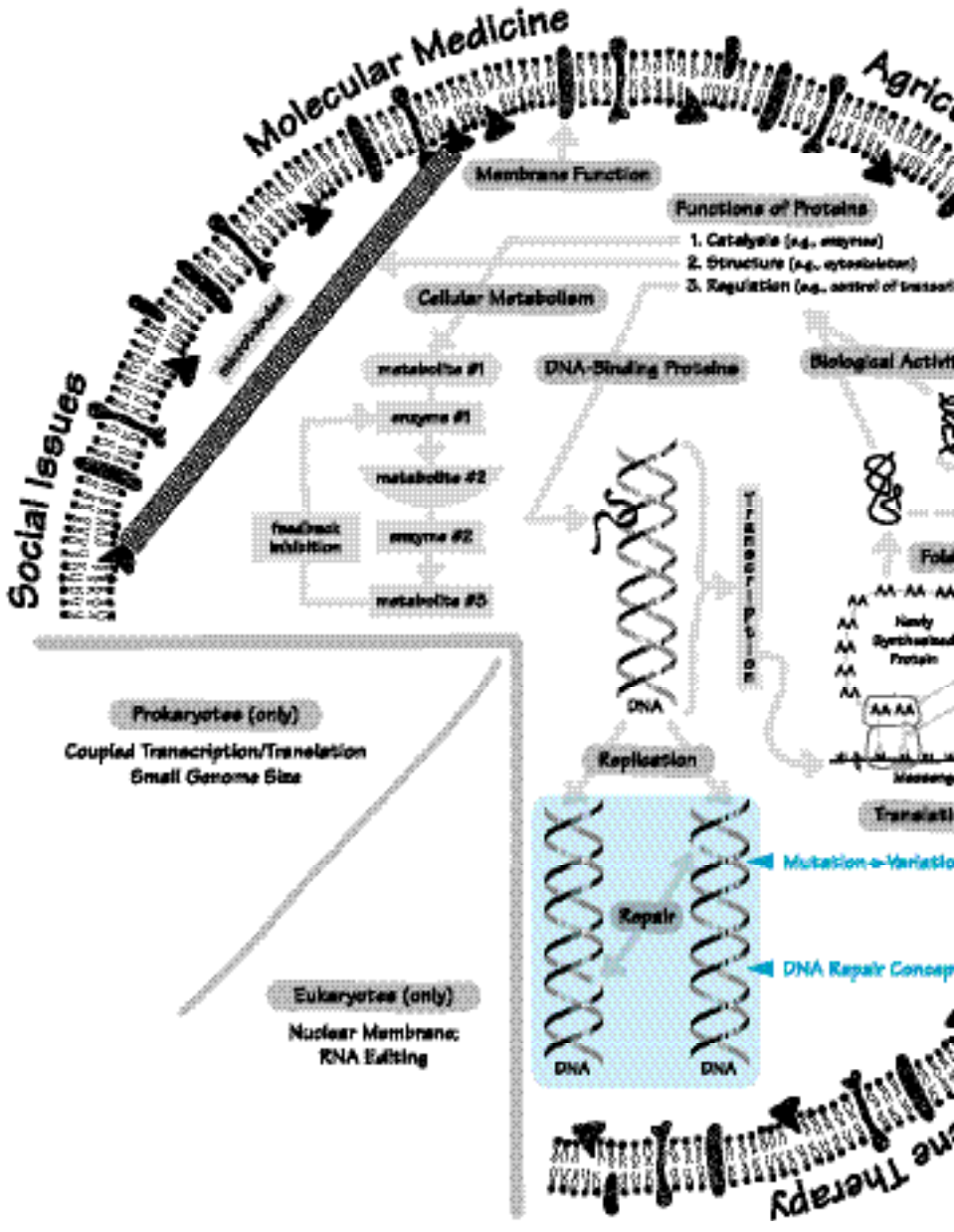
CHAPTER

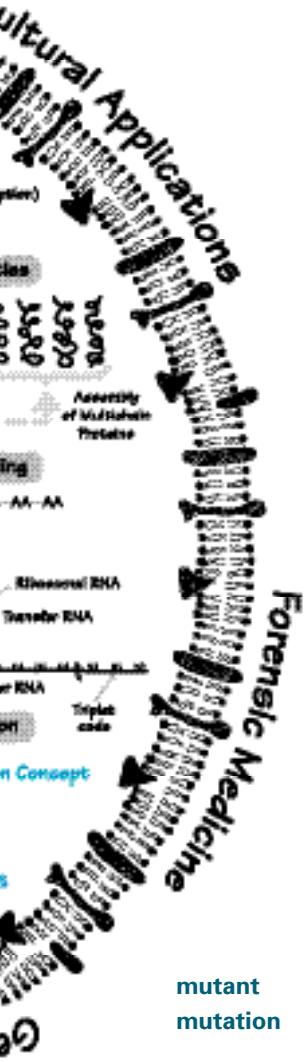
10

in this chapter you will learn

1. About different types of mutations
2. How mutations are produced
3. How damage to DNA is repaired



Mutations, Mutagenesis, and DNA Repair



There is no single molecule whose integrity is as vital to the cell as DNA. Indeed, survival of the species depends upon maintaining the nucleotide sequence intact. The genetic material is, however, subject to constant challenge. During DNA replication, RNA transcription, and even while in an inactive, resting state, damage can occur. Copying errors, breaks, and damage to nucleotide bases, if not corrected, lead to permanent change—mutation. Most mutations are potentially dangerous, since they lead to either blocks in DNA replication or the production of defective proteins. Thus, in the course of hundreds of millions of years there have evolved efficient systems for correcting occasional replication errors and for eliminating damage to DNA caused by environmental agents and intracellular chemicals.

We will begin our study of mutations and DNA repair by defining what it means to be a "mutant" and briefly surveying the various types of mutations that occur in living organisms. Next, we will consider some of the ways in which both spontaneous and induced mutation may arise, and examine the mechanisms of action of specific mutagenic agents. Finally, we will review some of the systems that have evolved to maintain the integrity of the DNA and to repair DNA damage once it does occur.

Types of Mutations

Mutant refers to an organism or a gene that is different from the normal or wild type. His$^-$ yeast or white-eyed *Drosophila* are examples of mutants. Sometimes, the "normal" form of an organism found in nature **lacks** the ability to carry out a certain biochemical reaction (for example, *E. coli* isolated from nature is unable to metabolize lactose, or has a *lac*$^-$ phenotype). In many such instances, the "+" form is referred to as the "wild type," and any "–" form is referred to as a mutant.

Mutation refers to any heritable change in the base sequence of a DNA molecule. The most common type of change is a substitution, addition, or deletion of one or more bases (Table 10-1).

A **mutagen** is a physical or chemical agent that causes mutations to occur (or increases the frequency of their occurrence).

mutant
mutation

mutagen

TABLE 10-1 EXAMPLES OF TYPES OF POINT MUTATIONS

Type of Mutation	Result at Molecular Level	Example
A. *Base Substitution Mutations:*		
Changes in DNA		
1. Transition:	One purine replaced by a different purine; *or* one pyrimidine replaced by a different pyrimidine.	A-T → G-C
2. Transversion:	A purine replaced by a pyrimidine or vice versa.	A-T → T-A
Changes in protein		
1. Silent mutation:	Altered codon codes for same amino acid.	GAG → GAA Glu Glu
2. Neutral mutation:	Altered codon codes for a different but functionally similar amino acid. (Protein *may* be functional.)	GAG → GAS Glu Asp
3. Missense mutation:	Altered codon codes for a different, dissimilar amino acid. (Protein often nonfunctional.)	GAG → AAG Glu Lys
4. Nonsense mutation (= Chain termination mutation):	New codon is a termination codon. (Protein synthesis stops. Protein is nonfunctional.)	GAG → UAG Glu Stop
B. *Frameshift Mutations:* Addition or deletion of one or more base pairs will result in a shift in the reading frame of the resulting mRNA molecule, and lead to production of a nonfunctional protein.		
1. Wild type base sequence:	ATG ACC AGG TC	
2. Base addition:	ATG ACA CAG GTC * ⌊_____⌋	
3. Base deletion:	ATG ACA GGT C Missing C ↑ * ⌊_____⌋	

*Horizontal brackets indicate the affected segment.

mutagenesis

Mutagenesis is the process of producing a mutation. If the mutation occurs in nature without the addition of a mutagen, it is referred to as a **spontaneous mutation,** and the resulting mutants are **spontaneous mutants.** If it is caused by a mutagen, the process is **induced mutagenesis.** In this chapter, we will concentrate on mutations occurring in living organisms. Later, in Chapter 16, we will discuss **site-specific mutagenesis,** a form of induced mutagenesis in which genetic engineering techniques are used to construct mutant DNA molecules containing mutations at specific, preselected locations.

Mutations are classified in several ways. One distinction is based upon the number of bases changed in the DNA molecule. A **point mutation,** for example, involves a single changed base pair, while a **multiple mutation** involves alterations in two or more base pairs. A point mutation may be a **base substitution,** a

base addition, or a **base deletion** (Table 10-1), but the term is usually reserved for base substitutions.

A second distinction is based upon the consequence of the change in terms of the amino acid sequence affected. As described in Table 10-1, a single-base substitution may cause little or no change in the amino acid sequence of the protein or in the ability of the protein to function normally **(silent mutations, neutral mutations),** or it may lead to the production of a nonfunctional and/or truncated polypeptide **(missense mutations, nonsense mutations).**

missense mutations
nonsense mutations

If the substitution produces a protein that is active at one temperature (typically at 30°C or lower) and inactive at a higher temperature (usually 40°–42°C), the mutation is called a **temperature-sensitive** or **ts mutant.** If the mutation generates a stop codon, causing protein synthesis to cease, the mutation is called a **chain-termination** or **nonsense mutation.** Temperature-sensitive and chain-termination mutations are considered to be **conditional mutations** because they exhibit the mutant phenotype only under certain conditions. Such mutations are extremely useful to the molecular biologist, since they can facilitate the study of essential processes, such as DNA replication, without which a mutant organism could not survive.

temperature-sensitive

conditional mutations

With microorganisms, the phenotype is capitalized (Lac^+ or Lac^-), while the genotype is written in lower-case italics (*lac*$^+$ or *lac*$^-$). This convention does not apply to higher organisms. The + or – notation also indicates whether the bacterium in question is able (+) or unable (–) to synthesize or utilize a specific substance. (His^+ bacteria can synthesize histidine, while His^- bacteria cannot.) Another commonly used notation indicates whether a particular bacterial strain is resistant (Amp-r) or sensitive (Amp-s) to a particular antibiotic, ampicillin in this instance. The corresponding genotypes of organisms having these phenotypes would be designated *amp-r* and *amp-s* respectively.

Biochemical Basis of Mutants

A mutant may be defined as an organism in which either the base sequence of the DNA or the phenotype has been changed. These definitions are often the same, since a single base change in a DNA molecule can lead to an alteration in the amino acid sequence of a protein. (Exceptions to this general rule occur in the cases of **silent mutations** and **neutral mutations,** which will be discussed soon.) The chemical and physical properties of each protein are determined by its amino acid sequence, so that a single amino acid change is capable of inactivating a protein.

From the discussion of protein structure in Chapter 4, it is easy to understand how an amino acid substitution can change the structure and biological activity of a protein. For instance, consider a hypothetical protein whose three-dimensional structure is determined entirely by an interaction between one positively charged amino acid (for example, lysine) and one negatively charged amino acid (glutamic acid). A substitution of methionine, which is uncharged, for lysine would clearly destroy the three-dimensional structure, as would the substitution of histidine, which is positively charged, for glutamic acid. Similarly, a protein might be stabilized by a hydrophobic cluster, in which case substitution of glutamine (polar) for leucine (nonpolar) would also be disruptive.

A base substitution does not always yield a mutant phenotype. Because of the redundancy in the code, some base changes do not alter the amino acid sequence **(silent mutations),** and some amino acid changes do not significantly affect the structure of the protein **(neutral mutations).** (See Table 10-1.)

The shapes of proteins are determined by such a variety of interactions that sometimes an amino acid substitution is only partially disruptive. For instance, an isoleucine might substitute successfully for leucine (and be neutral), but replacement with a more bulky amino acid such as phenylalanine might cause subtle stereochemical changes, although the hydrophobic cluster is preserved. This could be manifested as a reduction, rather than a loss, of activity of an enzyme. A bacterium carrying such a mutation in the enzyme that synthesizes adenine might grow very slowly (but it would grow), unless adenine were provided in the growth medium. Such a mutation is called a **leaky mutation;** these mutations are not particularly useful for most genetic studies. However, several hereditary disorders in humans are due to such leaky mutations. For example, in certain individuals, the gene that codes for the essential enzyme, glucose-6-phosphate dehydrogenase, contains a point mutation, resulting in the production of a defective form of the enzyme having greatly reduced catalytic activity. Individuals who possess this mutant gene develop severe hemolytic anemia when exposed to a variety of common substances, including sulfonamide antibiotics, antimalarial drugs, mothballs, and even certain types of dried beans (fava beans).

Generally speaking, the following types of amino acid substitutions are expressed as nonleaky mutations:

Polar amino acid ↔ Nonpolar amino acid

Change of sign of amino acid (+ ↔ −)

Small side chain ↔ bulky side chain

Sulfhydryl ↔ any other side chain

Hydrogen-bonding amino acid ↔ Non-hydrogen-bonding amino acid

Proline (changes the shape of the polypeptide backbone) ↔ any other amino acid

Any change in the substrate-binding site

So far, only amino acid substitutions have been discussed. Certain other types of mutations tend to eliminate the activity of the protein totally. These include **base additions** and **base deletions (= frameshift mutations),** in which all of the amino acids starting from the mutant site are different, and **chain termination mutations (= nonsense mutations),** in which a protein chain is terminated prematurely. Both will be discussed in greater detail later in this chapter.

Mutagenesis

The production of a mutant requires that a change occur in the DNA base sequence. There are a number of distinct mechanisms for altering the structure of

DNA. These include base substitutions, additions, or deletions during replication, base changes resulting from the inherent chemical instability of the bases or of the *N*-glycosylic bond, and alterations resulting from the action of other chemicals and environmental agents. These mechanisms are summarized in Table 10-2.

TABLE 10-2 COMMON DEFECTS IN DNA AND THEIR ORIGINS

Type of Defect	How Does This Type of Change Arise?	
1. Incorrect base in one strand cannot hydrogen bond with corresponding base in the opposite strand	Normal base tautomerizes (i.e., isomerizes in such a way that it is capable of an alternative form of hydrogen bonding); base substitution occurs during subsequent DNA replication	
2. Missing bases	Depurination: *N*-glycosylic bond joining purine base to deoxyribose is spontaneously broken without breaking DNA backbone	
3. Altered bases	Alkylating agents add methyl or ethyl groups to existing bases	
4. Addition or deletion of one or more bases	May occur spontaneously, or be induced by chemical mutagens (intercalating agents) or biological agents (transposable elements)	
5. Single-strand breaks	Phosphodiester bond is broken as a result of exposure to chemical agents or ionizing radiation	Peroxides, metal ions such as Fe^{2+} and Cu^{2+}, x-rays
6. Double-strand breaks	Phosphodiester bonds on opposite DNA strands are broken as a result of exposure to high doses of chemical agents or ionizing radiation	Peroxides, metal ions such as Fe^{2+} and Cu^{2+}, x-rays
7. Cross-linking of complementary DNA strands	Certain antibiotics (mitomycin-C) or reagents (nitrite ions) form covalent bonds between two bases on complementary DNA strands, preventing strand separation during DNA replication	Mitomycin-C, psoralen, *cis*-platinum

Induced Mutations

During its lifetime, an organism may be exposed to a variety of physical, chemical, and biological agents that are capable of causing damage to its genetic material. Each agent has its own characteristic mechanism of action, and each has a tendency to cause specific types of damage to the DNA molecule. With the widespread use of site-specific mutagenesis, as discussed in Chapter 16, chemical mutagens are no longer as widely used experimentally as they were in the past. Nevertheless, they have played a key role in helping us to gain an understanding of mutations and mutagenesis. We will briefly consider some of these classical mutagenic agents. An overview of the mode of action of a number of chemical mutagens is presented in Table 10-3.

Ultraviolet Irradiation

Ultraviolet light is a fairly potent mutagen. The main DNA lesions are two chemically different types of covalent dimers: cyclobutane-pyrimidine dimers and (604) pyrimidine dimers (**Figure 10-1**). The latter only account for 20% of the total dimers, but are correspondingly more mutagenic. In *E. coli* the number of mutants induced by ultraviolet (UV) light can be reduced by exposure to visible light **(photoreactivation),** which implicates cyclobutane pyrimidine dimers since (6-4) dimers cannot be photoreactivated. Bacterial mutants that lack the ability to carry out SOS repair (*lexA*$^-$ and *recA*$^-$ mutants) are not mutagenized by ultraviolet light. These and other experimental results have made it clear that mutagenesis by UV light is almost exclusively a result of error-prone SOS repair. Error-prone repair of pyrimidine dimers leads to the production of both transitions and transversions. Photoreactivation, SOS repair, and two other repair systems employed by cells to deal with UV damage are discussed later in this chapter.

photoreactivation

Mutagenesis by Insertion of Long Segments of DNA (Transposable Elements)

E. coli, and many other organisms as well, contain mobile DNA segments that are hundreds to thousands of base-pairs long (called **transposable elements**). When a transposable element replicates, one replica often remains in the original insertion site, and the other replica is inserted in another region of the chromosome in a complex process discussed in Chapter 14. This process of insertion of a replica at a second site is called **transposition.** When transposition occurs, the sequence frequently inserts itself into a bacterial gene, thereby mutating that gene.

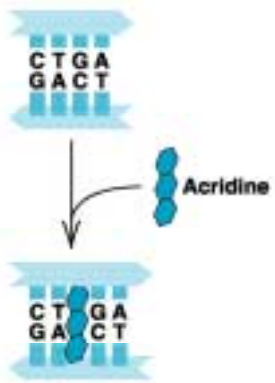
transposition

Some transposable elements contain sequences for termination of transcription. If such an element inserts between two genes that are transcribed as a polycistronic mRNA, or between the promoter and the first gene, all genes downstream from the insertion site will not be transcribed. These mutations belong to a class called **polar mutations.**

Mutator Genes

In *E. coli* there are genes that, when present in a mutant state, cause mutations to appear very frequently in other genes throughout the genetic map. These genes

TABLE 10-3 SOME COMMONLY USED CHEMICAL MUTAGENS AND THEIR MECHANISMS OF ACTION

Type of Mutagenic Agent	Example of This Type of Agent	Mode of Action of This Agent	Diagram of Mode of Action
1. Base analogue: Resembles a normal base found in DNA and is incorporated into DNA; later undergoes a tautomeric shift and mispairs, causing a transition mutation.	5-Bromouracil (base analogue of T) Structure:	Normally pairs with A, but can udergo a tautomeric shift and pair with G. This results in incorporation of C into the daughter DNA strands during subsequent rounds of DNA replication.	AT → GC
2. Nitrous acid	Structure:	Converts amino groups to keto groups by oxidative deamination. C → uracil (U) (pairs with A) A → hypoxanthine (H) (pairs with C) G → xanthine (X) (pairs with C)	AT → GC
3. Hydroxylamine	Structure:	Reacts with C and converts it to a modified base that pairs only with A.	CG → TA
4. Alkylating agents	EMS (ethylmethane sulfonate) and MMS (methylmethane sulfonate) Structure:	Add alkyl groups (ethyl or methyl) to the hydrogen-bonding oxygen of G and T, producing O-6 alkylguanine (pairs with T) and O-4-alkylthymine (pairs with G).	GC → AT
5. Intercalating agents: Planar, three-ringed molecules whose dimensions are roughly the same as those of a purine-pyrimidine base pair.	Proflavine, acridine orange Structures:	Insert between two adjacent base pairs in a DNA molecule, causing insertions or	

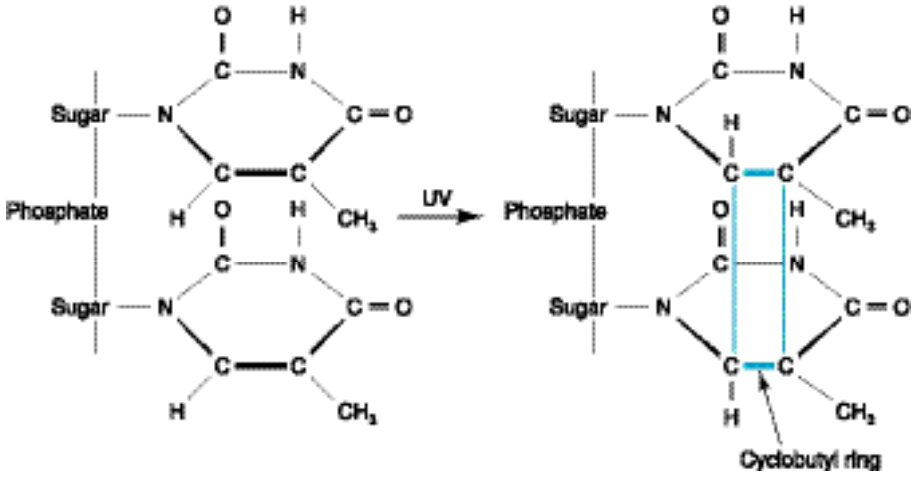


Figure 10-1 Structure of a cyclobutylthymine dimer. Following ultraviolet (UV) irradiation, adjacent thymine residues in a DNA strand are joined by formation of the bond shown in red. Although not drawn to scale, these bonds are considerably shorter than the spacing between the planes of adjacent thymines, so that the double-stranded structure becomes distorted. The shape of the thymine ring also changes as the C = C double bond of each thymine is converted to a C – C single bond in each cyclobutyl ring.

mutator genes

are called **mutator genes.** This is a misnomer, because the function of these genes is probably to keep the mutation frequency low; only when the product of the mutator gene is itself defective will there be widespread production of mutations.

Of the many mutator genes that have been observed, four types that yield strong phenotypic effects are:

1. A mutant DNA polymerase that reduces or eliminates the 3′ → 5′ exonuclease activity of the editing function.
2. A mutant methylating enzyme (the *dam* enzyme) responsible for methylation of the sequences that the mismatch repair system uses to discriminate parental strands from daughter strands.
3. A mutant enzyme that cannot carry out the excision step of mismatch repair.
4. Mutations in the regulatory circuits that maintain the error-prone SOS repair system in an "off" state.

KEY CONCEPT

The "Mutation → Variation" Concept

Mutations, which represent changes in the nucleotide sequence of genes, generate the wide-ranging diversity that characterizes the biological kingdom.

Reversion

reversion

So far we have discussed changes from the wild type to the mutant state. The reverse process, in which the wild type phenotype is regained, also occurs; this process is called **back mutation, reverse mutation,** or, most commonly, **reversion.** It may occur spontaneously, or may be induced by a variety of mutagenic agents. There are several different ways in which reversion may occur. In some cases, the back mutation occurs at the **same site** as the original mutation and restores the wild-type base sequences. This represents a **"true reversion."** In other instances, the wild-type phenotype is restored as a result of a second mutation occurring at a *different* site. Such an event may be thought of as a **"pseudoreversion,"** and mutations of this latter type are described as **second-site** or **suppressor**

suppressor mutations

one lin

mutations. This second site may be situated elsewhere within the *same* gene **(intragenic suppression)** or within a *different* gene **(intergenic suppression).** We will first consider the intragenic type of revertants.

intragenic suppression
intergenic suppression

Intragenic Revertants

Let's consider a hypothetical protein containing 97 amino acids whose structure is determined entirely by an ionic interaction between a positively charged (+) amino acid at position 18 and a negative one (–) at position 64 (**Figure 10-2**). If the (+) amino acid is replaced by a (–) amino acid, the protein is clearly inactive. Three kinds of reversion or suppression events would restore activity (Figure 10-12(a)):

1. The original (+) amino acid could be put back, either as a result of restoration of the *identical* base sequence or because the genetic code is redundant, by substitution of a base sequence that codes for the same amino acid.
2. A different (+) amino acid could be put in position 18.
3. The (–) amino acid at position 64 could be replaced by a (+) amino acid; this second mutation would restore the activity of the protein. A possibility that would not generally work, but which might work in a specific case, is to insert a (+) amino acid at position 17 or 19.

Figure 10-2(b) shows another more complicated example of an intragenic reversion. In this case, the structure of a protein is maintained by a hydrophobic interaction. The replacement of an amino acid with a small side chain by a bulky phenylalanine changes the shape of that region of the protein. A second amino acid substitution providing space for the phenylalanine might restore the protein structure.

The analysis of second-site amino acid substitutions has been an important aid in determining the three-dimensional structure of proteins because the following rule is always obeyed: If a substitution of amino acid A by amino acid X, which creates a mutant, is compensated for by a substitution of amino acid B by amino acid Y, then A and B are either three-dimensional neighbors, or both are contained in two interacting regions.

Revertants of frameshift mutations usually occur at a second site. It is, of course, possible that a particular added base could be removed or a particular deleted base could be replaced by a spontaneous event, but this would not occur very often. Second-site reversion of a frameshift mutation has two requirements, illustrated in **Figure 10-3**:

1. The reverting event must be very near the original site of mutation so that very few amino acids are altered between the two sites.
2. The segment of the polypeptide chain in which both changes occur must be able to withstand substantial alteration.

Intergenic Suppression

Intergenic suppression refers to a mutational change that eliminates or suppresses the mutant phenotype. One type, which occurs when two proteins interact, is a mutation in the binding site of a protein A that prevents the protein from interacting with protein B. A second mutation, in the binding site of protein B, alters this binding site so that the mutant protein B can bind to the mutant protein A. As a result, the interaction between the two proteins is restored. The occurrence of

e long

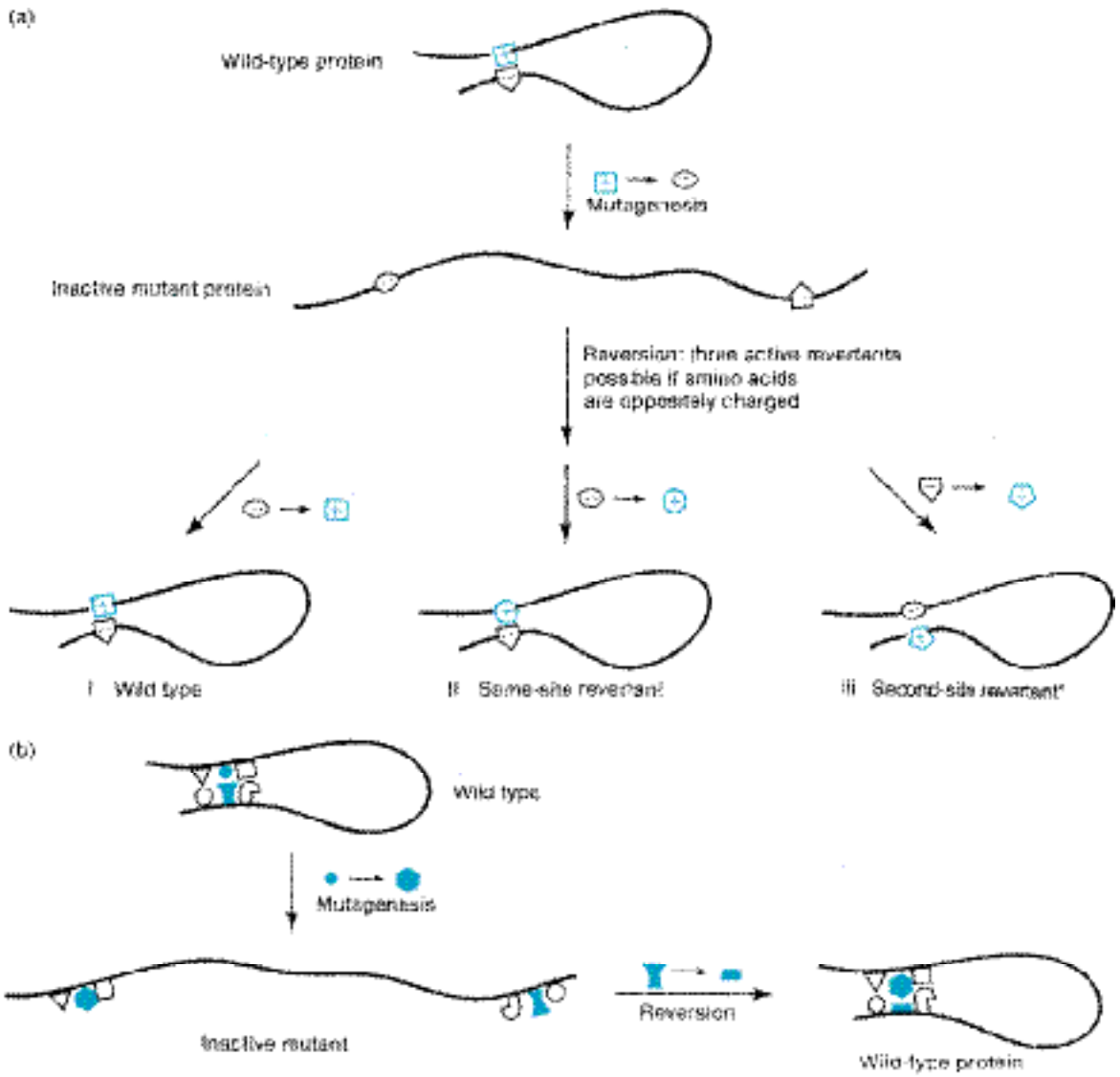


Figure 10-2 Several mechanisms of reversion. In panel (a) the charge of one amino acid is changed, and the protein loses activity. The activity is returned by (I) restoring the original amino acid, (II) replacing the (–) amino acid by another (+) amino acid, or (III) reversing the charge of the original (–) amino acid. In each case the attraction of opposite charges is restored. In panel (b) the structure of the protein is determined by interactions between six hydrophobic amino acids. Activity is lost when the small circular amino acid is replaced by the bulky hexagonal one, and is restored when space is made by replacing the convex amino acid with the small rectangle.

intergenic suppression of this kind is an important indicator of the interaction between two proteins and is a splendid example of a genetic result giving information about molecular structure.

A second type of intergenic suppression has the remarkable property that the second-site mutation not only eliminates the effect of the original mutation, but

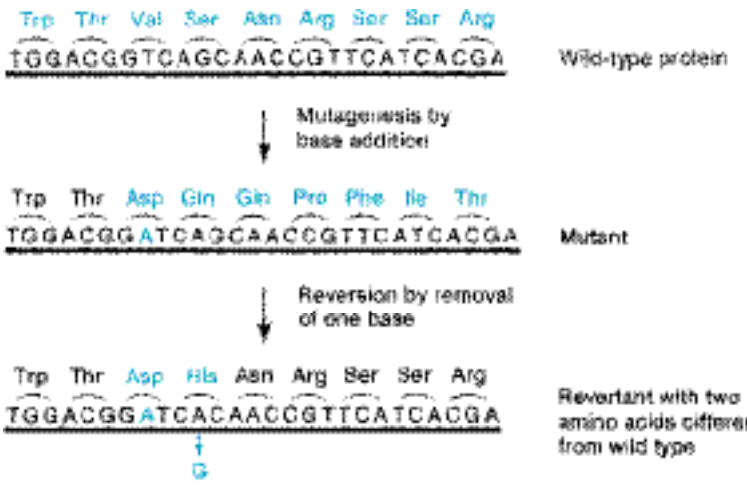


Figure 10-3 Reversion by base deletion from an acridine-induced, base-addition mutant.

also suppresses mutations in many other genes as well. This type of suppression, which is produced by mutations in certain tRNA molecules and aminoacyl tRNA synthetases, can be seen most clearly when suppression of chain-termination mutations is examined.

Chain-termination mutations (also called **nonsense mutations**) are common, for they can arise in many ways. For example, a single-base change in any of the codons AAG, CAG, GAG, UCG, UUG, UGG, UAC, and UAU can give rise to the chain-termination codon UAG. If such a mutation occurs within a gene, a mutant protein with little or no function will result because no normal tRNA molecule exists whose anticodon is complementary to UAG. Only a fragment of the wild-type protein is produced, and this usually fails to function unless the mutation is very near the carboxyl terminus of the protein.

In certain bacterial strains, the presence of a chain-termination mutation is not sufficient to stop polypeptide synthesis. For example, a phage may have acquired a UAG codon in a gene encoding a critical protein. When the phage infects most host bacteria, no phage progeny will be produced. In a particular bacterial strain, however, the mutant phage may grow normally, indicating that the mutation is made silent by some element in the bacterium. Such a bacterium is said to be able to **suppress** the mutation and to contain a **suppressor.** A bacterium able to suppress a particular type of chain-termination mutation—a UAG codon in the example—is usually able to do so with a large number of mutations in that class, whether the mutation is in a phage or in the bacterium itself. In general, other types of chain-termination mutations, for example, UGA, will not be suppressed. The explanation for this phenomenon is that the bacterium (called a **suppressor mutant**) contains an altered tRNA molecule that can respond to a particular stop codon. For UAG, the altered tRNA might contain the anticodon CUA, which can pair with that codon. Such a tRNA molecule is called a **suppressor tRNA** (or a **nonsense suppressor tRNA**), and the mutation on which it can act is said to be suppressor sensitive.

suppressor

suppressor mutant

suppressor tRNA

nonsense suppressor tRNA

What has been mutated in the production of a suppressor tRNA? Clearly, it must be a normal tRNA gene. Therefore, in the example just given, a $tRNA^{Lys}$

molecule whose anticodon is CUU has been altered to have an anticodon CUA, which can hydrogen bond to the codon UAG.

Inasmuch as a single-base change is sufficient to alter the complementarity of an anticodon and a codon, there are (at most) eight tRNA molecules having a complementary anticodon that, with a single changed base, will also suppress a UAG codon. Thus, the following amino acids (whose codons are also indicated) can be put in the site of a chain-termination codon: Lys (AAG), Gln (CAG), Glu (GAG), Ser (UCG), Trp (UGG), Leu (UUG), and Tyr (UAC and UAU). Note that these are the same amino acid codons that can be altered by mutation to form a UAG site. Suppressors also exist for chain-termination mutants of the UAA and UGA types. These too are mutant tRNA molecules whose anticodons are altered by a single-base change.

In conventional notation, suppressors are given the genetic symbol *sup*, followed by a number (or occasionally a letter) that distinguishes one suppressor from another. A cell lacking a suppressor is designated *sup0* and sup^-.

Several features of nonsense suppressors should be recognized:

1. Not every UAG suppressor can restore a functional protein by suppressing each UAG chain-termination codon. Thus, a UAG codon produced by mutating the leucine UUG codon might be suppressed by a suppressor tRNA that inserts tyrosine, serine, or tryptophan, but might not be able to tolerate a substitution by the electrically charged amino acids lysine, glutamine, or glutamic acid.
2. Suppression may be incomplete in that the activity of the suppressed mutant protein may not be as great as that of the wild-type protein, and the stop codon may not always be read as a sense codon.
3. A cell can survive the presence of a suppressor only if the cell also contains two or more copies of the tRNA gene. Clearly, if a $tRNA^{Ser}$ molecule that reads the UCG codon is mutated, then the UCG codon can no longer be read as a sense codon. This will lead to chain termination wherever UCG occurs, and a cell harboring such a mutant tRNA molecule will fail to complete virtually every protein made by the cell. Therefore, in any living cell containing a suppressor tRNA, there must always be an additional copy of a wild-type tRNA that can function in normal translation.

If a cell contains a UAG suppressor, the proteins terminated by a single UAG codon will not be terminated, and the existence of a suppressor tRNA should be lethal. There are two ways that this problem can be avoided:

1. Protein factors active in termination (see Chapter 9) respond to chain-termination codons even though a tRNA molecule that recognizes the codon is present; i.e., suppression is weak.
2. Normal chain termination often uses pairs of distinct termination codons such as the sequence UAG-UAA. The existence of a UAG suppressor, thus, would not prevent termination of a double-terminated protein.

Suppression of missense mutations also occurs. For example, a protein in which valine (nonpolar) has been mutated to aspartic acid (polar), resulting in loss of activity, is restored to the wild-type phenotype by a missense suppressor that substitutes alanine (nonpolar) for aspartic acid. Such a mutation can occur in three ways: (1) a mutant tRNA molecule may recognize two codons, possibly by a

change in the anticodon loop; (2) a mutant tRNA can be recognized by a noncognate aminoacyl tRNA synthetase and be misacylated; and (3) a mutant synthetase can charge a noncognate tRNA molecule. Examples of each type of suppression are known. Suppression of missense mutations is necessarily inefficient. If a suppressor that substitutes alanine for aspartic acid worked with, say, 20% efficiency, then, in virtually every protein molecule synthesized by the cell, at least one aspartic acid would be replaced and the cell could not possibly survive. The usual frequency of missense suppression is about 1%.

Reversion As a Means of Detecting Mutagens and Carcinogens

In view of the increased number of chemicals present as environmental contaminants, and because many cancer-causing agents **(carcinogens)** are also mutagens, tests for the mutagenicity of these substances have become important. One simple method for screening large numbers of substances for mutagenicity is a reversion test using nutritional mutants of bacteria. In the simplest reversion test, known numbers of a mutant bacterium are plated on growth medium containing a potential mutagen, and the number of revertant colonies that arise is counted. If the substance is a mutagen, the number of colonies will be greater than that obtained in the absence of the mutagen. However, such simple tests fail to demonstrate the mutagenicity of many carcinogens. These substances are not directly mutagenic (or carcinogenic), but are converted to actively mutagenic compounds by enzymatic reactions that occur in the livers of mammals, and that have no counterpart in bacteria. The normal function of these enzymes is to protect the organism from various noxious substances that occur naturally by enzymatically converting them to nontoxic substances. However, when the enzymes encounter certain man-made and natural compounds, they convert these substances (which may not themselves be directly harmful) to mutagens or carcinogens. The enzymes are contained in the microsomal fraction of liver cells. Addition of the microsomal fraction to the bacterial growth medium allows these substances to undergo enzymatic "activation," and enables researchers to determine their mutagenicity under conditions more closely resembling those found in the liver. This is the basis of the **Ames test** for potential carcinogens.

carcinogens

Ames test

In the Ames test, histidine-requiring (His^-) mutants of the bacterium *Salmonella typhimurium*, which contain either a base substitution or a frameshift mutation, are used to test for reversion to His^+. The frequency of spontaneous His^+ revertants is low in this mutant, but mutants are readily produced in one or both of these mutants by most known mutagens. Solid medium is prepared containing a very small amount of histidine, sufficient to initiate the growth of individual cells, but not enough to support colony formation. Minimal histidine is necessary because many mutagenic agents act only on cells that are actively involved in DNA replication, and two rounds of DNA replication and cell division are often required for the new mutant phenotype to be expressed. A small amount of rat liver extract and about 10^8 His^- mutants are spread on each plate. Then the substance to be tested is applied to one set of plates (Group A = experimental group),

while distilled water is applied to the plates in the control group (Group B), which lacks the carcinogen. The number of colonies appearing on the plates in Group B (control) is usually about 5 to 10 (these are the spontaneous revertants). When a known mutagen is present, more colonies will be present on the plates in Group A. The number of colonies on the Group A plates depends upon the concentration of the substance being tested and, for a known mutagen, correlates roughly with its known effectiveness as a mutagen (and, hence, as a potential carcinogen).

The Ames test has not been used with thousands of substances and mixtures (industrial chemical, food additives, pesticides, hair dyes, and the like) and numerous unsuspected substances have been found to stimulate reversion in this test. This does not mean that the substance is definitely a carcinogen, but only that it has a high probability of being one. As a result of these tests, many industries have reformulated their products to render them nonmutagenic. Ultimate proof of carcinogenicity is determined by testing for tumor formation in laboratory animals. The Ames test and several other microbiological tests are used to reduce the number of substances that need to be tested in animals, since, to date, only a few percent of the more than 300 substances known to be carcinogens from animal experiments have failed to increase the reversion frequency in the Ames test.

DNA Repair Mechanisms

Maintaining the integrity of the genetic information is essential to the survival of the cell. Without a doubt, this is the reason that living organisms have evolved such a diverse repertoire of DNA repair mechanisms. The process begins during DNA replication, and continues, in various forms, throughout the postreplication period. Let us now begin to consider the intricacies of DNA repair.

Spontaneous Mutations and Their Repair

Spontaneous mutations are those that arise naturally, not through the action of a mutagenic agent. They may arise through (1) errors in DNA replication or (2) spontaneous alteration of a nucleotide within an existing DNA molecule.

As has already been described in Chapter 7, DNA polymerases occasionally catalyze the incorporation of an incorrect base that cannot form a hydrogen bond with the template base in the parent strand. Such errors are usually corrected by the editing (or proofreading) function of these enzymes. In the event that an occasional misincorporated base is overlooked during proofreading, a second correction system, called **mismatch repair,** exists to deal with such mispaired bases (**Figure 10-4**). In mismatch repair, a pair of non–hydrogen-bonded bases is recognized as incorrect, and a polynucleotide segment is excised from one strand, thereby removing one member of the unmatched pair.

mismatch repair

If it is to eliminate errors, the mismatch repair system must be able to distinguish the correct base in the parental strand from the incorrect base in the daughter strand. In *E. coli*, the critical information is provided by the presence or absence of **adenine methylation** in the base sequence G-A-T-C. These bases carry methyl groups not found elsewhere in the DNA. Methylation of the adenines in this

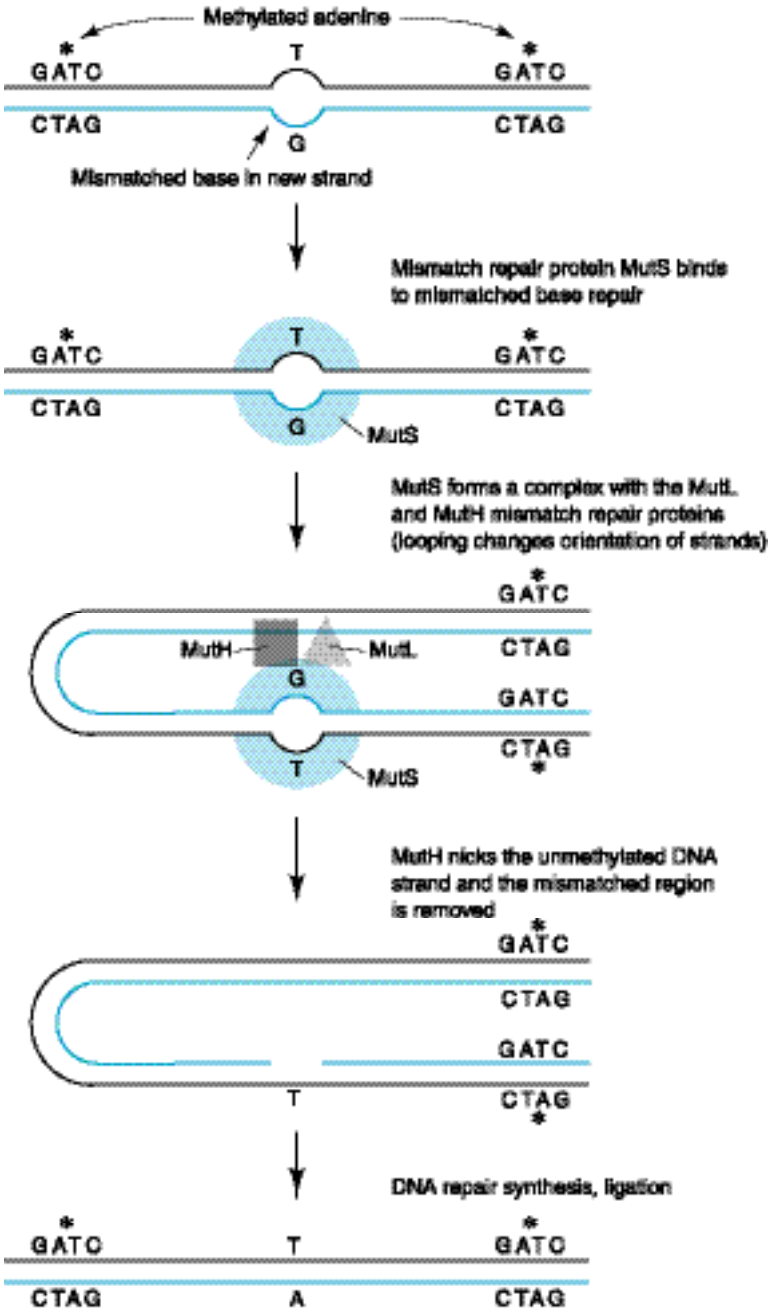


Figure 10-4 Possible mechanism of mismatch repair in *E. coli*. Mismatch repair enzyme recognizes mismatched base pair and undermethylated daughter DNA strand. Mismatched region is excised and repaired.

sequence is associated with DNA replication. However, it does not take place in the replication fork, but is somewhat delayed. The effect is that, while parental strands are fully methylated, the newly synthesized daughter strands near the replication fork are not yet methylated. The mismatch repair system recognizes the degree of methylation of each strand, and when a mismatch is found, it preferentially excises nucleotides from the *undermethylated* strand—that is, from the daughter strand (see Figure 10-3). Thus, the parent strand is always the template, enabling the repair system to correct misincorporation errors. However, despite of the existence of proofreading and mismatch repair, errors occasionally escape detection and give rise to mutations.

Spontaneous mutations may also arise as a consequence of tautomeric shifts in the nucleotide bases. Each of the four nitrogen bases is capable of existing in a rare, alternate isomeric form **(tautomer),** having different hydrogen-bonding properties (Figure 10-1). The rare imino forms of adenine **(A*)** and cytosine **(C*)** base pair with cytosine and adenine, respectively, while the rare enol forms of thymine **(T*)** and guanine **(G*)** hydrogen bond with guanine and thymine, respectively. A base might be incorporated in its normal form, then undergo a tautomeric shift after incorporation. During a subsequent round of DNA replication, the presence of one of these rare tautomers in the template could result in the incorporation of an incorrect base into the daughter DNA molecule. (A similar process occurs when mutagenic agents called **base analogues** undergo tautomerization, as we will see later, in **Figure 10-5**.)

base analogues

Depurination refers to the breakage of the *N*-glycosylic bond between a base (in this case a purine) and the deoxyribose of the nucleotide, with the subsequent loss of the nitrogen base (Figure 10-2). Unless such damage is repaired prior to DNA replication, there is a great likelihood that a mutation will arise, since the **apurinic site** lacks the "information" necessary to specify the insertion of the correct complementary base into the new DNA strand. If the DNA polymerase should happen to reach the apurinic site before repair has taken place, replication might come to a halt, and/or an incorrect base might be inserted into the new DNA strand opposite the apurinic site. Once again, specific repair mechanisms involving enzymes called **AP endonucleases (AP = apurinic)** exist to deal with such damage, and will be discussed in the next section.

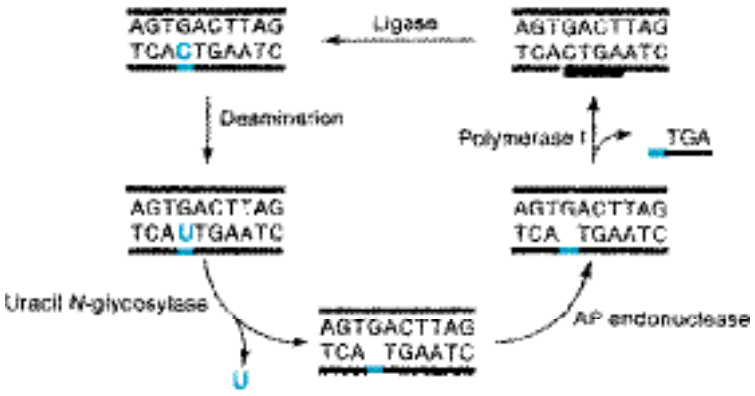


Figure 10-5 Scheme for repair of cytosine deamination. The same mechanism could remove a uracil that is accidentally incorporated.

Deamination

Deamination is another process that can give rise to mutations. It has been estimated that, in an average human cell, cytosine deamination occurs at a rate of 100 cytosines per genome each day. When cytosine loses an amino group, it becomes uracil. After one round of replication, this would lead to replacement of the original G-C pair with an A-U pair, which would then become an A-T pair after another round of replication. Since this would be mutagenic, cells have evolved a mechanism for replacing the unwanted U by a C. The first step in their repair cycle (Figure 10-5) is removal of the uracil by the enzyme **uracil *N*-glycosylase.** This enzyme cleaves the *N*-glycosylic bond and leaves the deoxyribose in the backbone. A second enzyme, **AP endonuclease** (which acts on apurinic or "baseless" sites in general), makes a single cut in the DNA backbone on the 5′ side of the damaged site, and the sugar phosphate residue is removed by a **phosphodiesterase.** DNA polymerase then fills the gap with the correct nucleotide and ligase seals the gap. (This sequence, endonuclease-exo-nuclease-polymerase, is an example of a general repair mechanism called excision repair, described later in the chapter.) A specific glycosylase is also available to remove hypoxanthine, the consequence of adenine deamination.

uracil *N*-glycosylase

The latter part of this repair mechanism, namely that beginning with AP endonuclease (shown in Figure 10-5), is also used in the repair of missing bases arising by depurination.

Mutational Hot Spots and 5′-Methycytosine Deamination

If a hundred mutations in a single gene are mapped, they are, for the most part, distributed roughly equally over the mutated sites. However, a few sites are represented by as many as 100 times the typical number of mutations; these sites are called **hot spots.**

hot spots

About 4% of the cytosines in a typical DNA molecule are in the methylated form 5′-methylcytosine (MeC). The role of methylcytosine (and other methylated bases) is not clearly understood in many cases. Earlier in this chapter, we described the role of adenine methylation in mismatch repair, and we shall examine the additional role of cytosine methylation in regulation of gene expression in vertebrates in Chapter 13. At any rate, methylated cytosines are not inherently harmful and do not change the hydrogen bonding properties of the base. MeC pairs with guanine just as cytosine does.

MeC is also subject to alteration by spontaneous deamination, and this leads to a curious situation. When MeC is deaminated, it becomes 5-methyluracil, which is another name for the normal base thymine. Therefore, the G-MeC base pair becomes a G-T pair, which in subsequent replication yields an A-T pair. Note that the main mismatch repair system is certainly able to convert the G-T pair back to a correct G-C pair, but *only* if it is able to distinguish between the parental and daughter strand on the basis of degree of adenine methylation. Unfortunately, spontaneous deamination can also occur in nonreplicating DNA (e.g., in a resting cell or in a phage), where both strands have been equally methylated by the adenine-methylating system. The mismatch repair system, thus, receives no signal, indicating that the G-C pair is the correct one, and could just as well convert the G-T

pair to an A-T pair. The mutation frequency consequently can be quite high at a MeC site. A given G-MeC → A-T transition will, of course, only produce a mutation if the change causes an amino acid substitution that affects the activity of the gene product. Such MeC sites do not occur very often, so hot spots should not be particularly frequent. Direct determination of the base sequence of several genes and of hot spot mutants has shown that, indeed, MeC accounts for most of the hot spots for spontaneous mutagenesis.

A repair system does exist that is capable of correcting the results of MeC deamination. A specific mismatch repair enzyme recognizes mismatched G-T base pair and *always* removes the mispaired thymine, rather than the guanine, so that the correct (G-C) base-pairing sequence can be restored. The new cytosine can be remethylated at a later time.

Repair by Direct Reversal

The simplest mechanism for repair of altered bases is to restore them to normal without major surgery on the DNA molecule. In the case of pyrimidine dimers formed in DNA by ultraviolet light (see Figure 10-1), this reversal can be accomplished very simply by an enzyme that recognizes and binds to these dimers. This enzyme, called **photolyase,** is activated by visible light and cleaves the dimers to yield intact pyrimidines. Photolyase is found in many types of cells, but operates only on pyrimidine dimers and only when activated by light (300–600 nm).

photolyase

Another example of direct reversal is the methyltransferase (an alkyltransferase) that recognizes 0^6 methyl guanine in DNA and removes the offending methyl group by attaching it instead to its own cysteines. The exchange reaction restores the normal guanine in the DNA, but inactivates the enzyme. The importance of this particular repair scheme, which is used to repair damage caused by alkylating agents such as MMS, is underscored by the fact that an entire protein molecule is expended for each 0^6 methyl guanine repaired.

Excision Repair

The most ubiquitous repair scheme, and one that can deal with a large variety of structural defects in DNA, is called excision repair. This multistep enzymatic process is best understood in *E. coli* (in which it was originally discovered). It is a mechanism by which pyrimidine dimers can be repaired in the DNA without photolyase. The essential steps of this process, as it occurs in *E. coli*, are illustrated in **Figure 10-6**. Variations on this basic "cut and patch" theme occur in other organisms. In the first step, **incision,** a repair endonuclease recognizes the distortion produced by a thymine dimer (or other bulky lesions) and makes two cuts in the sugar-phosphate backbone, 12 to 13 nucleotides apart, on each side of the dimer; (one cut is located eight nucleotides to the 5′ side, the other four to five nucleotides to the 3′ side). In *E. coli*, the product of three genes (*uvrA, uvrB,* and *uvrC*) act in concert to produce the incisions. In other systems, a single incision is adequate to initiate the repair sequence. At the incision site illustrated in Figure 10-14, there is a 5′-P group on the side of the cut containing the dimer, and a

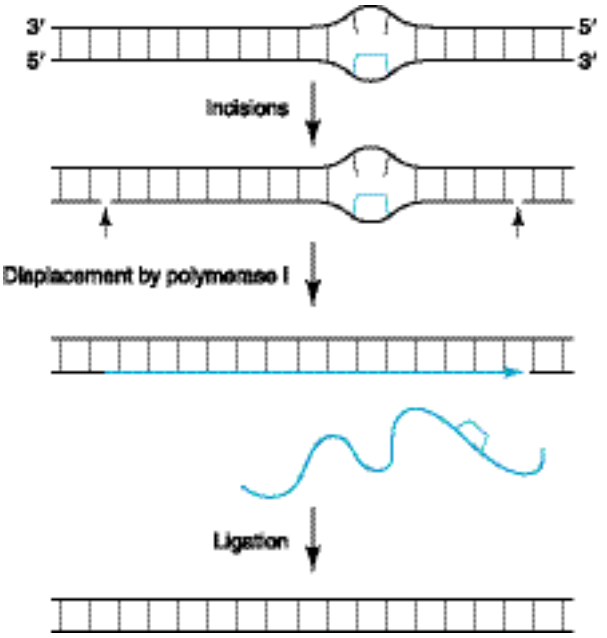


Figure 10-6 Scheme for excision repair of a thymine dimer by the "cut and patch" mechanism. The thymine dimer and the repair patch are both shown in red.

3′-OH group on the other side. The 3′-OH group is recognized by polymerase I, which can then synthesize a new strand while simultaneously displacing the DNA segment carrying the thymine dimer. In the final step of the process, ligase joins the newly synthesized segment to the original DNA strand.

The excision-repair system has been shown to preferentially repair pyrimidine dimers that are located in certain essential regions of the genome, notably within expressed genes. In genes that are being actively transcribed, the repair is somehow targeted selectively to the strand that is being used as a template for mRNA synthesis. This special treatment makes sense because a pyrimidine dimer poses an absolute block to the process of transcription and could kill a cell quite directly if it prevented expression of an essential gene.

In mammalian cells, recall that chromatin structure is much more complex than it is in bacteria such as *E. coli.* Thus, the process of excision repair involves many more enzymes. At least a dozen different genes are involved in recognizing and excising the damaged DNA segment. Individuals who suffer from a rare genetic disorder known as **xeroderma pigmentosum (XP)** are extremely sensitive to sunlight and have a tendency to develop skin cancer after very short exposures to the sun. Recent evidence has shown that xeroderma pigmentosum can result from a mutation in any one of seven different genes involved in various steps in the human excision-repair process. Cells from XP patients are sensitive to UV light and have a reduced ability to remove thymine dimers from their DNA. Another known genetic disease, Cockayne's syndrome, is also characterized by sensitivity to UV light, but without the tendency to develop skin cancer. Its victims do suffer from a variety of other problems, including neurological and developmental abnormalities. It now appears that some of the genes involved in the excision-repair pathway may play a dual role. In fact, one of the proteins involved

in the unwinding of the damaged segment of DNA (the *XPB* gene product, which functions as a helicase) also functions as a component of a crucial transcription factor (TFIIH), which is involved in regulating the activity of all protein-encoding genes. This could explain the observation that the defect in Cockayne's syndrome affects the preferential repair of expressed genes, as previously discussed.

Recombinational Repair

The excision-repair systems just described are responsible for the removal of many thymine dimers and other lesions. However, before sufficient time has elapsed for their repair, many dimers have interfered with various cellular processes. The deleterious effects of some of the remaining dimers are eliminated by recombinational repair, which is carried out by the system responsible for genetic recombination.

In order to discuss the mechanism of recombinational repair, it is necessary to know the effect of a thymine dimer on DNA replication. When polymerase III reaches a thymine dimer, the replication fork fails to advance. A thymine dimer is still capable of forming hydrogen bonds with two adenines because the chemical change in dimerization does not alter the groups that engage in hydrogen bonding. However, the dimer introduces a distortion into the helix, and when an adenine is added to the growing chain, polymerase III reacts to the distorted region as if a mispaired base had been added; the editing function then removes the adenine. The cycle begins again—an adenine is added and then it is removed; the net result is that the polymerase is stalled at the site of the dimer. (The same effect would occur if, instead of a dimer, radiation or chemical damage resulted in the formation of a base with which no nucleoside triphosphate could base pair.) Evidence that such a phenomenon occurs after ultraviolet irradiation is the existence of a UV-light-induced idling process—that is, rapid cleavage of deoxynucleoside triphosphates to monophosphates without any net DNA synthesis (i.e., without advance of the replication fork). A cell in which DNA synthesis is permanently stalled cannot complete a round of replication, and does not divide.

There are basically two different ways in which DNA synthesis can get going again—**postdimer initiation** and **transdimer synthesis.**

One way a cell could deal with a thymine dimer is to pass it by and reinitiate chain growth beyond the block, perhaps at the starting point for the next Okazaki fragment (**Figure 10-7**). The result of this process is that the daughter strands would have large gaps; one for each unexcised thymine dimer. There is no way to produce viable daughter cells by continued replication alone, because the strands with

Figure 10-7 Blockage of replication by thymine dimers (represented by joined lines) followed by restarts several bases beyond the dimer. The black region is a segment of ultraviolet-light-irradiated parent DNA. The blue region represents synthesis of a daughter molecule from right to left. The daughter strand contains gaps.

the thymine dimers would continue to turn out gapped daughter strands, and the first set of gapped daughter strands would be fragmented when the growing fork entered a gap. However, by a recombination mechanism known as **sister-strand exchange,** proper double-stranded DNA molecules known as **sister-strand exchange,** proper double-stranded DNA molecules can be produced.

The essential idea in sister-strand exchange is that a single-stranded segment free of any defects is excised from a "good" strand on the homologous DNA segment and inserted into the gap created by the bypassing of the thymine dimer during replication (**Figure 10-8**). This recombinational event requires the RecA protein. The gap created in the donor molecule by excision of the "patch" is then repaired by polymerase I and ligase. If this exchange and gap filling are done for each thymine dimer, two complete daughter strands can be formed, and each can serve in the *next round* of replication as a template for synthesis of normal DNA molecules. Note that the system fails if two thymine dimers in opposite strands are very near to one another because then no undamaged sister-strand segments are available to be excised. Many of the molecular details of recombinational repair are not yet known, so the model shown in Figure 10-8 must be considered to be a working hypothesis reflecting our present state of knowledge.

Recombinational repair is an important mechanism because it eliminates the necessity for delaying replication for the many hours that would be needed for excision repair to remove all thymine dimers. It may also be the case that some kinds of damage cannot be removed by excision repair—for example, alterations that do not cause helix distortion, but do stop DNA synthesis.

Since recombinational repair occurs after DNA replication, in contrast with excision repair, it is often called **postreplicational repair.** It should be noted, however, that what is repaired is the gap in the daughter strand, rather than the thymine dimer.

Despite the evidence that thymine dimers and other helix-distorting lesions block the progression of DNA polymerase, it is clear that most types of cells, including mammalian cells, can tolerate large numbers of persisting lesions without

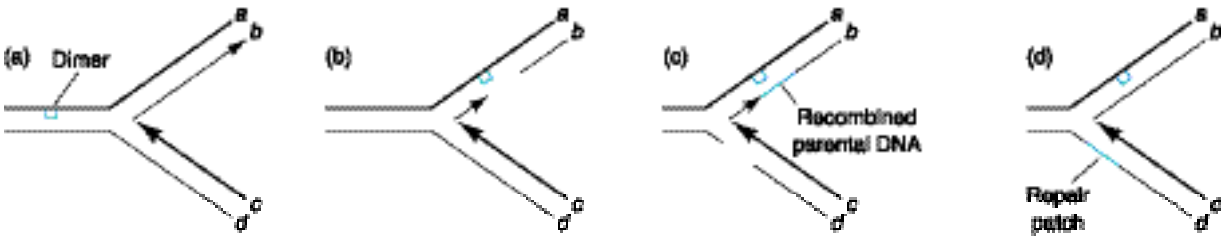


Figure 10-8 Recombinational bypass. (a) In strand *a,* a Molecule containing a thymine dimer (red box) is being replicated. (b) Replication is blocked at the site of the dimer, leaving a gap because an Okazaki fragment cannot be completed. (c) That gap is filled by a sister–strand recombination event in which a segment of the parent strand *d* is utilized. That now leaves a gap in strand *d.* (d) The daughter strand *c* can serve as a template for repair DNA synthesis to fill the gap in strand *d.* DNA synthesized after irradiation is shown in red. Heavy and thin lines are used only to identify strands of the same polarity.

resorting to recombinational strand exchange. Somehow replication is able to overcome the lesions, albeit with an expected high error rate at those noncoding sites.

The SOS Response

SOS response system

While most DNA repair mechanisms are constitutive (i.e., active all the time), a few are activated in response to some signal such, as a blocked replication fork. Most notable is the **SOS response system,** characterized in *E. coli* as a complex regulatory scheme in which the products of two genes, *recA* and *lexA*, govern the expression of a number of other genes involved in DNA repair (**Figure 10-9**). Some of these genes (*uvrA* and *uvrB*) are involved in excision repair, while others like *umuC* and *umuD* are required to help replication bypass the offending lesion. In order to continue DNA replication across a region containing a thymine

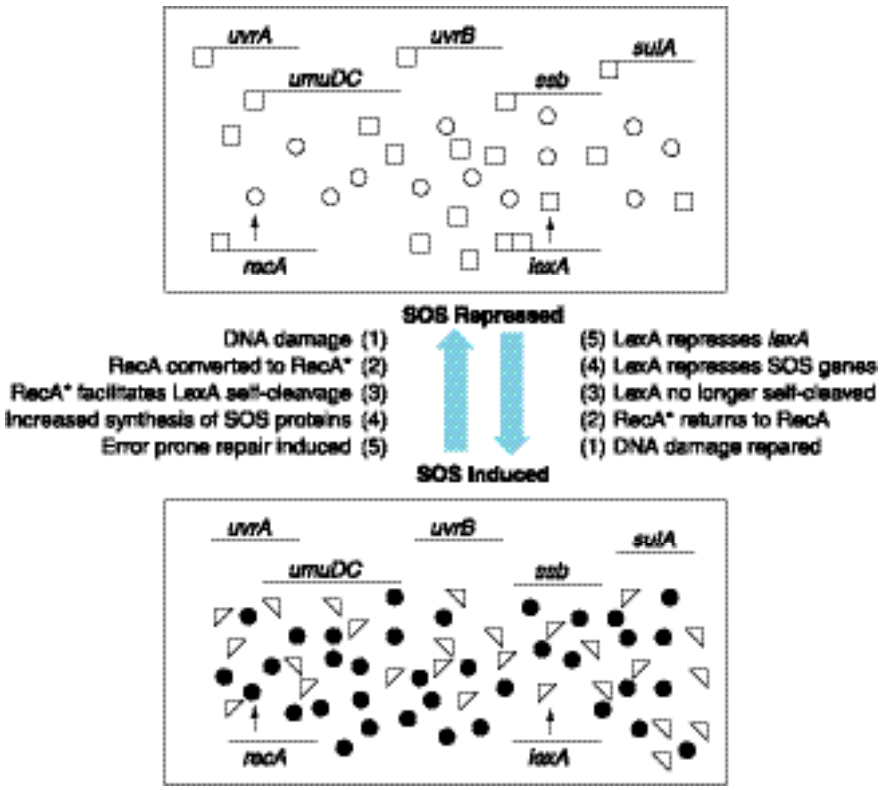


Figure 10-9 The SOS response. In uninduced cells, the RecA protein [◯] is not activated and, thus, does not facilitate self-proteolysis of the LexA protein [□]. The LexA protein functions as a repressor, turning off transcription of many different genes, including the *recA* gene and the *lexA* gene itself (i.e., it is autoregulatory). DNA-damaging agents induce the SOS response by activating RecA protein to RecA*[●], which facilitates the self-cleavage of LexA protein [◹] and several other proteins, including the λ repressor. After LexA protein is cleaved, it cannot function as a repressor, resulting in transcription of all of the genes regulated by the LexA protein. As the SOS genes repair the DNA damage, RecA* returns to RecA, LexA is no longer cleaved, and accumulation of LexA represses the SOS genes.

dimer, the editing function of polymerase III must be relaxed, otherwise the helical distortion caused by the dimer would trigger pol III's 3′ → 5′ proofreading activity, and the replication fork would become stalled. Relaxation of proofreading is not without its own inherent hazards. It leads to **error-prone DNA replication,** and results in a higher-than-normal level of mutagenesis. Because this is detrimental to the cell, expression of the genes involved in the SOS response must be carefully regulated in order to maintain the fidelity of DNA replication under normal circumstances. In the presence of certain types of potentially lethal DNA damage, the system is switched "on" as a last-ditch effort to allow the cell to survive; (hence the name "SOS" response). The enhanced mutagenesis that results may be seen as a benefit in that it may result in progeny that, through mutation, are better adapted to live in the noxious environment that led to the induction of the SOS system.

The genes of the SOS response are ordinarily maintained in an "off" state. This is accomplished by the *lexA* gene product, the LexA repressor, which binds to an operator sequence near each gene and prevents its transcription. The LexA repressor even represses its own expression, so that only small amounts of this protein are synthesized routinely. (Regulation of gene expression in prokaryotes will be discussed in greater detail in Chapter 11.) The RecA protein plays two different, complementary roles in the SOS response. First, it inhibits the editing function of DNA polymerase III. It binds to the distorted region of the DNA molecule containing the dimer; when pol III encounters this region, the RecA protein interacts with the polymerase subunit involved in proofreading. This results in inhibition of pol III's editing function and allows the replication to proceed.

Binding of the RecA protein to the single-stranded DNA also has a second effect: it causes a conformational change in the RecA protein (→ RecA*). This activated form of the RecA* protein interacts with the LexA repressor protein, stimulating the LexA repressor to inactivate itself by proteolytic cleavage. With the repressor inactivated, all of the genes involved in the SOS response can be expressed, and the enzymes required for DNA repair are synthesized. (Although this also leads to high levels of expression of the *lexA* gene, the repressor will continue to be cleaved as long as the activated RecA* is present.) Eventually, DNA repair will be completed and the RecA protein will once again lose its proteolytic activity. This will allow LexA repressor levels to increase, shutting off the expression of the genes involved in the SOS response, and allowing pol III to resume normal proofreading.

The first evidence for the SOS system was obtained 25 years before the mechanism began to be understood. The experiment, illustrated in **Figure 10-10**, involved an analysis of the survival of UV-irradiated bacteriophage λ plated on *E. coli*. The survival of the UV-irradiated phage was markedly enhanced if the bacterial host had also been irradiated with a low UV dose. However, in that case, a much higher percentage of mutants were obtained among the surviving phages. This phenomenon (termed **UV reactivation**) is now understood in terms of the SOS system. The low UV dose to the bacteria activates the SOS error-prone replication of the damaged phage. (As you recall, bacteriophages are dependent upon *host enzymes* for bacteriophage DNA replication.)

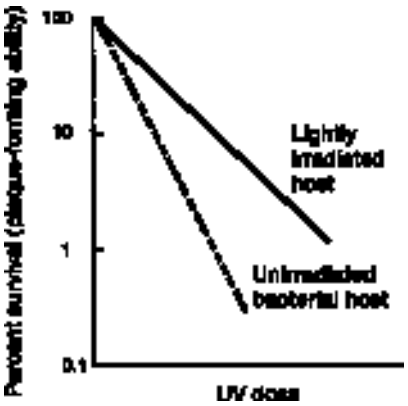


Figure 10-10 Ultraviolet (UV) reactivation of UV-light-irradiated phage λ. The dashed line shows the survival curve (for plaque-forming ability) obtained when λ phage irradiated with various doses of UV light are plated on unirradiated bacteria. The solid line represents survival of plaque-forming ability, when UV-light irradiated λ are plated on lightly irradiated bacteria.

There are examples of inducible cellular responses to other types of environmental stress, such as heat shock, oxidative DNA damage, or agents that produce 0^6 methyl guanine. In the latter case, there is an *adaptive response* in *E. coli* that increases the level of the methyltransferase enzyme 100-fold when the cells are "conditioned" by growth in the presence of agents that produce this lesion.

Future Practical Applications?

Mutations in one particular gene, the p53 gene (a 53,000 dalton molecular weight protein) are known to be responsible for more than 50% of human cancers, including breast, prostate, brain, lung, liver, pancreatic, colorectal and adrenal cancers, soft tissue sarcomas, osteosarcoma, lymphoma, and melanoma. The protein product of the normal, nonmutant p53 gene acts as a transcription factor and tumor suppressor, and is involved in regulation of the cell cycle, DNA repair, and programmed cell death. When DNA damage occurs, expression of p53 increases, leading to the arrest of the cell cycle. This gives the cell time to repair its DNA before it is replicated and passed along to a new generation of daughter cells. In cells in which extensive DNA damage has occurred, the p53 gene triggers the cell to undergo apoptosis, or programmed cell death.

If the p53 gene is mutated, it is no longer able to arrest the cell cycle in response to DNA damage. Mutations begin to accumulate in a variety of other genes, the end result being development of certain forms of cancer. As our understanding of the role of the p53 gene has increased, it has become possible to use this knowledge to predict the likelihood that a patient will develop cancer or that he will respond well to chemotherapy. We can even design a gene therapy protocol that may be effective against his particular form of cancer. Overexpression of the p53 gene is associated with an increased risk of developing cancer, and screening for levels of p53 expression may prove to be a valuable diagnostic tool in individuals who are at high risk of developing cancer.

Once the patient has been diagnosed with a malignancy, analysis of his levels of p53 expression may enable physicians to determine whether traditional chemo-

KEY CONCEPT

The "DNA Repair" Concept

Despite the inherent chemical stability of double-stranded DNA, from time to time it sustains damage, and, hence, its continual repair is necessary for cells to function properly.

therapy is likely to be an effective treatment. Conventional chemotherapy is largely dependent upon the ability of the anticancer drug to induce tumor regression by stimulating the tumor cells to undergo apoptosis. In patients possessing a mutant p53 gene and overexpressing the p53 protein, tumors are likely to be resistant to treatment involving conventional chemotherapy drugs.

Now that the p53 gene has been characterized, clinical trials involving its use in gene therapy have been initiated. Liver, head and neck cancer, lung cancer, and even oral cancers have been targeted in recent human gene therapy trials involving recombinant adenoviral vectors carrying the human p53 allele. Preliminary results are encouraging. In the case of gene therapy directed against head and neck cancer, for example, tumors shrank in 25 out of 30 patients. Researchers are optimistic that p53 gene therapy may offer new hope of successful treatment of a variety of cancers, particularly when used in conjunction with chemotherapy.

SUMMARY

Mutagenesis is the process by which mutations (changes in base sequence of a DNA molecule) are created. An organism carrying a mutation is called a mutant. A base substitution is a change in a single nucleotide, and can be a missense mutation (causing the change of a single amino acid) or a nonsense mutation (causing protein-synthesis termination). Base substitutions may be either transitions, e.g., Py-Pu to Py-Pu, or transversions, e.g., Py-Pu to Pu-Py. A frameshift mutation is an addition or loss of a nucleotide causing an alteration of the reading frame. A silent mutation causes no change in the phenotype. The phenotype of a conditional mutant is seen only under certain conditions. Mutations may be spontaneous (due to uncorrected replication errors or natural changes), or induced by a chemical or physical mutagen. Base-analogue mutagens are incorporated into DNA via normal base-pairing interactions during DNA replication, but in future rounds of replication they form incorrect base pairs, leading to a mutation. Chemical mutagens alter existing bases in DNA, changing their base-pairing properties or leading to error-prone DNA replication. Ultraviolet radiation can also induce error-prone replication. Intercalating agents insert between adjacent base pairs in DNA and cause frameshift mutations during DNA replication. Mutations can also be caused by transposable elements that insert into genes, or by mutator genes that, if mutated, lead to mutation in other genes. Mutational hot spots are highly mutable regions of DNA. Reversion or suppression is the process by which a mutant phenotype is returned to wild type, and may occur by a change at another site in the same gene (intragenic) or in a different gene (intergenic). The Ames test uses reversion to test for the possible mutagenic effects of chemical compounds.

Processes exist to repair the various DNA alterations that may occur in living organisms. Mismatch repair is used to replace an incorrect base that was not corrected by proofreading. This repair system recognizes a pair of non-hydrogen-bonded bases, excises the one that is incorrect, and replaces it with the proper base. A pathway involving the enzyme uracil-*N*-glycosylase is responsible for repair of cytosine deamination. Four mechanisms exist for the repair of thymine dimers. Photolyase is an enzyme that cleaves the cyclobutyl ring of the dimer when activated by visible light. The excision repair system makes a cut in the backbone on either side of the dimer, allowing the displacement and subsequent replacement of the portion of the strand containing the dimer. Recombinational repair involves postdimer reinitiation of DNA synthesis and sister-strand exchange, and prevents long delays of replication. The SOS response system also prevents delay of replication; this system allows for transdimer synthesis—at the expense of a higher rate of base misincorporation by the replication enzymes.

DRILL QUESTIONS

1. In what way does 5-bromouracil (BU) function as a mutagen? What type of mutations would occur from its use?
2. Several hundred independent missense mutants, altered in the A protein of tryptophan synthetase, have been collected. Originally, it was hoped that at least one mutant for each of the 186 amino acid positions in the protein would be found. However, fewer than 30 of the positions were represented with one or more mutants. Suggest some possibilities to explain why this set of missense mutants was so limited.
3. Since nonsense suppressors are mutant tRNA molecules, how does the cell survive loss of a needed tRNA by such a mutation?
4. Suppose a hypothetical enzyme contains 156 amino acids. (This number has no significance in the problem.) Assume amino acid 28, which is glutamic acid, is replaced in a mutant by asparagine and, as a result, all enzymic activity is lost. Suppose in this mutant protein, amino acid 76, which is asparagine, is replaced by glutamic acid and full activity of the enzyme is restored. What can you say about amino acids 28 and 76 in the normal protein?
5. Two DNA alterations occur so frequently that they are considered to be the weak points of a DNA molecule. What are these two changes?
6. Uvr^+ bacteria possess the excision-repair system. The ability of UV-irradiated T4 phage to form plaques is the same on both Uvr^+ and Uvr^- bacteria. How might you explain this fact?
7. Which of the enzymes listed below is involved in repair of *both* thymine dimers and deaminate cytosine in *E. coli?*
   (a) DNA polymerase III
   (b) Photolyase
   (c) Uracil *N*-glycosylase
   (d) DNA polymerase I
   (e) AP endonuclease

PROBLEMS

1. Which of the following amino acid substitutions would probably yield a mutant phenotype? Explain your reasoning.
   (a) Pro to His
   (b) Lys to Arg
   (c) Ile to Thr
   (d) Ile to Val
   (e) Ala to Gly
   (f) Phe to Leu
   (g) Try to His
   (h) Arg to Ser
2. Consider a bacterial gene containing 1,000 base pairs. As a result of treatment of a bacterial culture with a mutagen, mutations in this gene are recovered at a frequency of one mutant per 10^5 cells. One of these mutants is grown, and a pure culture of this mutant is obtained. This culture is then

treated with the same mutagen and revertants are found at a frequency of one per 10^5 cells. Would you expect the gene product obtained from the revertant to have the same amino acid sequence as the wild-type cell? Explain.

3. A particular mutant shows absolutely no activity. Despite an exhausting search for revertants by your colleagues, using such mutagens as 5-bromouracil, nitrous acid, and UV light, none is found. What sort of mutation do you think exists in the original mutant? Why? What mutagen would you use to try to obtain a revertant?
4. A bacterial repair system called X removes thymine dimers. You have in your bacterial collection the wild-type (X^+) and an X^- mutant. Phage, when UV-irradiated and then plated, gives a larger number of plaques on X^+ than on X^- bacteria. It has been proposed on the basis of survival curve analysis that the X enzyme is inducible. To test this proposal, UV-irradiated phage are adsorbed to both X^+ and X^- bacteria in the presence of the antibiotic chloramphenicol (which inhibits protein synthesis). No thymine dimers are removed in the X^+ cell and 50% are removed in the X^- cell. The same results were obtained in the absence of chloramphenicol.
   (a) Is X an inducible system?
   (b) Suppose 5% of the thymine dimers had been removed in the presence of chloramphenicol and 50% in its absence; how would your conclusion be changed?

CONCEPTUAL QUESTIONS

1. Consider the effects of life on Earth from the breakdown of the ozone layer (which would allow more UV light to penetrate to the Earth's surface). What might be the eventual outcome of that situation?
2. Would different types of organisms—for example, bacteria, fruit flies, frogs, and humans—be expected to sustain different types of DNA damage?
3. How might it be established whether DNA damage-repair mechanisms evolved first to repair replication errors, and only secondarily to repair damage from environmental influences, or vice versa?