** and **RPA** verify and stabilise the open complex.
- The structure-specific endonucleases **XPF-ERCC1** and **XPG** make incisions 5' and 3' of the lesion respectively, excising an oligonucleotide of about 25-30 nucleotides — considerably longer than the bacterial excision patch.
- Replicative polymerases **Pol δ** or **Pol ε**, together with the sliding clamp **PCNA**, fill the resulting gap, and **DNA ligase I** (or **LigIII** in TC-NER) seals the nick.

Feature	Prokaryotic NER (UvrABC)	Eukaryotic NER
Damage sensor	UvrA ₂ B complex	XPC-RAD23B (GG-NER) or stalled RNA Pol II + CSA/CSB (TC-NER)
Helicase(s)	UvrB, UvrD	TFIIH (XPB, XPD subunits)
Incision enzyme(s)	UvrC (single protein, two active sites)	XPF-ERCC1 (5') and XPG (3') — two separate nucleases
Excised patch size	~12-13 nucleotides	~25-30 nucleotides
Gap-filling polymerase	DNA polymerase I	DNA polymerase δ or ε, with PCNA
Human disease link	—	Xeroderma pigmentosum (XP genes); Cockayne syndrome (CSA/CSB)

5. Mismatch Repair (MMR)

Mismatch repair corrects **base-base mismatches and small insertion/deletion loops** that escape the proofreading activity of replicative DNA polymerases. Because both strands of a freshly replicated mismatch are chemically normal bases, the key challenge for MMR is not recognising damage but distinguishing the **newly synthesised (error-containing) strand** from the original template strand.

(a) Prokaryotic MMR (*E. coli*)

- **MutS** recognises and binds the mismatch; **MutL** then joins to form a complex that activates **MutH**.

- Strand discrimination uses **Dam methylation**: immediately after replication the parental strand is methylated at GATC sequences (by Dam methyltransferase) but the new strand is transiently unmethylated. MutH nicks the unmethylated (new) strand at a nearby hemimethylated GATC site.
- A helicase (UvrD) and exonucleases degrade the new strand from the nick back through the mismatch; **DNA polymerase III** resynthesises it, and **DNA ligase** seals the nick.

(b) Eukaryotic MMR

- Homologues of MutS form heterodimers: **MSH2-MSH6** (MutS α) recognises single base mismatches and single-nucleotide loops, while **MSH2-MSH3** (MutS β) recognises larger insertion/deletion loops.
- Homologues of MutL similarly form **MLH1-PMS2** (MutL α), which nicks the new strand.
- Strand discrimination does not use Dam methylation (eukaryotes lack it); instead the new strand is identified through strand discontinuities, such as the 3' end at an Okazaki fragment or the nick left after ligation, along with an association with **PCNA**, which is loaded in an orientation that marks the newly made strand.
- **Exonuclease 1 (EXO1)** degrades the marked strand past the mismatch, and **DNA polymerase δ** , again aided by PCNA, resynthesises it before ligation.
- Defective MMR in humans causes microsatellite instability and is the molecular basis of **Lynch syndrome (hereditary non-polyposis colorectal cancer, HNPCC)**, most often due to inherited mutations in MLH1 or MSH2.

Feature	Prokaryotic MMR	Eukaryotic MMR
Mismatch recognition	MutS	MSH2-MSH6 (MutS α) or MSH2-MSH3 (MutS β)
Coordinating factor	MutL	MLH1-PMS2 (MutL α)
Strand discrimination	Dam methylation (GATC hemimethylation)	Strand discontinuities / PCNA orientation
Nicking enzyme	MutH	MutL α endonuclease activity (PMS2)
Resynthesis	DNA polymerase III	DNA polymerase δ , with PCNA

6. Double-Strand Break (DSB) Repair

A double-strand break severs both strands of the helix at once, with no intact complementary strand immediately available as a template. Eukaryotic cells use two major, mechanistically distinct pathways to repair DSBs.

Double-Strand Break Repair: Homologous Recombination vs NHEJ

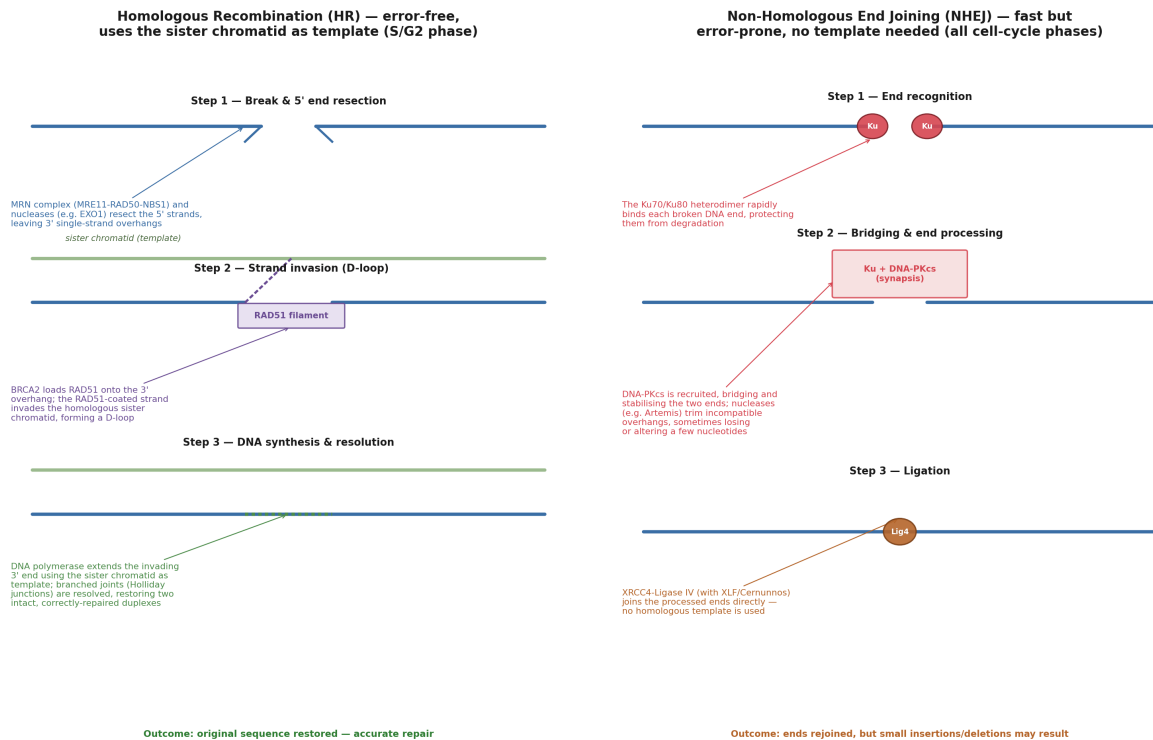


Fig. 3 — Double-strand break repair by homologous recombination (accurate, uses a sister chromatid template) versus non-homologous end joining (fast, template-free, but error-prone).

(a) Homologous Recombination (HR)

- The **MRN complex** (MRE11-RAD50-NBS1) and nucleases such as EXO1 resect the 5' ends at the break, generating long 3' single-stranded overhangs.
- BRCA2** loads the recombinase **RAD51** onto the single-stranded 3' overhang, forming a nucleoprotein filament that searches for and invades a homologous duplex — normally the **sister chromatid** — forming a **D-loop**.
- DNA polymerase extends the invading strand using the sister chromatid as template; the resulting branched joints (Holliday junctions) are resolved to yield two intact, correctly repaired duplexes.
- Because it uses an undamaged homologous template, HR is essentially **error-free**, but it requires a sister chromatid and is therefore restricted to the **S and G2 phases** of the cell cycle.

(b) Non-Homologous End Joining (NHEJ)

- The **Ku70/Ku80** heterodimer rapidly binds each broken DNA end, protecting them from degradation and recruiting the kinase **DNA-PKcs**, which bridges and synapses the two ends together.
- Nucleases (e.g. **Artemis**) and polymerases process incompatible end structures as needed, which can add or remove a small number of nucleotides.
- XRCC4-DNA Ligase IV**, assisted by **XLF (Cernunnos)**, directly ligates the processed ends — no homologous template is required.
- NHEJ is fast and active throughout the cell cycle (dominant in G1), but because end processing can alter the sequence at the junction, it is comparatively **error-prone**, sometimes producing small insertions or deletions.

Feature	Homologous Recombination (HR)	NHEJ
Template required	Yes — sister chromatid	No
Cell-cycle timing	S and G2 phases	All phases (dominant in G1)
Accuracy	High — essentially error-free	Lower — can introduce indels
Key factors	MRN complex, BRCA1/2, RAD51	Ku70/80, DNA-PKcs, XRCC4-Ligase IV
Speed	Slower	Fast
Relevant disease genes	BRCA1/BRCA2 (breast/ovarian cancer risk)	Defects cause radiosensitivity, immunodeficiency (e.g. in V(D)J recombination)

7. Diseases Associated with Defective DNA Repair

Because DNA repair pathways guard against mutation and genome instability, inherited defects in their components are strongly associated with cancer predisposition and, in several cases, additional developmental or neurological features.

Disorder	Pathway affected	Key features
Xeroderma pigmentosum (XP)	Nucleotide excision repair (XPA-XPG genes)	Extreme UV sensitivity, severe sunburn, greatly elevated skin cancer risk
Cockayne syndrome	Transcription-coupled NER (CSA, CSB)	UV sensitivity with developmental delay and premature ageing features, but not increased skin cancer
Lynch syndrome (HNPCC)	Mismatch repair (MLH1, MSH2, and related genes)	Hereditary predisposition to colorectal and other cancers; microsatellite instability
Hereditary breast/ovarian cancer	Homologous recombination (BRCA1, BRCA2)	Elevated lifetime risk of breast and ovarian cancer
Ataxia telangiectasia	DNA damage signalling/checkpoint response (ATM kinase)	Neurodegeneration, immunodeficiency, radiosensitivity, cancer predisposition

8. Quick Revision Summary

Pathway	Key point
Direct reversal	Photolyase (UV dimers, absent in humans); MGMT (alkylation) — single-step, no excision
Base excision repair (BER)	Glycosylase removes damaged base → AP endonuclease nicks → polymerase fills → ligase seals
Nucleotide excision repair (NER)	UvrABC (bacteria, ~12-13 nt patch) vs XPC/TFIIH/XPF-XPG (eukaryotes, ~25-30 nt patch); repairs bulky, helix-distorting lesions
Mismatch repair (MMR)	MutSLH (bacteria, Dam methylation marks new strand) vs MSH/MLH homologues (eukaryotes, strand discontinuity/PCNA marks new strand)
Homologous recombination (HR)	Uses sister chromatid as template via RAD51; accurate; restricted to S/G2
Non-homologous end joining (NHEJ)	Ku70/80 + DNA-PKcs + Ligase IV directly join broken ends; fast but error-prone; active in all phases
XP / HNPCC / BRCA	Classic human diseases of defective NER / MMR / HR respectively, all with elevated cancer risk

Compiled study note — molecular biology reference. Figures are original schematic diagrams prepared for this note and are not to scale.