arginine (see Figure 7.15). On this medium, wild-type and arginine auxotrophs grow, but no other auxotrophs grow. The pattern of the colonies that grow is transferred onto sterile velveteen cloth, and replicas of the colony pattern on the cloth are then made by gently pressing new plates onto the velveteen. If the new plate contains minimal medium, the wild-type colonies can grow but the arginine auxotrophs cannot. By comparing the patterns on the original minimal medium plus arginine master plate with those on the minimal medium replica plate, researchers can readily identify the potential arginine auxotrophs. They can then be picked from the original master plate and cultured for further study.

Conditional Mutants. The products of many genes—DNA polymerases and RNA polymerases, for example—are important for the growth and division of cells, and knocking out the functions of such genes by introducing mutations typically is lethal. The structure and function of such genes can be studied by inducing **conditional mutants**, which reduce the activity of gene products only under certain conditions. A common type of conditional mutation is a temperature-sensitive mutation. In yeast, for instance, many **temperature-sensitive mutants** that grow normally at 23°C but grow very slowly or not at all at 36°C can be isolated. Heat sensitivity typically results from a missense mutation causing a change in the amino acid sequence of a protein so that, at the higher temperature, the protein assumes a nonfunctional shape.

Essentially the same procedures are used to screen for heat-sensitive mutations of microorganisms as for auxotrophic mutations. For example, replica plating can be used to screen for temperature-sensitive mutants when the replica plate is incubated at a higher temperature than the master plate. That is, such mutants grow on the master plate, but not on the replica plate.

Resistance Mutants. In microorganisms such as *E. coli*, yeast, and cells in tissue culture, mutations can be induced for resistance to particular viruses, chemicals, or drugs. For example, in *E. coli*, mutants resistant to phage T1 have been induced (recall the discussion at the beginning of this chapter), and some mutants are resistant to antibiotics such as streptomycin. In yeast, for example, some mutants are resistant to antifungals such as nystatin.

Selecting resistance mutants is straightforward. To isolate azide-resistant mutants of *E. coli*, for example, mutagenized cells are plated on a medium containing azide, and the colonies that grow are resistant to azide. Similarly, antibiotic-resistant *E. coli* mutants can be selected by plating on antibiotic-containing medium.

Keynote

A number of screening procedures have been developed to isolate mutants of interest from a heterogeneous mixture of cells in a mutagenized population of cells.

Repair of DNA Damage

Mutagenesis involves damage to DNA. Especially with high doses of mutagens, the mutational damage can be considerable. What we see as mutations are DNA alterations that are not corrected by various DNA damage repair systems; that is, "mutations = DNA damage − DNA repair." Both prokaryotic and eukaryotic cells have a number of enzyme-based systems that repair DNA damage. If the repair systems cannot correct all the lesions, the result is a mutant cell (or organism) or, if too many mutations remain, death of the cell (or organism).

There are two general categories of repair systems, based on the way they function. *Direct reversal repair systems* correct damaged areas by reversing the damage, whereas *excision repair systems* cut out a damaged area and then repair the gap by new DNA synthesis. Selected repair systems are described in this section.

Direct Reversal Repair of DNA Damage

Mismatch Repair by DNA Polymerase Proofreading. The frequency of base-pair substitution mutations in bacterial genes ranges from 10^{-7} to 10^{-11} errors per generation. However, DNA polymerase inserts incorrect nucleotides at a frequency of 10^{-5}. Most of the difference between the two values is accounted for by the 3′-to-5′ exonuclease proofreading activity of the DNA polymerase in both bacteria and eukaryotes (see Chapter 3, p. 40). When an incorrect nucleotide is inserted, the polymerase often detects the mismatched base pair and corrects the area by "backspacing" to remove the wrong nucleotide and then resuming synthesis in the forward direction.

The *mutator* mutations in *E. coli* illustrate the importance of the 3′-to-5′ exonuclease activity of DNA polymerase for maintaining a low mutation rate. Mutator mutants have a much higher than normal mutation frequency for all genes. These mutants have mutations in genes for proteins whose normal functions are required for accurate DNA replication. For example, the *mutD* mutator gene of *E. coli* encodes the ε (epsilon) subunit of DNA polymerase III, the primary replication enzyme of *E. coli*. The *mutD* mutants are defective in 3′-to-5′ proofreading activity, so that many incorrectly inserted nucleotides are left unrepaired.

Repair of UV-Induced Pyrimidine Dimers. Through **photoreactivation**, or **light repair**, UV light-induced thymine (or other pyrimidine) dimers (see Figure 7.10) are reverted directly to the original form by exposure to near-UV light in the wavelength range from 320 to 370 nm. Photoreactivation occurs when an enzyme called *photolyase* (encoded by the *phr* gene) is activated by a photon of light and splits the dimers apart. Strains with mutations in the *phr* gene are defective in light repair. Photolyase has been found in bacteria and in simple eukaryotes, but not in humans.

Repair of Alkylation Damage. Alkylating agents transfer alkyl groups (usually methyl or ethyl groups) onto the bases. The mutagen MMS methylates the oxygen of carbon-6 in guanine, for example (see Figure 7.12c). In *E. coli*, this alkylation damage is repaired by an enzyme called (O^6-methylguanine methyltransferase, encoded by the *ada* gene. The enzyme removes the methyl group from the guanine, thereby changing the base back to its original form. A similar specific system exists to repair alkylated thymine. Mutations of the genes encoding these repair enzymes result in a much higher rate of spontaneous mutations.

Excision Repair of DNA Damage

Many mutations affect only one of the two strands. In such cases, the DNA damage can be excised and the normal strand used as a template for producing a corrected strand. Depending on the damage, excision may involve a single base or nucleotide, or two or more nucleotides. Each excision repair system involves a mechanism to recognize the specific DNA damage it repairs.

Base Excision Repair. Damaged single bases or nucleotides are most commonly repaired by removing the base or the nucleotide involved and then inserting the correct base or nucleotide. In **base excision repair,** a repair glycosylase enzyme removes the damaged base from the DNA by cleaving the bond between the base and the deoxyribose sugar. Other enzymes then cleave the sugar–phosphate backbone before and after the now baseless sugar, releasing the sugar and leaving a gap in the DNA chain. The gap is filled with the correct nucleotide by a repair DNA polymerase and DNA ligase, with the opposite DNA strand used as the template. Mutations caused by depurination or deamination are examples of damage that may be repaired by base excision repair.

Nucleotide Excision Repair. In 1964, two groups of scientists—R. P. Boyce and P. Howard-Flanders, and R. Setlow and W. Carrier—isolated mutants of *E. coli* that, after UV irradiation, showed a higher than normal rate of induced mutation in the dark. These UV-sensitive mutants were called *uvrA* mutants (*uvr* for "*UV r*epair"). The *uvrA* mutants can repair thymine dimers only with the input of light, meaning they have a normal photoreactivation repair system. However, *uvrA*$^+$ (wild-type) *E. coli* can repair thymine dimers in the dark. Because the normal photoreactive repair system cannot operate in the dark, the investigators hypothesized that there must be another light-independent repair system. They called this system the **dark repair** or **excision repair system,** now typically referred to as the **nucleotide excision repair (NER)** system. The NER system in *E. coli* also corrects other serious damage-induced distortions of the DNA helix.

The NER system involves four proteins—UvrA, UvrB, UvrC, and UvrD—encoded by the genes *uvrA*, *uvrB*, *uvrC*, and *uvrD* (Figure 7.16). A complex of two UvrA proteins and one UvrB protein slides along the DNA (Figure 7.16, step 1). When the complex recognizes a pyrimidine dimer or another serious distortion in the DNA, the UvrA subunits dissociate and a UvrC protein binds to the UvrB protein at the lesion (Figure 7.16, step 2). The resulting UvrBC protein bound to the lesion makes one cut about four nucleotides to the 3′ side in the damaged DNA strand (done by UvrB) and about seven nucleotides to the 5′ side of the lesion (done by UvrC) (Figure 7.16, step 3). UvrB is then released, and UvrD binds to the 5′ cut (Figure 7.16, step 4). UvrD is a helicase that unwinds the region between the cuts, releasing the short single-stranded segment. DNA polymerase I fills in the gap in the 5′-to-3′ direction (Figure 7.16, step 5), and DNA ligase seals the final gap (Figure 7.16, step 6).

Nucleotide excision repair systems have been found in most organisms that have been studied. In yeast and mammalian systems, about 12 genes encode proteins involved in nucleotide excision repair.

Methyl-Directed Mismatch Repair. Despite proofreading by DNA polymerase, a number of mismatched base pairs remain uncorrected after replication has been completed. In the next round of replication, these errors will become fixed as mutations if they are not repaired.

Many mismatched base pairs left after DNA replication can be corrected by **methyl-directed mismatch repair**. This system recognizes mismatched base pairs, excises the incorrect bases, and then carries out repair synthesis. In *E. coli*, the products of three genes—*mutS*, *mutL*, and *mutH*—are involved in the initial stages of mismatch repair (Figure 7.17, p. 149). First, the *mutS*-encoded protein, MutS, binds to the mismatch (Figure 7.17, step 1). Then the repair system determines which base is the correct one (the base on the parental DNA strand) and which is the erroneous one (the base on the new DNA strand). In *E. coli*, the two strands are distinguished by methylation of the A nucleotide in the sequence GATC. This sequence has an axis of symmetry; that is, the same sequence is present 5′-to-3′ on both DNA strands to give

5′-GATC-3′
3′-CTAG-5′.

Both A nucleotides in the sequence usually are methylated. However, after replication, the parental DNA strand has a methylated A in the GATC sequence, whereas the A in the GATC of the *newly replicated* DNA strand is not methylated until a short time after its synthesis. Therefore, the MutS protein bound to the mismatch forms a complex with the *mutL*- and *mutH*-encoded proteins, MutL and MutH, to bring the unmethylated GATC sequence close to the mismatch (Figure 7.17, step 2). The MutH protein then nicks the unmethylated DNA strand at the GATC site, the mismatch is removed by an exonuclease (Figure 7.17, step 3), and the gap is repaired by DNA polymerase III and ligase (Figure 7.17, step 4).

Mismatch repair also takes place in eukaryotes. However, it is unclear how the new DNA strand is distinguished from the parental DNA strand (no methylation is involved). In humans, four genes, respectively named

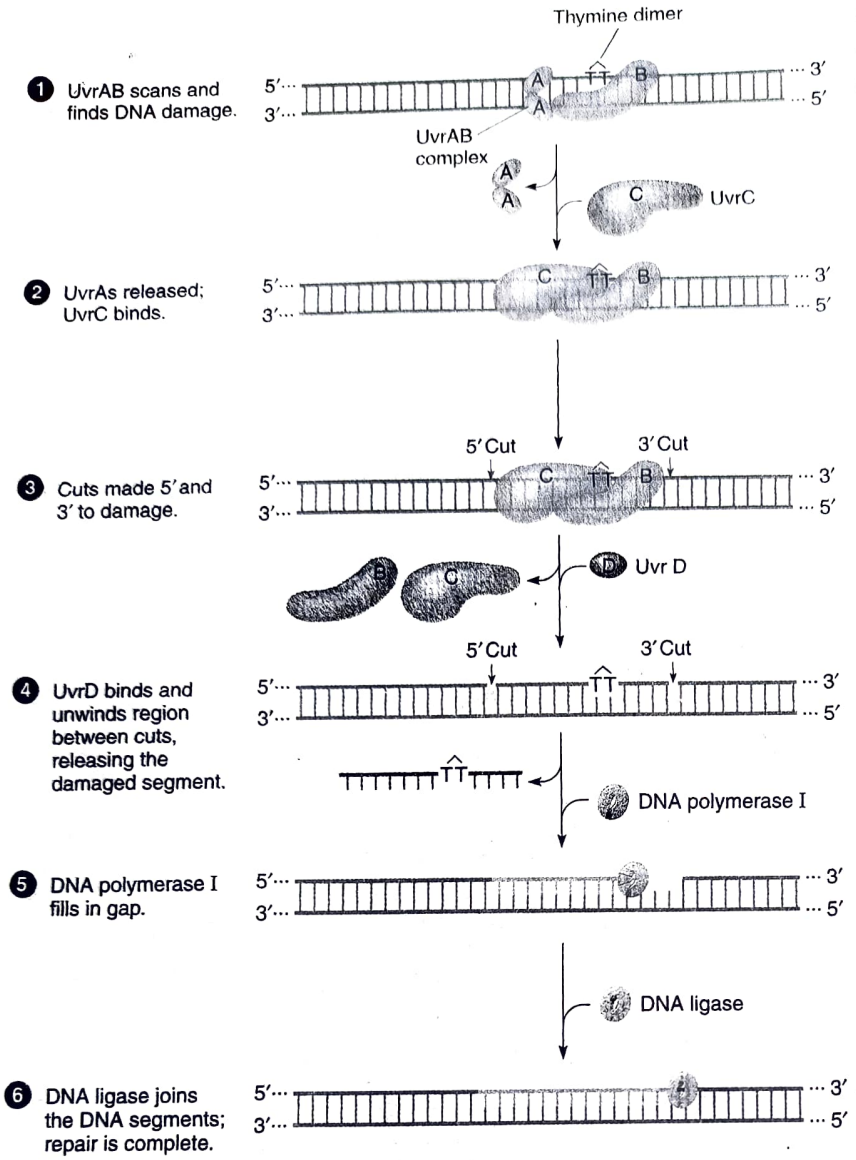


Figure 7.16
Nucleotide excision repair (NER) of pyrimidine dimer and other damage-induced distortions of DNA.

hMSH2, hMLH1, hPMS1, and *hPMS2,* have been identified; *hMSH2* is homologous to *E. coli mutS,* and the other three genes have homologies to *E. coli mutL.* The genes are known as *mutator genes,* because loss of function of such a gene results in an increased accumulation of mutations in the genome. Mutations in any one of the four human mismatch repair genes confer a phenotype of hereditary predisposition to a form of colon cancer called hereditary nonpolyposis colon cancer (HNPCC: OMIM 120435). The role of mutator genes in cancer is described in Chapter 20, p. 594.

Translesion DNA Synthesis and the SOS Response. Lesions that block the replication machinery from proceeding past that point can be lethal if unrepaired. Fortunately, a last-resort process called **translesion DNA synthesis** allows replication to continue past the lesions. The process involves a special class of DNA polymerases that are synthesized only in response to DNA damage. In *E. coli,* such DNA damage activates a complex system called the *SOS response.* (The system is called "SOS" because it is induced as a last-resort, emergency response to mutational damage.) The SOS response allows the cell to survive otherwise lethal events, although often at the expense of generating new mutations.

In *E. coli,* two genes are key to controlling the SOS system: *lexA* and *recA.* The SOS response works as follows: When there is no DNA damage, the *lexA*-encoded protein, LexA, represses the transcription of about 17 genes whose protein products are involved in repairing

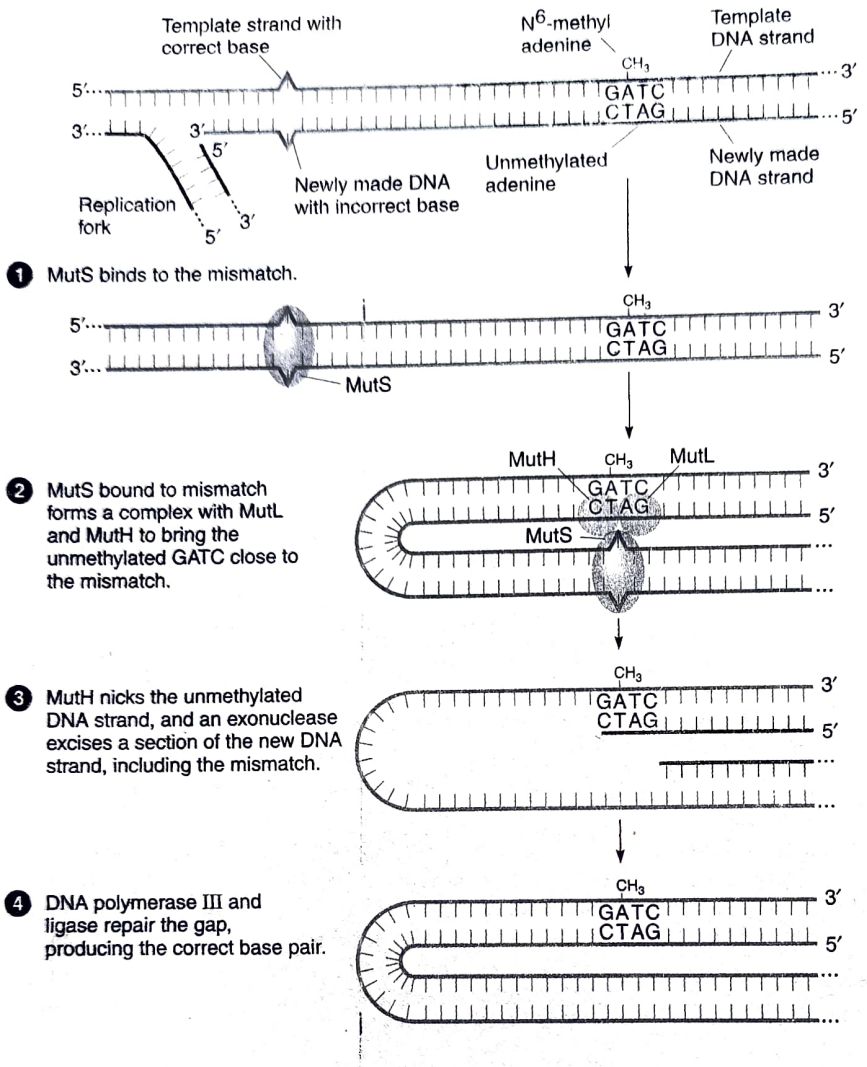


Figure 7.17
Mechanism of mismatch repair. The mismatch correction enzyme recognizes which strand the base mismatch is on by reading the methylation state of a nearby GATC sequence. If the sequence is unmethylated, a segment of that DNA strand containing the mismatch is excised and new DNA is inserted.

and dealing with various kinds of DNA damage. Upon sufficient damage to DNA, the *recA*-encoded protein, RecA, is activated. Activated RecA stimulates the LexA protein to cleave itself, which in turn relieves the repression of the DNA repair genes. As a result, the DNA repair genes are expressed, and DNA repair proceeds. After the DNA damage is dealt with, RecA is inactivated, and newly synthesized LexA protein again represses the DNA repair genes.

Among the gene products made during the SOS response is the DNA polymerase for translesion DNA synthesis. This polymerase continues replication over and past the lesion, although it does so by incorporating one or more nucleotides that are not specified by the template strand into the new DNA across from the lesion. These nucleotides may not match the wild-type template sequence; therefore, the SOS response itself is a mutagenic system because mutations will be introduced into the DNA as a result of its activation. Such mutations are less harmful than the potentially lethal alternative caused by incompletely replicated DNA.

Keynote

Mutations constitute damage to the DNA. Both prokaryotes and eukaryotes have a number of repair systems that deal with different kinds of DNA damage. All the systems use enzymes to make the correction. Without such repair systems, lesions would accumulate and be lethal to the cell or organism. Not all lesions are repaired, and mutations do appear, but at low frequencies. At high doses of mutagens, repair systems are unable to correct all of the damage, and cell death may result.

Human Genetic Diseases Resulting from DNA Replication and Repair Mutations

Some human genetic diseases are attributed to defects in DNA replication or repair; examples are listed in Table 7.1. For instance, *xeroderma pigmentosum*, or XP (OMIM 278700; Figure 7.18) is caused by homozygosity for a recessive mutation in a repair gene. Individuals with this lethal affliction are photosensitive, and portions of their

Base Excision Repair

A separate excision repair system operates to remove altered nucleotides generated by reactive chemicals present in the diet or produced by metabolism. The steps in this repair pathway in eukaryotes, which is called **base excision repair** (**BER**), are shown in Figure 7.27. BER is initiated by a *DNA glycosylase* that recognizes the alteration (step 1, Figure 7.27) and removes the base by cleavage of the glycosidic bond holding the base to the deoxyribose sugar (step 2). A number of different DNA glycosylases have been identified, each more-or-less specific for a particular type of altered base, including uracil (formed by the hydrolytic removal of the amino group of cytosine), 8-oxoguanine (caused by damage from oxygen free radicals, page 34), and 3-methyladenine (produced by transfer of a methyl group from a methyl donor, page 169).

Structural studies of the DNA glycosylase that removes the highly mutagenic 8-oxoguanine (oxoG) indicate that this enzyme diffuses rapidly along the DNA "inspecting" each of the G-C base pairs within the DNA duplex (Figure 7.28, step 1). In step 2, the enzyme has come across an oxoG-C base pair. When this occurs, the enzyme inserts a specific amino acid side chain into the DNA helix, causing the nucleotide to rotate ("flip") 180 degrees out of the DNA helix and into the body of the enzyme (step 2). If the nucleotide does, in fact, contain an oxoG, the base fits into the active site of the enzyme (step 3) and is cleaved from its associated sugar. In contrast, if the extruded nucleotide contains a normal guanine, which only differs in structure by two atoms from oxoG, it is unable to fit into the enzyme's active site (step 4) and it is returned to its appropriate position within the stack of bases. Once an altered purine or pyrimidine is removed by a glycosylase, the "beheaded" deoxyribose phosphate remaining in the site is excised by the combined action of a specialized (AP) endonuclease and a DNA polymerase. AP endonuclease cleaves the DNA backbone (Figure 7.27, step 3) and a phosphodiesterase activity of polymerase β removes the sugar–phosphate remnant that had been attached to the excised base (step 4). Polymerase β then fills the gap by inserting a nucleotide complementary to the undamaged strand (step 5), and the strand is sealed by DNA ligase III (step 6).

The fact that cytosine can be converted to uracil may explain why natural selection favored the use of thymine, rather than uracil, as a base in DNA, even though uracil was presumably present in RNA when it served as genetic material during the early evolution of life (page 186). If uracil had been retained as a DNA base, it would have caused difficulty for repair systems to distinguish between a uracil that "belonged" at a particular site and one that resulted from an alteration of cytosine.

Mismatch Repair

It was noted earlier that cells can remove mismatched bases that are incorporated by the DNA polymerase and escape the enzyme's proofreading exonuclease. This process is called **mismatch repair** (**MMR**). A mismatched base pair causes a distortion in the geometry of the double helix that can be recognized by a repair enzyme. But how does the enzyme "recognize"

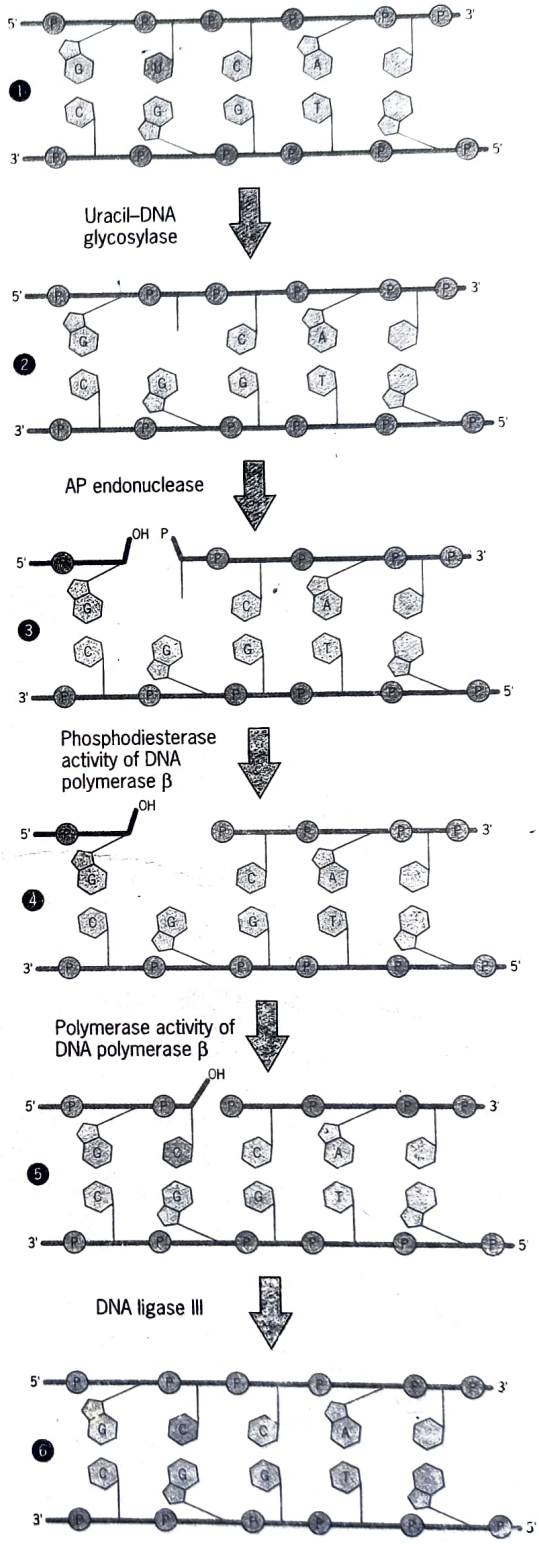


FIGURE 7.27 Base excision repair. The steps are described in the text. Other pathways for BER are known, and BER also has been shown to have distinct transcription-coupled and global repair pathways.

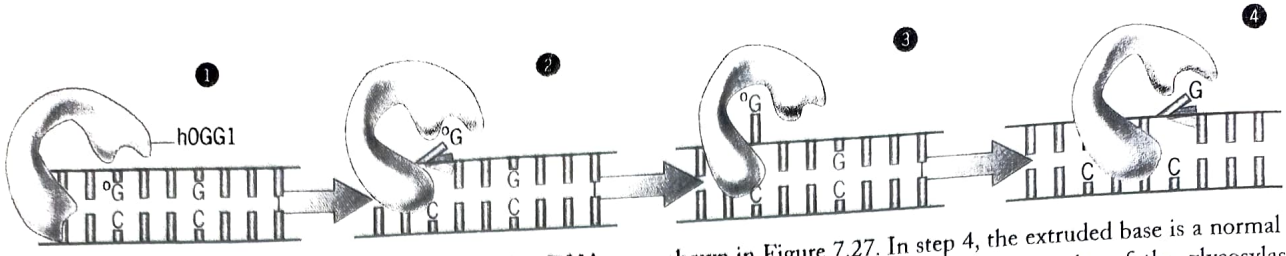
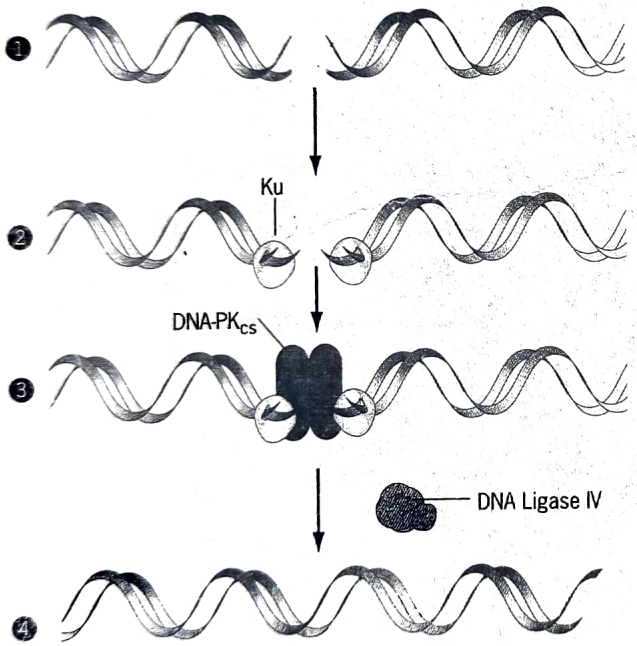


FIGURE 7.28 Detecting damaged bases during BER. In step 1, a DNA glycosylase (named hOGG1) is inspecting a nucleotide that is paired to a cytosine. In step 2, the nucleotide is flipped out of the DNA duplex. In this case, the base is an oxidized version of guanine, 8-oxoguanine, and it is able to fit into the active site of the enzyme (step 3) where it is cleaved from its attached sugar. The subsequent steps in BER were shown in Figure 7.27. In step 4, the extruded base is a normal guanine, which is unable to fit into the active site of the glycosylase and is returned to the base stack. Failure to remove oxoG would have resulted in a G-to-T mutation. (BASED ON S. S. DAVID, WITH PERMISSION FROM NATURE 434:569, 2005; © COPYRIGHT 2005, BY MACMILLAN MAGAZINES LIMITED.)

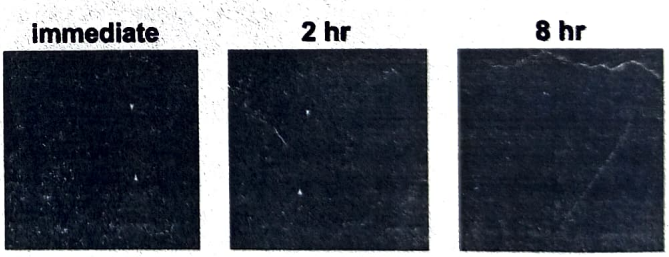
which member of the mismatched pair is the incorrect nucleotide? If it were to remove one of the nucleotides at random, it would make the wrong choice 50 percent of the time, creating a permanent mutation at that site. Thus, for a mismatch to be repaired after the DNA polymerase has moved past a site, it is important that the repair system distinguish the newly synthesized strand, which contains the incorrect nucleotide, from the parental strand, which contains the correct nucleotide. In *E. coli*, the two strands are distinguished by the presence of methylated adenosine residues on the parental strand. DNA methylation does not appear to be utilized by the MMR system in eukaryotes, and the mechanism of identification of the newly synthesized strand remains unclear. Several different MMR pathways have been identified and will not be discussed.

Double-Strand Breakage Repair

X-rays, gamma rays, and particles released by radioactive atoms are all described as *ionizing radiation* because they generate ions as they pass through matter. Millions of gamma rays pass through our bodies every minute. When these forms of radiation collide with a fragile DNA molecule, they often break both strands of the double helix. **Double-strand breaks (DSBs)** can also be caused by certain chemicals, including several (e.g., bleomycin) used in cancer chemotherapy, and free radicals produced by normal cellular metabolism (page 34). DSBs are also introduced during replication of damaged DNA. A single double-strand break can cause serious chromosome abnormalities, which can have grave consequences for the cell. DSBs can be repaired by several alternate pathways. The predominant pathway in mammalian cells is called *nonhomologous end joining (NHEJ)*, in which a complex of proteins bind to the broken ends of the DNA duplex and catalyze a series of reactions that rejoin the broken strands. The major steps that occur during NHEJ are shown in Figure 7.29*a* and described in the accompanying legend. Figure 7.29*b* shows the nuclei of human



(*a*)



(*b*)

FIGURE 7.29 Repairing double-strand breaks (DSBs) by nonhomologous end joining. (*a*) In this simplified model of double-strand break repair, the lesion (step 1) is detected by a heterodimeric, ring-shaped protein called Ku, that binds to the broken ends of the DNA (step 2). The DNA-bound Ku recruits another protein, called DNA-PK$_{cs}$, which is the catalytic subunit of a DNA-dependent protein kinase (step 3). Most of the substrates phosphorylated by this protein kinase have not been identified. These proteins bring the ends of the broken DNA together in such a way that they can be joined by DNA ligase IV to regenerate an intact DNA duplex (step 4). The NHEJ pathway may also involve the activities of nucleases and polymerases (not shown) and is more error prone than is the homologous recombination pathway of DSB repair. (*b*) Time course analysis of Ku localization at sites of DSB formation induced by laser microbeam irradiation at a site indicated by the arrowheads. The NHEJ protein Ku becomes localized at the damage site immediately following irradiation but remains there just briefly as the damage is presumably repaired. Micrographs were taken (1) immediately, (2) 2 hours, and (3) 8 hours after irradiation. (B: FROM JONG-SOO KIM ET AL, COURTESY OF KYOKO YOKOMORI, J. CELL BIOL. 170:344, 2005; BY COPYRIGHT PERMISSION OF THE ROCKEFELLER UNIVERSITY PRESS.)

systems that correct virtually any type of damage to which a DNA molecule is vulnerable. It is estimated that less than one base change in a thousand escapes a cell's repair systems. The existence of these systems provides an excellent example of the molecular mechanisms that maintain cellular homeostasis. The importance of DNA repair can be appreciated by examining the effects on humans that result from DNA repair deficiencies, a subject discussed in the Human Perspective on page 294.

Both prokaryotic and eukaryotic cells possess a variety of proteins that patrol vast stretches of DNA, searching for subtle chemical modifications or distortions of the DNA duplex. In some cases, damage can be repaired directly. Humans, for example, possess enzymes that can directly repair damage from cancer-producing alkylating agents. Most repair systems, however, require that a damaged section of the DNA be *excised*, that is, selectively removed. One of the great virtues of the DNA duplex is that each strand contains the information required for constructing its partner. Consequently, if one or more nucleotides is removed from one strand, the complementary strand can serve as a template for reconstruction of the duplex. The repair of DNA damage in eukaryotic cells is complicated by the relative inaccessibility of DNA within the folded chromatin fibers of the nucleus. As in the case of transcription, DNA repair involves the participation of chromatin-reshaping machines, such as the histone modifying enzymes and nucleosome remodeling complexes discussed on page 255. Although presumably important in DNA repair, the roles of these proteins will not be considered in the following discussion.

Nucleotide Excision Repair

Nucleotide excision repair (NER) operates by a cut-and-patch mechanism that removes a variety of bulky lesions, including pyrimidine dimers and nucleotides to which various chemical groups have become attached. Two distinct NER pathways can be distinguished:

1. A *transcription-coupled pathway* in which the template strands of genes that are being actively transcribed are preferentially repaired. Repair of a template strand is thought to occur as the DNA is being transcribed, and the presence of the lesion may be signaled by a stalled RNA polymerase. This preferential repair pathway ensures that those genes of greatest importance to the cell, which are the genes the cell is actively transcribing, receive the highest priority on the "repair list."
2. A slower, less efficient *global genomic pathway* that corrects DNA strands in the remainder of the genome.

Although recognition of the lesion is probably accomplished by different proteins in the two NER pathways (step 1, Figure 7.26), the steps that occur during repair of the lesion are thought to be very similar, as indicated in steps 2–6 of Figure 7.26. One of the key components of the NER repair machinery is TFIIH, a huge protein that also participates in the initiation of transcription. The discovery of the involvement of TFIIH established a crucial link between transcription and DNA repair, two processes that were previously assumed to be independent of one another (discussed in the Experimental Pathways, which can be accessed on the Web at www.wiley.com/college/karp). Included among the various subunits of TFIIH are two subunits (XPB and XPD) that possess helicase activity; these enzymes separate the two strands of the duplex (step 2, Figure 7.26) in preparation for removal of the lesion. The damaged strand is then cut on both sides of the lesion by a pair of endonucleases (step 3), and the segment of DNA between the incisions is released (step 4). Once excised, the gap is filled by a DNA polymerase (step 5), and the strand is sealed by DNA ligase (step 6).

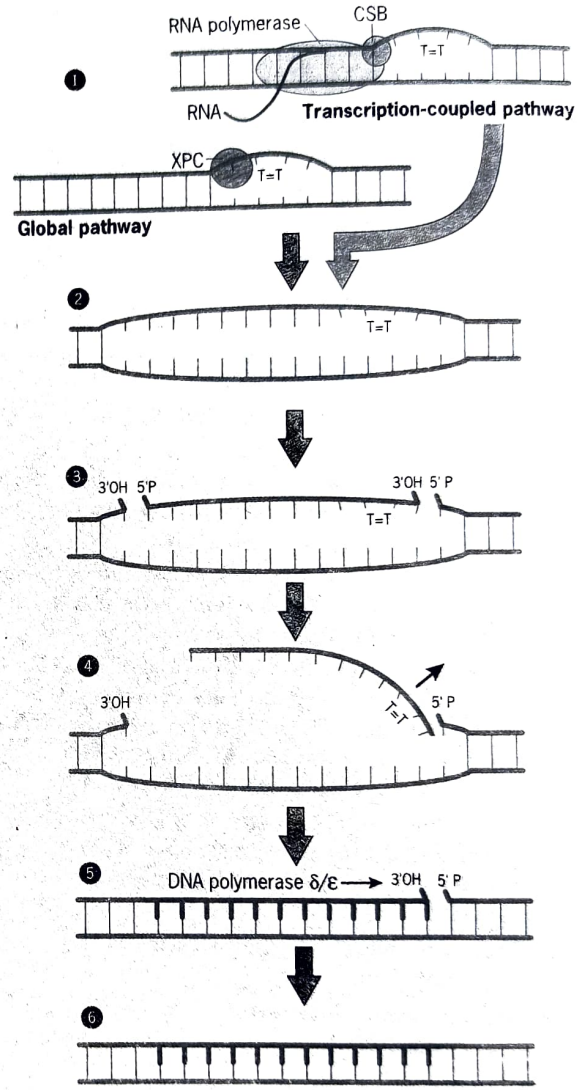


FIGURE 7.26 Nucleotide excision repair. The following steps are depicted in the drawing and discussed in the text: (1) damage recognition in the global pathway is mediated by an XPC-containing protein complex, whereas damage recognition in the transcription-coupled pathway is thought to be mediated by a stalled RNA polymerase in conjunction with a CSB protein; (2) DNA strand separation (by XPB and XPD proteins, two helicase subunits of TFIIH); (3) incision (by XPG on the 3′ side and the XPF–ERCC1 complex on the 5′ side); (4) excision, (5) DNA repair synthesis (by DNA polymerase δ and/or ε); and (6) ligation (by DNA ligase I).

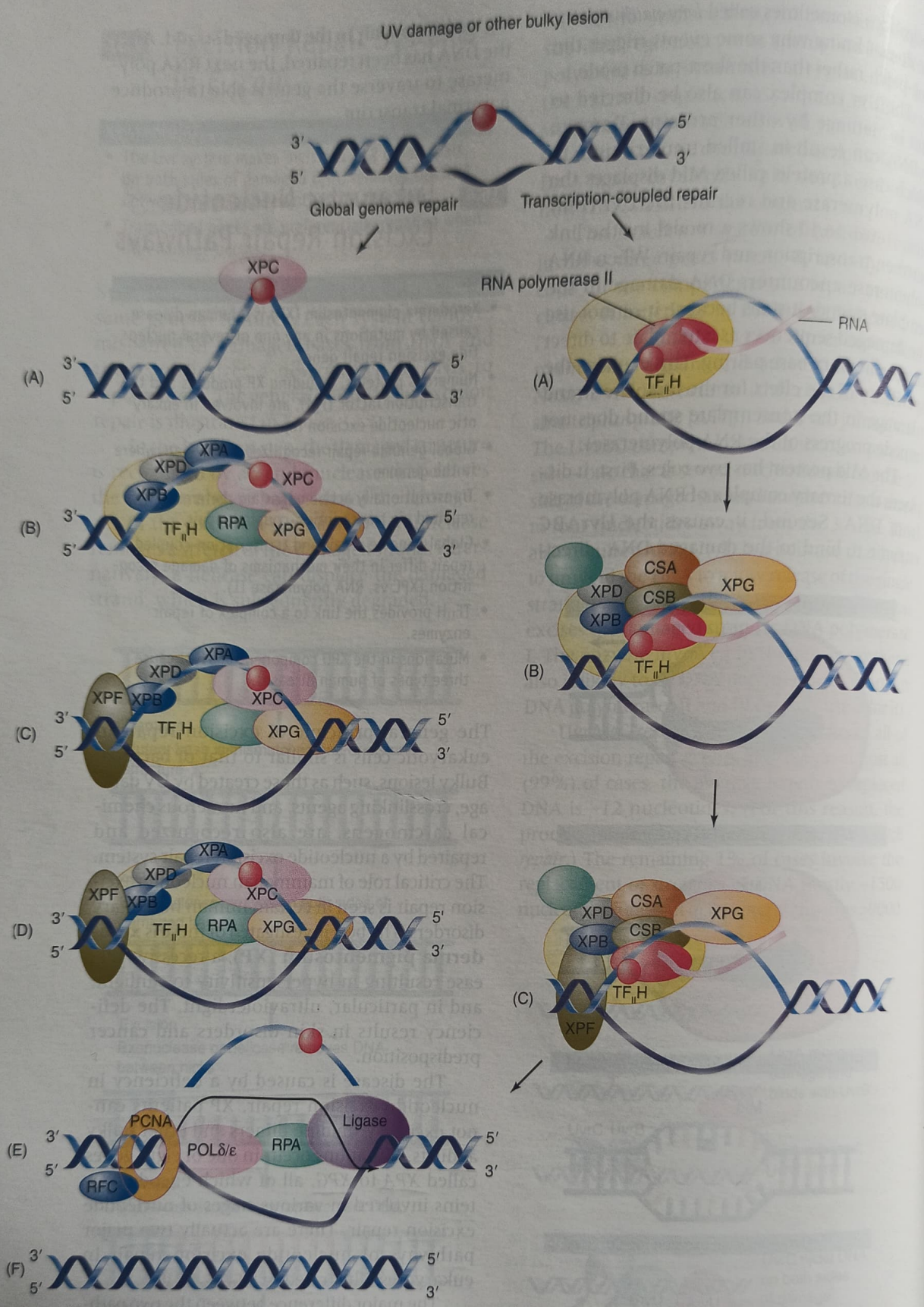
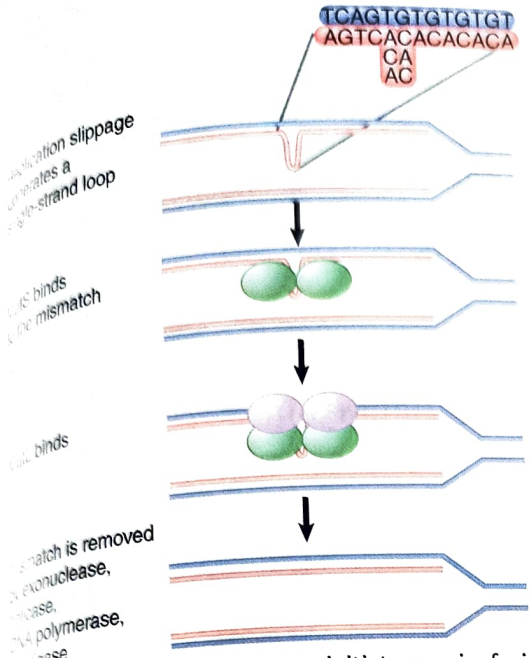


FIGURE 16.11 Nucleotide excision repair occurs via two major pathways: global genome repair, in which XPC recognizes damage anywhere in the genome, and transcription-coupled repair, in which the transcribed strand of active genes is preferentially repaired, and the damage is recognized by an elongating RNA polymerase. Adapted from E. C. Friedberg, et al., *Nature Rev. Cancer* 1 (2001): 22–23.



RE 16.19 The MutS/MutL system initiates repair of mis-ches produced by replication slippage.

bers of repeats) of microsatellite sequences change apidly in the tumor cells due to the loss of the mismatch repair system to correct replication slip-page in these sequences. This instability can be used diagnostically to identify HNPCC.

16.8 Recombination-Repair Systems in *E. coli*

Key concepts

- The *rec* genes of *E. coli* code for the principal recombination-repair system.
- The recombination-repair system functions when replication leaves a gap in a newly synthesized strand that is opposite a damaged sequence.
- The single strand of another duplex is used to replace the gap.
- The damaged sequence is then removed and resynthesized.

Recombination-repair systems use activities that overlap with those involved in genetic recombination. They are also sometimes called "post-replication repair" because they function after replication. Such systems are effective in dealing with the defects produced in daughter duplexes by replication of a template that contains damaged bases. An example is illustrated in FIGURE 16.20.

Consider a structural distortion, such as a pyrimidine dimer, on one strand of a double helix. When the DNA is replicated, the dimer prevents the damaged site from

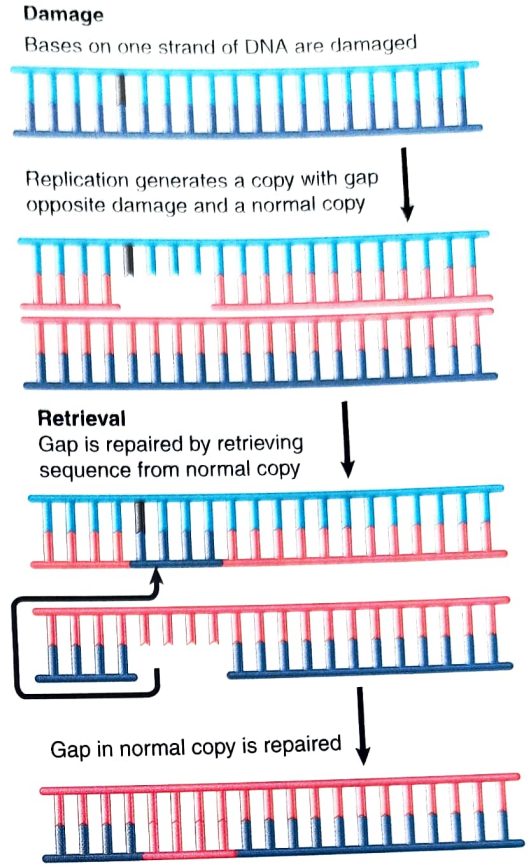


FIGURE 16.20 An *E. coli* retrieval system uses a normal strand of DNA to replace the gap left in a newly synthesized strand opposite a site of unrepaired damage.

acting as a template. Replication is forced to skip past it.

DNA polymerase probably proceeds up to or close to the pyrimidine dimer. The polymerase then ceases synthesis of the corresponding daughter strand. Replication restarts some distance farther along. This replication may be performed by translesion polymerases, which can replace the main DNA polymerase at such sites of unrepaired damage (see *Section 16.6, Error-Prone Repair*). A substantial gap is left in the newly synthesized strand.

The resulting daughter duplexes are different in nature. One has the parental strand containing the damaged adduct, which faces a newly synthesized strand with a lengthy gap. The other duplicate has the undamaged parental strand, which has been copied into a normal complementary strand. The retrieval system takes advantage of the normal daughter.

The gap opposite the damaged site in the first duplex is filled by utilizing the homologous single strand of DNA from the normal duplex. Following this **single-strand exchange**, the recipient duplex has a parental (damaged)

strand facing a wild-type strand. The donor duplex has a normal parental strand facing a gap; the gap can be filled by repair synthesis in the usual way, generating a normal duplex. Thus the damage is confined to the original distortion (although the same recombination-repair events must be repeated after every replication cycle unless and until the damage is removed by an excision repair system).

The principal pathway for recombination-repair in *E. coli* is identified by the *rec* genes (see Figures 15.17 and 15.18). In *E. coli* deficient in excision repair, mutation of the *recA* gene essentially abolishes all the remaining repair and recovery facilities. Attempts to replicate DNA in *uvr*$^-$ *recA*$^-$ cells produce fragments of DNA whose size corresponds with the expected distance between thymine dimers. This result implies that the dimers provide a lethal obstacle to replication in the absence of RecA function. It explains why the double mutant cannot tolerate >1 to 2 dimers in its genome (compared with the ability of a wild-type bacterium to handle as many as 50).

One *rec* pathway involves the *recBC* genes and is well characterized; the other involves *recF* and is not so well defined. They fulfill different functions *in vivo*. The RecBC pathway is involved in restarting stalled replication forks (see *Section 16.9, Recombination Is an Important Mechanism to Recover from Replication Errors*). The RecF pathway is involved in repairing the gaps in a daughter strand that are left after replicating past a pyrimidine dimer.

The RecBC and RecF pathways both function prior to the action of RecA (although in different ways). They lead to the association of RecA with a single-stranded DNA. The ability of RecA to exchange single strands allows it to perform the retrieval step in Figure 16.20. Nuclease and polymerase activities then complete the repair action.

The RecF pathway contains a group of three genes: *recF*, *recO*, and *recR*. The proteins form two types of complex, RecOR and RecOF. They promote the formation of RecA filaments on single-stranded DNA. One of their functions is to make it possible for the filaments to assemble in spite of the presence of single strand binding protein (SSB), which is inhibitory to RecA assembly.

The designations of repair and recombination genes are based on the phenotypes of the mutants, but sometimes a mutation isolated in one set of conditions and named as a *uvr* gene turns out to have been isolated in another set of conditions as a *rec* gene. This illustrates the point that the *uvr* and *rec* pathways are not independent, because *uvr* mutants show reduced efficiency in recombination-repair. We must expect to find a network of nuclease, polymerase, and other activities, which constitute repair systems that are partially overlapping (or in which an enzyme usually used to provide some function can be substituted by another from a different pathway).

16.9 Recombination Is an Important Mechanism to Recover from Replication Errors

Key concepts

- A replication fork may stall when it encounters a damaged site or a nick in DNA.
- A stalled fork may reverse by pairing between the two newly synthesized strands.
- A stalled fork may restart after repairing the damage and use a helicase to move the fork forward.
- The structure of the stalled fork is the same as a Holliday junction and may be converted to a duplex and DSB by resolvases.

In many cases, rather than skipping a DNA lesion, DNA polymerase instead stops replicating when it encounters DNA damage. FIGURE 16.21 shows one possible outcome when a replication fork stalls. The fork stops moving forward when it encounters the damage. The replication apparatus disassembles, at least partially. This allows branch migration to occur, when the fork effectively moves backward, and the new daughter strands pair to form a duplex structure. After the damage has been repaired, a helicase rolls the fork forward to restore its structure. Then the replication apparatus can reassemble, and replication is restarted (see *Section 14.16, The Primosome Is Needed to Restart Replication*).

The pathway for handling a stalled replication fork requires repair enzymes, and restarting stalled replication forks is thought to be a major role of the recombination-repair systems. In *E. coli*, the RecA and RecBC systems have an important role in this reaction (in fact, this may be their major function in the bacterium). One possible pathway is for RecA to stabilize single-stranded DNA by binding to it at the stalled replication fork and possibly acting as the sensor that detects the stalling event. RecBC is involved in excision repair of the damage. After the damage has been repaired, replication can resume.

Another pathway may use recombination-repair—possibly the strand-exchange reactions

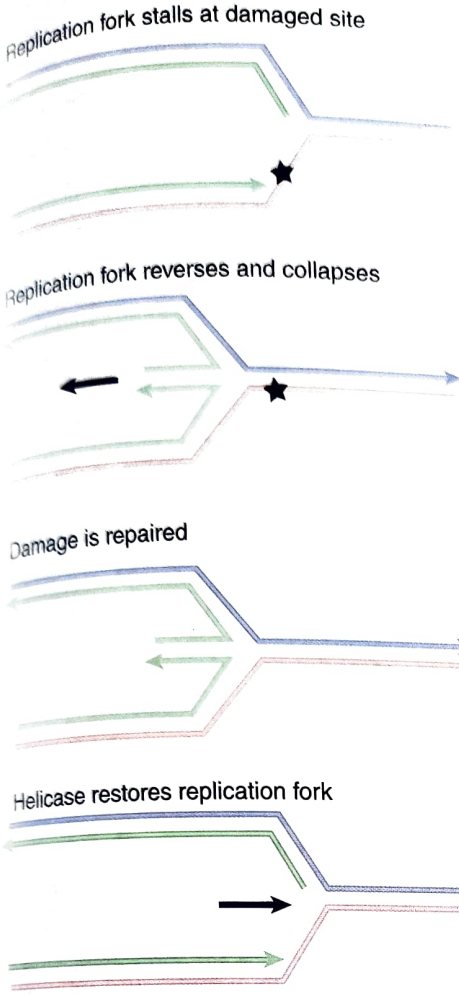


FIGURE 16.21 A replication fork stalls when it reaches a damaged site in DNA. Reversing the fork allows the two daughter strands to pair. After the damage has been repaired, the fork is restored by forward-branch migration catalyzed by a helicase. Arrowheads indicate 3′ ends.

of RecA. FIGURE 16.22 shows that the structure of the stalled fork is essentially the same as a Holliday junction created by recombination between two duplex DNAs (see *Section 15.3, Double-Strand Breaks Initiate Recombination*). This makes it a target for resolvases. A double-strand break is generated if a resolvase cleaves either pair of complementary strands. In addition, if the damage is in fact a nick, another double-strand break is created at this site.

Stalled replication forks can be rescued by recombination-repair. We don't know the exact sequence of events, but one possible scenario is outlined in FIGURE 16.23. The principle is that a recombination event occurs on either side of the damaged site, allowing an undamaged single strand to pair with the damaged strand. This allows the replication fork to be reconstructed so that replication can continue, effectively bypassing the damaged site.

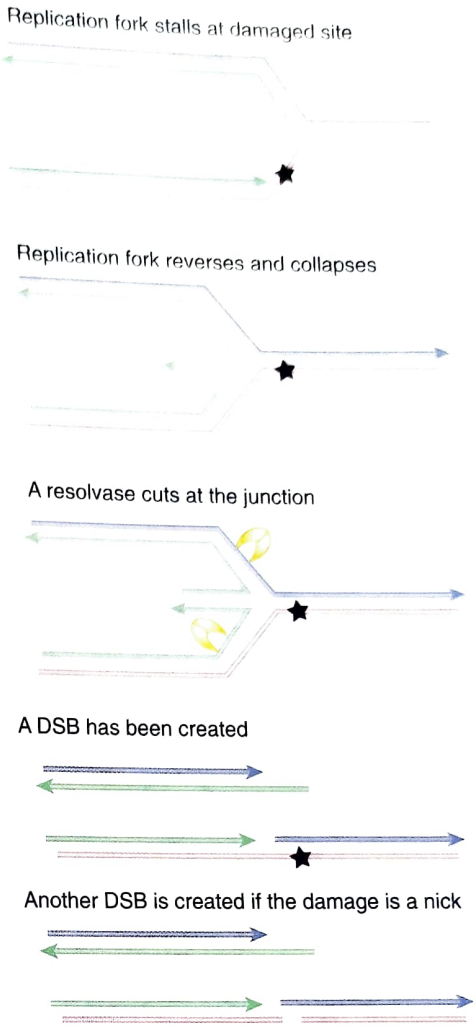


FIGURE 16.22 The structure of a stalled replication fork resembles a Holliday junction and can be resolved in the same way by resolvases. The results depend on whether the site of damage contains a nick. Result 1 shows that a double-strand break is generated by cutting a pair of strands at the junction. Result 2 shows a second DSB is generated at the site of damage if it contains a nick. Arrowheads indicate 3′ ends.

16.10 Recombination-Repair of Double-Strand Breaks in Eukaryotes

Key concepts

- The yeast *RAD* mutations, identified by radiation-sensitive phenotypes, are in genes that code for repair systems.
- The *RAD52* group of genes is required for recombination repair.
- The MRX (yeast) or MRN (mammals) complex is required to form a single-stranded region at each DNA end.
- The RecA homolog Rad51 forms a nucleoprotein filament on the single-stranded regions, assisted by Rad52 and Rad55/57.
- Rad54 and Rdh54/Rad54B are involved in homology search and strand invasion.

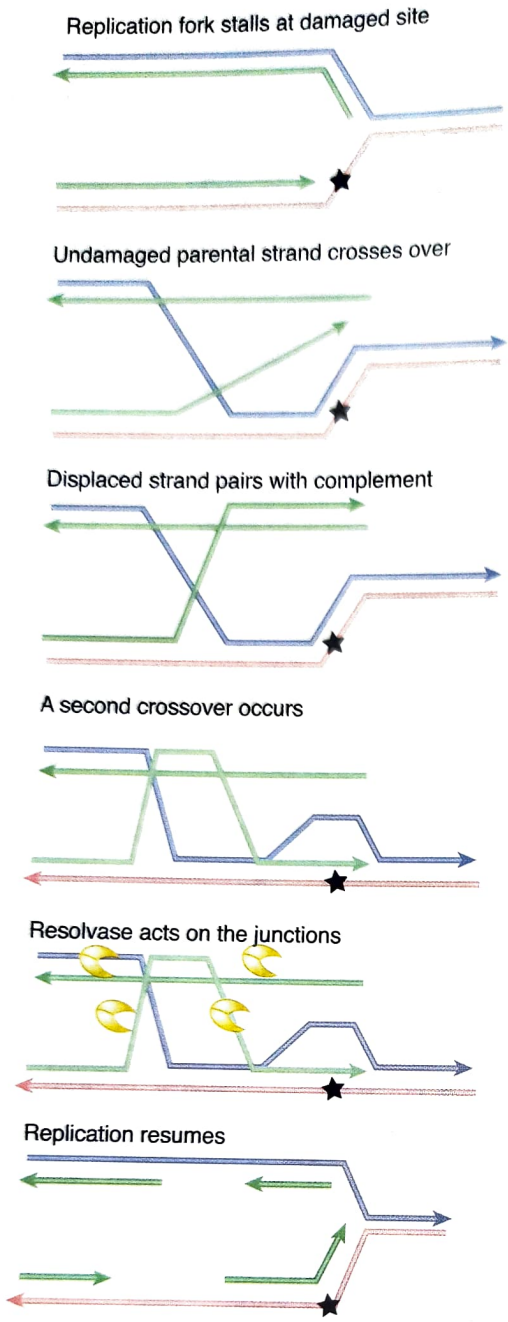


FIGURE 16.23 When a replication fork stalls, recombination-repair can place an undamaged strand opposite the damaged site. This allows replication to continue.

When a replication fork encounters a lesion in a single stand, it can result in the formation of a double-strand break (DSB). DSBs are one of the most severe types of DNA damage that can occur, particularly in eukaryotes. If a DSB on a linear chromosome is not repaired, the portion of the chromosome lacking a centromere will not be segregated at the next cell division. In addition to their occurrence during replication, DSBs can be generated in a number of other ways, including ionizing radiation, oxygen radicals generated by cellular metabolism, or action of endonucleases. The preferred [illegible] for repairing DSBs is to use recom[illegible]ism repair, as this ensures that no critic[illegible]om- information is lost due to sequence [illegible]tic breakpoint. [illegible]the

Several of the genes required for [illegible] nation repair in eukaryotes have alre[illegible]bi- discussed in the context of homologou[illegible]een bination (see *Section 15.15, Eukaryot[illegible]m- Involved in Homologous Recombination*). Many eukaryotic repair genes are named *RAD* genes; they were initially characterized genetically in yeast by virtue of their sensitivity to radiation. There are three general groups of repair genes in the yeast *S. cerevisiae*, identified by the *RAD3* group (involved in excision repair), the *RAD6* group (required for postreplication repair), and the *RAD52* group (concerned with recombination-like mechanisms). Homologs of these genes are present in higher eukaryotes as well.

The *RAD52* group plays essential roles in homologous recombination, and includes a large number of genes such as *RAD50*, *RAD51*, *RAD54*, *RAD55*, *RAD57*, and *RAD59*. These Rad proteins are all required at different stages of repair of a double-strand break. As occurs during meiotic recombination, the Mre11/Rad50/Xbs1 (MRX) complex (MRN in mammals) binds to the free DNA ends, and may tether the ends together, as shown in FIGURE 16.24. In concert with exonucleases and helicases, the MRX complex is required to resect the ends of the double-strand break to generate single-stranded tails with 3′–OH overhangs. This single-stranded DNA serves to activate a DNA damage checkpoint, stopping cell division until the damage can be repaired. The RecA homolog Rad51 binds to the single-stranded DNA to form a nucleoprotein filament, which is used for strand invasion of a homologous sequence. Rad52 and the Rad55/57 complex are required to form a stable Rad51 filament, and Rad54 and its homolog Rdh54 (Rad54B in mammals) assist in the search for homologous donor DNA and subsequent strand invasion. Rad54 and Rdh54 are members of the SWI2/SNF2 superfamily of chromatin remodeling enzymes (see *Section 28.7, Chromatin Remodeling Is an Active Process*), and may be necessary for reconfiguring chromatin structure at both the damage site and at the donor DNA. Following repair synthesis, the resulting structure (which resembles a Holliday junction) is resolved (see Figure 15.4 in *Section 15.3, Double-Strand Breaks Initiate Recombination*, for an illustration of these events).

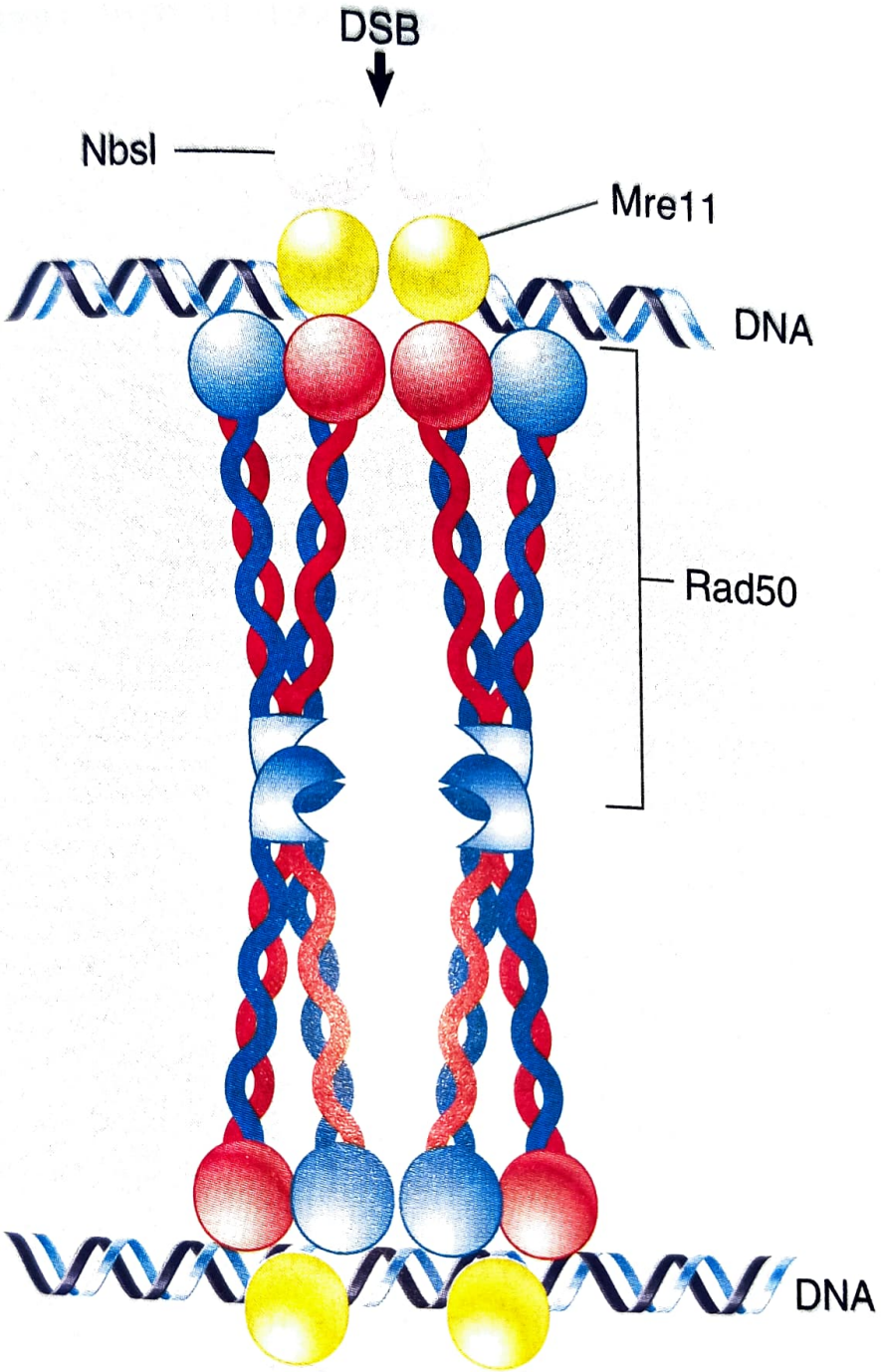
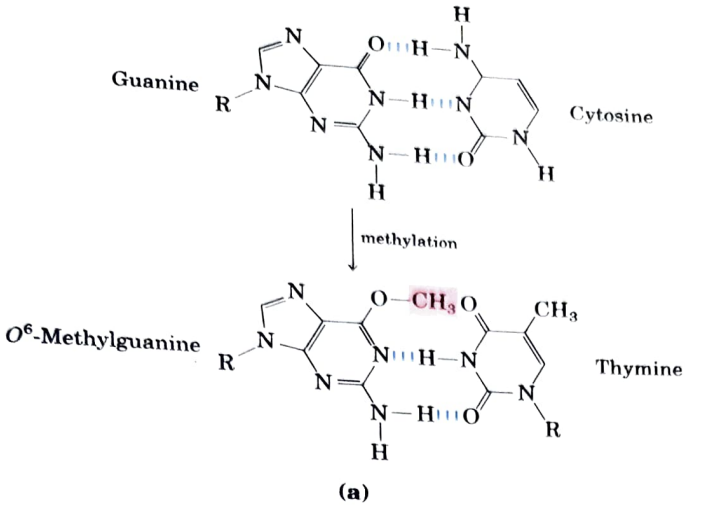
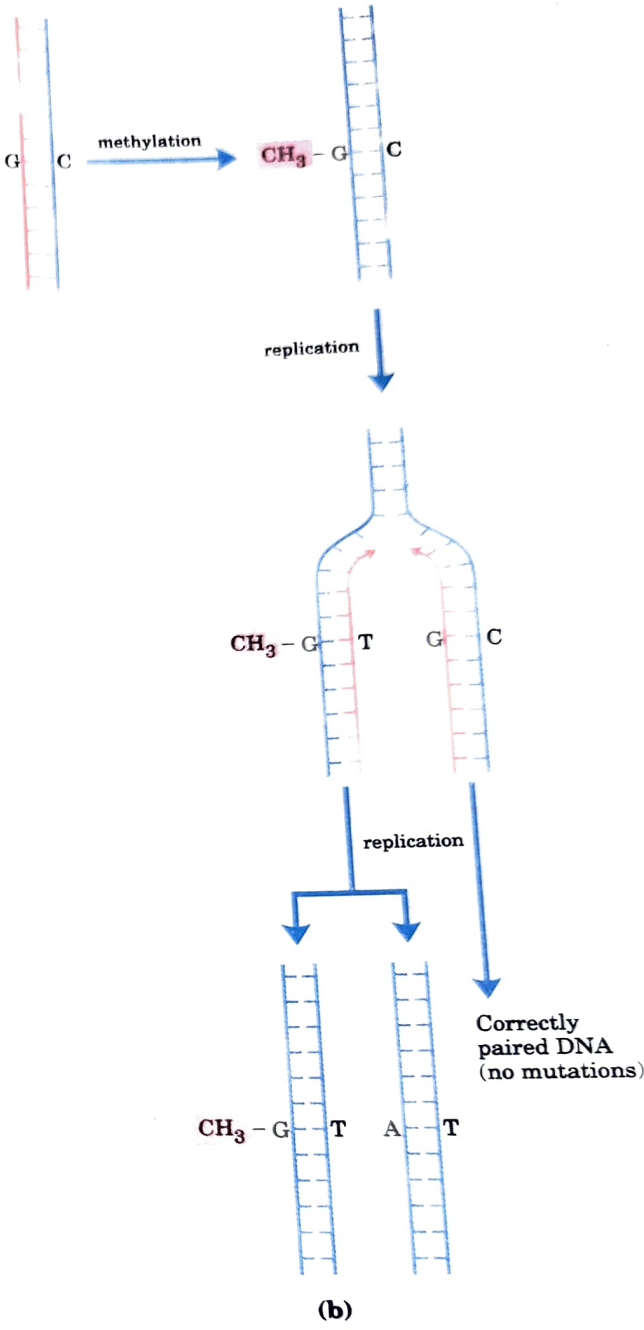


FIGURE 16.24 The MRN complex, required for 5′ end resection, also serves as a DNA bridge to prevent broken ends from separating. The "head" region of Rad50, bound to Mre11, binds DNA, while the extensive coiled coil region of Rad50 ends with a "zinc hook" that mediates interaction with another MRN complex. The precise position of Nbs1 within the complex is unknown but it interacts directly with Mre11.



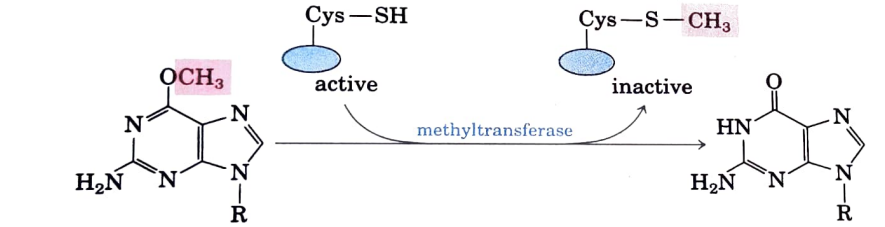
(a)



(b)

Figure 24–23 An example of how DNA damage results in mutations. The methylation product O^6-methylguanine pairs with thymine rather than cytosine **(a)**. If not repaired, this leads to a G≡C to A=T mutation after replication **(b)**.

Direct repair of O^6-methylguanine is carried out by O^6-methylguanine-DNA methyltransferase, which catalyzes the transfer of the methyl group of O^6-methylguanine to a specific Cys residue on the same protein. This methyltransferase is not strictly an enzyme, because a single methyl transfer event inactivates the protein. The consumption of an entire protein molecule to correct a single damaged base is another vivid illustration of the central importance of maintaining the integrity of cellular DNA.



O^6-Methylguanine nucleotide

Guanine nucleotide

When DNA Damage Is Extensive, Repair Becomes Error-Prone

Up to this point, our discussion has focused on the accurate repair of the relatively rare DNA lesions that occur daily in any cell. However, in *E. coli,* when the chromosome is subjected to heavy damage through exposure to UV light or a DNA-damaging reagent, DNA repair becomes significantly less accurate and a high mutation rate is observed. This is referred to as **error-prone repair,** a distinct and unusual pathway.

Given the energetic investment made to maintain the structural and sequence integrity of cellular DNA, it may seem incongruous that mechanisms exist to *increase* mutation rates. As is often the case in biochemistry, however, an examination of an apparent exception to a

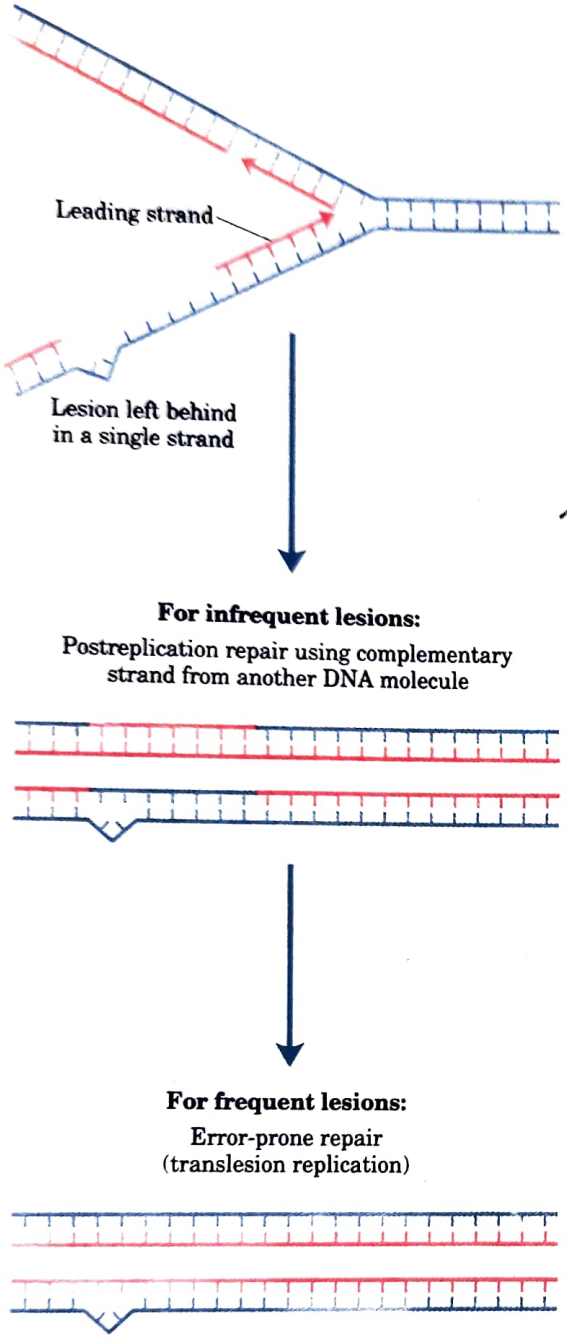


Figure 24–24 DNA damage and its effect on DNA replication. If an unrepaired lesion is encountered at the replication fork, replication generally stops and is resumed farther along the chromosome. The lesion is left behind in an unreplicated, single-stranded segment of the DNA. There are two possible avenues for repair. The recombinational pathway, called postreplication repair, is described in Fig. 24–34. When lesions are so numerous that normal replication is inhibited, a second repair mechanism operates. The specialized system uses DNA polymerase II and can replicate over many types of lesions. This is called error-prone repair because mutations often result.

general rule can throw light on the rule itself. In this instance we must examine the complex interrelationships among repair, replication, and recombination. In *E. coli*, normal DNA replication with DNA polymerase III cannot proceed past many types of DNA lesions. Under normal circumstances, most lesions are repaired before the replication complex arrives. The occasional unrepaired lesion blocks replication, but replication begins again beyond the site of the lesion (Fig. 24–24) and the lesion itself can eventually be repaired with the aid of recombination processes (postreplication repair) described later in this chapter. Higher levels of DNA damage, however, effectively bring normal DNA replication to a halt and trigger a stress response in the cell involving a regulated increase (induction) in the levels of a number of proteins. This is called, appropriately enough, the **SOS response.** Some of the proteins induced, such as the UvrA and UvrB proteins, have roles in DNA repair (Table 24–6). A number of the induced proteins, however, are part of a specialized replication system that can replicate past the DNA lesions that block DNA polymerase III. Because proper base pairing is often impossible at the site of a lesion, this translesion replication is error-prone. The resulting increase in mutagenesis does not contradict the general principle that replication accuracy is important—the resulting mutations actually kill many cells. This is the biological price that is paid, however, to overcome the general barrier to replication and permit at least a few mutant cells to survive.

Table 24–6 Genes induced as part of the SOS response

Gene name	Role in DNA repair
Genes of known function	
*pol*B (*din*A)	Encodes polymerization subunit of DNA polymerase II, required for error-prone repair
*uvr*A, *uvr*B, *uvr*C	Encode ABC excinuclease
*umu*C, *umu*D	Encode proteins required for error-prone repair
*sul*A	Encodes protein that inhibits cell division, possibly to allow time for DNA repair
*rec*A	Encodes RecA protein required for error-prone repair and recombinational repair
Genes involved in DNA metabolism, but role in DNA repair unknown	
ssb	Encodes single-strand binding protein (SSB)
*uvr*D	Encodes DNA helicase II (DNA-unwinding protein)
*him*A	Encodes subunit of integration host factor, involved in site-specific recombination, replication, transposition, regulation of expression of a number of genes
*rec*N	Involved in recombinational repair
Genes of unknown function	
*din*B	
*din*D	
*din*F	

Translesion replication brings us back to a discussion of *E. coli* DNA polymerase II. This polymerase is induced as part of the SOS response and, unlike DNA polymerase III, it is capable of replication past lesions such as AP sites. This enzyme has some of the same subunits as DNA polymerase III, and at least some of these protein subunits are synthesized in larger amounts as part of the SOS response. In addition to this unusual polymerase activity, error-prone repair requires the activities of the UmuC, UmuD, and RecA proteins, although their precise molecular functions in mutagenesis are not understood.

The RecA protein merits some additional discussion because it has several distinct functions (besides mutagenesis) in the bacterial cell. RecA protein is involved in recombination and in the regulation of the SOS response, and in these cases its molecular function is well characterized. The regulation of the SOS response is described in Chapter 27. We now turn to a discussion of genetic recombination.

DNA Recombination

The rearrangement of genetic information in and among DNA molecules encompasses a variety of processes that are collectively placed under the heading of genetic recombination. An understanding of how DNA rearrangements occur is finding practical application as scientists explore new methods for altering the genomes of a variety of organisms (Chapter 28).

Genetic recombination events fall into at least three general classes. **Homologous genetic recombination** involves genetic exchanges between any two DNA molecules (or segments of the same molecule) that share an extended region with homologous sequences. The actual sequence of bases in the DNA is irrelevant as long as the two DNAs are similar. **Site-specific recombination** [illegible] defined DNA sequence,

Once a Holliday intermediate has been formed, enzymes involved in completing recombination include topoisomerases, a resolvase, other nucleases, DNA polymerase I or III, and DNA ligase. The RuvC protein (M_r 20,000) of *E. coli* cleaves Holliday intermediates in the manner depicted in Figure 24–28a. Many details of the reactions carried out by the recombination enzymes and the coordination of these reactions in the cell are not yet understood.

Homologous Recombination Is an Important Pathway for DNA Repair

Recombination provides an avenue for accurate DNA repair when the necessary sequence information is not available from a strand paired with the damaged strand (see Fig. 24–27). To illustrate the role of recombination in DNA repair, we will examine the fate of lesions encountered during normal replication and left behind unreplicated in single-stranded DNA (see Figs. 24–24, 24–27). Repair of these lesions is called postreplication repair, and in *E. coli* this process requires RecA protein.

A plausible pathway for postreplication repair is presented in Figure 24–34. A lesion in an unpaired DNA strand cannot be excised, because this would leave breaks in both DNA strands, an outcome that could be lethal to the cell. To prevent chromosomal breakage and allow for repair, the region containing the lesion must acquire a complementary strand. The recombination pathway makes use of the homologous DNA on the other leg of the replication fork. A RecA protein–mediated strand-exchange reaction transfers an undamaged complementary strand from the homologous DNA, converting the region containing the lesion into heteroduplex DNA. A notable property of RecA protein–mediated DNA strand exchange is that it proceeds efficiently past most DNA lesions with the aid of energy supplied by ATP hydrolysis. Once the lesion is made part of a duplex, the damage can be readily repaired. The repair of lesions of this type is clearly a major function of the homologous recombination system of every cell.

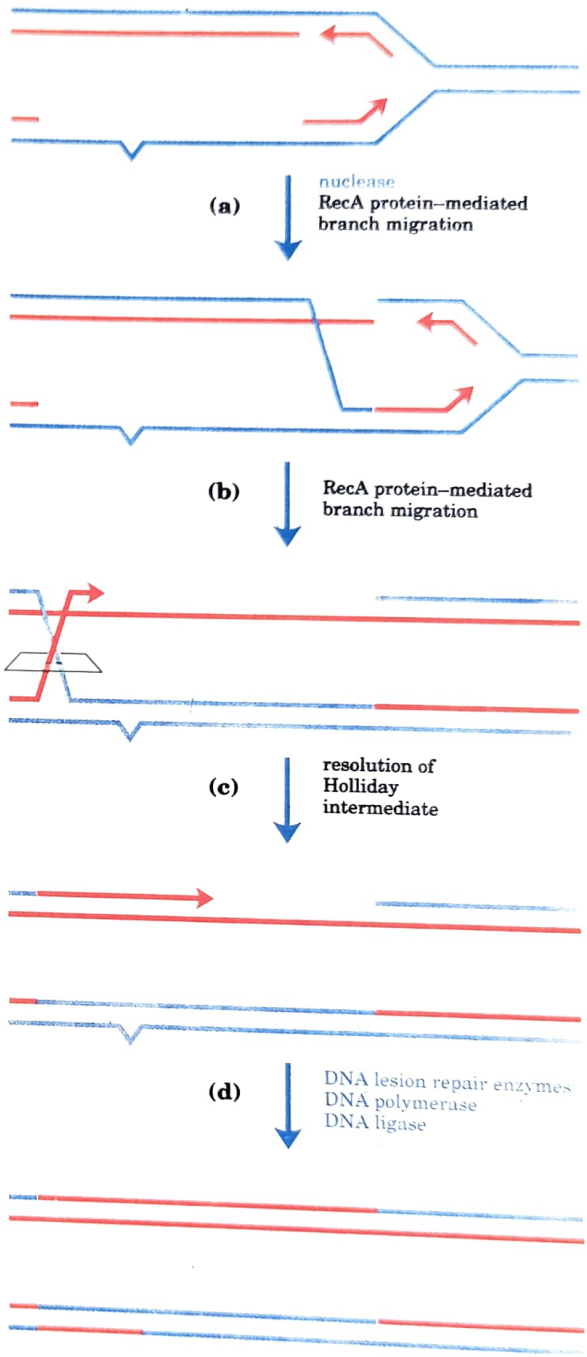


Figure 24–34 Model for the role of RecA protein in postreplication repair. **(a)** A region of single-stranded DNA containing a lesion remains unreplicated. **(b)** A RecA protein–mediated strand exchange transfers a complementary strand from the homologous DNA. **(c)** A RecA protein–mediated branch migration results in formation of a Holliday intermediate, which is then cleaved. **(d)** The lesion can now be repaired, and the transferred strand can be replaced by DNA polymerase and ligase activities.

Site-Specific Recombination Results in Precise DNA Rearrangements

We now turn from general recombination, which can involve any two homologous sequences, to a very different type of recombination that is limited to specific sequences. Site-specific recombination reactions occur in virtually every cell, but their functions are specialized and vary greatly from one species to the next. These functions include the regulation of expression of certain genes, the promotion of programmed DNA rearrangements that occur during development in many organisms, and DNA rearrangements tied to the replication cycle of some viral and plasmid DNAs, as illustrated later. A site-specific recombination system consists of an enzyme called a recombinase and a short (20 to 200 base pairs, depending on the system) unique DNA sequence where the recombinase acts (the recombination site). Some systems also include one or more auxiliary proteins that regulate the timing or outcome of the reaction.

From in vitro studies of more than a dozen site-specific recombination systems, some principles have emerged. The fundamental reaction pathway for many systems is illustrated in Figure 24–35. A recombinase recognizes and binds to each of two recombination sites on dif-

References

1. i Genetics - Russell

2. Cell & Molecular Biology - Karp

3. Gene X - Lewin

4. Principles of Biochemistry - Lehninger, Nelson & Cox