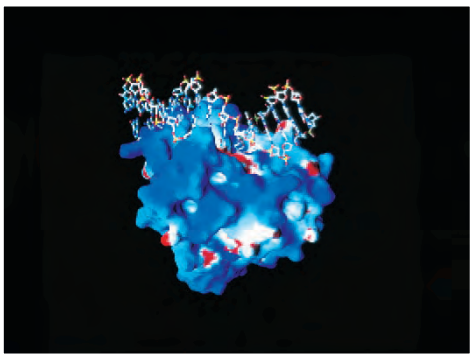
12

DNA
Its Mutation, Repair, and Recombination

STUDY OBJECTIVES

1. To look at the nature of mutation in prokaryotes 316
2. To analyze functional and structural allelism and examine the mapping of mutant sites within a gene 317
3. To verify the colinearity of gene and protein 324
4. To study mutagenesis 325
5. To investigate the processes of DNA repair and recombination 339

STUDY OUTLINE



Computer-generated space-filling model of a DNA enzyme repairing damaged DNA.

(© James King-Holmes/SPL/Photo Researchers, Inc.)

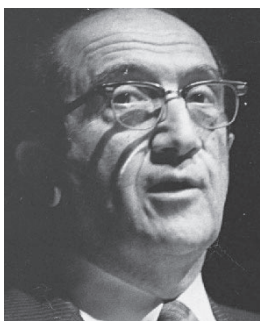
The mutation, repair, and recombination of DNA are treated together in this chapter because the three processes have much in common. The physical alteration of DNA is involved in each; repair and recombination share some of the same enzymes. We progress from mutation—the change in DNA—to repair of damaged DNA, and, finally, to recombination, the new arrangement of pieces of DNA.

MUTATION

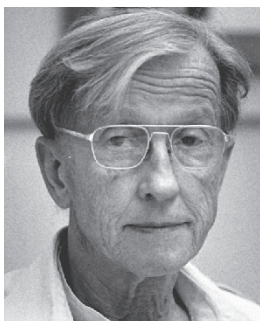
The concept of mutation (a term coined by de Vries, a rediscoverer of Mendel) is pervasive in genetics. **Mutation** is both the process by which a gene (or chromosome) changes structurally and the end result of that process. Without alternative forms of genes, the biological diversity that exists today could not have evolved. Without alternative forms of genes, it would have been virtually impossible for geneticists to determine which of an organism's characteristics are genetically controlled. Studies of mutation provided the background for our current knowledge in genetics.

Fluctuation Test

In 1943, Salvador Luria and Max Delbrück published a paper entitled "Mutations of Bacteria from Virus Sensitivity to Virus Resistance." This paper ushered in the era of bacterial genetics by demonstrating that the phenotypic variants found in bacteria are actually attributable to mutations rather than to induced physiological changes. Very little work had previously been done in bacterial genetics because of the feeling that bacteria did not have "normal" genetic systems like the systems of fruit flies and corn. Rather, bacteria were believed to respond to environmental change by physiological adaptation, a non-Darwinian view. As Luria said, bacteriology remained "the last stronghold of Lamarckism" (the belief that acquired characteristics are inherited).



Salvador E. Luria (1912–1991). (Courtesy of Dr. S. E. Luria.)



Max Delbrück (1906–1981). (Courtesy of Dr. Max Delbrück.)

What Causes Genetic Variation?

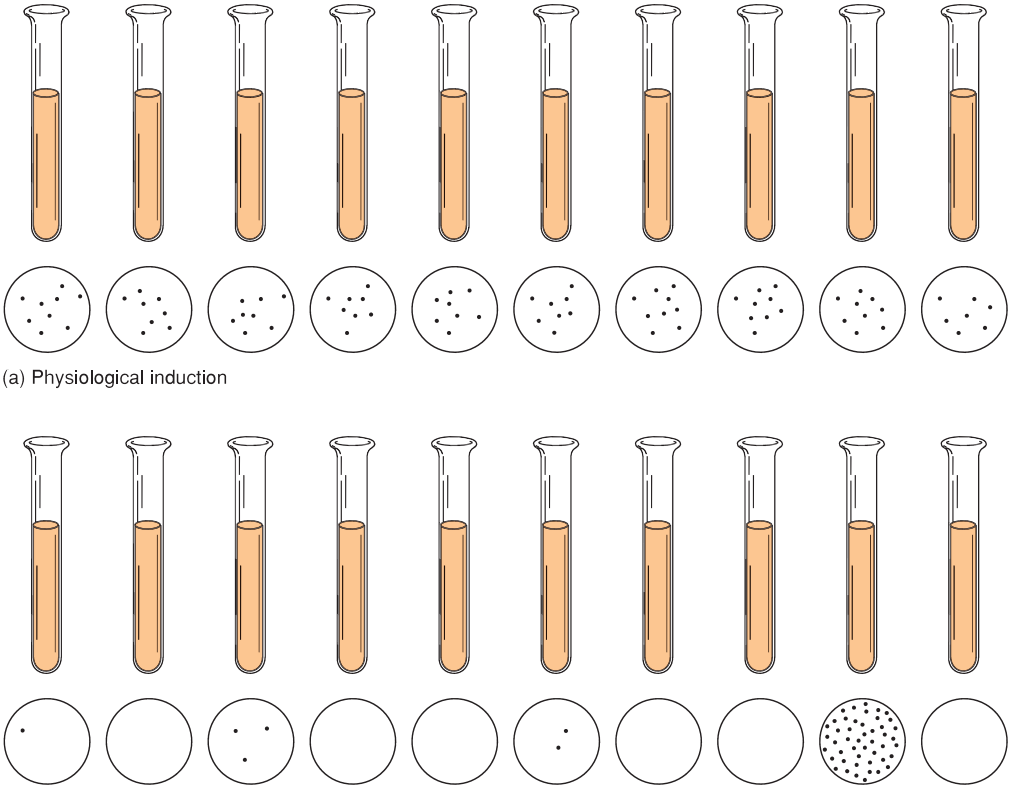
Luria and Delbrück studied the Ton^r (phage T1-resistant) mutants of a normal Ton^s (phage T1-sensitive) *Escherichia coli* strain. They used an enrichment experiment, as described in chapter 7, wherein a petri plate is spread with *E. coli* bacteria and T1 phages. Normally, no bacterial colonies grow on the plate: all the bacteria are lysed. However, if one of the bacterial cells is resistant to T1 phages, it produces a bacterial colony, and all descendants of this colony are T1 resistant. There are two possible explanations for the appearance of T1-resistant colonies:

1. Any *E. coli* cell may be induced to be resistant to phage T1, but only a very small number actually are. That is, all cells are genetically identical, each with a very low probability of exhibiting resistance in the presence of T1 phages. When resistance is induced, the cell and its progeny remain resistant.
2. In the culture, a small number of *E. coli* cells exist that are already resistant to phage T1; in the presence of phage T1, only these cells survive.

If the presumed rates of physiological induction and mutation are the same, determining which of the two mechanisms is operating is difficult. Luria and Delbrück, however, developed a means of distinguishing between these mechanisms. They reasoned as follows: If T1 resistance was physiologically induced, the relative frequency of resistant *E. coli* cells in a culture of the normal (Ton^s) strain should be a constant, independent of the number of cells in the culture or the length of time that the culture has been growing. If resistance was due to random mutation, the frequency of mutant (Ton^r) cells would depend on when the mutations occurred. In other words, the appearance of a mutant cell would be a random event. If a mutation occurs early in the growth of the culture, then many cells descend from the mutant cell, and therefore many resistant colonies develop. If the mutation does not occur until late in the growth of the culture, then the subsequent number of mutant cells is few. Thus, if the mutation hypothesis is correct, there should be considerable fluctuation from culture to culture in the number of resistant cells present (fig. 12.1).

Results of the Fluctuation Test

To distinguish between these hypotheses, Luria and Delbrück developed what is known as the **fluctuation test.** They counted the mutants both in small ("individual") cultures and in subsamples from a single large ("bulk") culture. All subsamples from a bulk culture should have the same number of resistant cells, differing only because of



(a) Physiological induction

(b) Random mutation

Figure 12.1 Occurrence of *E. coli* Ton^r colonies in Ton^s cultures. Ten cultures of *E. coli* cells were grown from a standard inoculum in separate test tubes in the absence of phage T1, then spread on petri plates in the presence of phage T1. The resistant cells grow into colonies on the plates. We expect a uniform distribution of resistant cells if the physiological induction hypothesis is correct (*a*) or a great fluctuation in the number of resistant cells if the random mutation hypothesis is correct (*b*).

random sampling error. If, however, mutation occurs, the number of resistant cells among the individual cultures should vary considerably from culture to culture; the number would be related to the time that the mutation occurred during the growth of each culture. If mutation arose early, there would be many resistant cells. If it arose late, there would be relatively few resistant cells. Under physiological induction, the distribution of resistant colonies should not differ between the individual and bulk cultures.

Luria and Delbrück inoculated twenty individual cultures and one bulk culture with *E. coli* cells and incubated them in the absence of phage T1. Each individual culture was then spread out on a petri plate containing a very high concentration of T1 phages; ten subsamples from the bulk culture were plated in the same way. We can see from the results (table 12.1) that there was minimal variation in the number of resistant cells among the bulk culture subsamples but a very large amount of variation, as predicted for random mutation, among the individual cultures.

If bacteria have "normal" genetic systems that undergo mutation, bacteria could then be used, along with higher organisms, to answer genetic questions. As we have pointed out, the modern era of molecular genetics began with the use of prokaryotic and viral systems in genetic research. In the next section, we turn our attention to several basic questions about the gene, questions whose answers were found in several instances only because prokaryotic systems were available.

Genetic Fine Structure

How do we determine the relationship among several mutations that cause the same phenotypic change? What are the smallest units of DNA capable of mutation and recombination? Are the gene and its protein product colinear? The answers to the latter two questions are important from a historical perspective. The answer to the first question is relevant to our current understanding of genetics.

Table 12.1 Results from the Luria and Delbrück Fluctuation Test

Individual Cultures*		Samples from Bulk Culture*	
Culture Number	Ton^r Colonies Found	Sample Number	Ton^r Colonies Found
1	1	1	14
2	0	2	15
3	3	3	13
4	0	4	21
5	0	5	15
6	5	6	14
7	0	7	26
8	5	8	16
9	0	9	20
10	6	10	13
11	107		
12	0		
13	0		
14	0		
15	1		
16	0		
17	0		
18	64		
19	0		
20	35		
Mean ($\bar{m}$)	11.4		16.7
Standard deviation	27.4		4.3

Source: From E. Luria and M. Delbrück, *Genetics,* 28: 491. Copyright © 1943 Genetics Society of America.

* Each culture and sample was 0.2 ml and contained about 2×10^7 *E. coli* cells.

Complementation

If two recessive mutations arise independently and both have the same phenotype, how do we know whether they are both mutations of the same gene? That is, how do we know whether they are alleles? To answer this question, we must construct a heterozygote and determine the **complementation** between the two mutations. A heterozygote with two mutations of the same gene will produce only mutant messenger RNAs, which result in mutant enzymes (fig. 12.2*a*). If, however, the mutations are not allelic, the gamete from the a_1 parent will also contain an a_2^+ allele, and the gamete from the a_2 parent will also contain the a_1^+ allele (fig. 12.2*b*). If the two mutant genes are truly alleles, then the phenotype of the heterozygote should be mutant. If, however, the two mutant genes are nonallelic, then the a_1 mutant will have contributed the wild-type allele at the A_2 locus, and the a_2 mutant will have contributed the wild-type allele at the A_1 locus to the heterozygote. Thus, the two mutations will complement each other and produce the wild-type. Mutations that fail to complement each other are termed **functional alleles.** The test for defining alleles strictly on this basis of functionality is termed the ***cis-trans* complementation test.**

There are two different configurations in which a heterozygous double mutant of functional alleles can form (fig. 12.3). In the *cis-trans* complementation test, only the *trans* configuration is used to determine whether the two mutations were allelic. In reality, the *cis* configuration is not tested; it is the conceptual control, in which wild-type activity (with recessive mutations) is always expected. The test is thus sometimes simply called a *trans* test. Functional alleles produce a wild-type phenotype in the *cis* configuration but a mutant phenotype in the *trans* configuration. This difference in phenotypes is called a *cis-trans* position effect.

From the terms *cis* and *trans,* Seymour Benzer coined the term **cistron** for the smallest genetic unit (length of genetic material) that exhibits a *cis-trans* position effect. We thus have a new word for the gene, one in which function is more explicit. We have, in essence, refined Beadle and Tatum's one-gene-one-enzyme hypothesis to a more accurate one-cistron-one-polypeptide concept. The cistron is the smallest unit that codes for a messenger RNA that is then translated into a single polypeptide or expressed directly (transfer RNA or ribosomal RNA).

From functional alleles, we can go one step further in recombinational analysis by determining whether two allelic mutations occur at exactly the same place in the cistron. In other words, when two mutations prove to be functional alleles, are they also **structural alleles?** The methods used to analyze complementation can be used here also. Crosses are carried out to form a mutant heterozygote (*trans* configuration) whose offspring are then tested for recombination between the two mutational sites. If no recombination occurs, then the two alleles probably contain the same

Seymour Benzer (1921–). (Courtesy of Dr. Seymour Benzer, 1970.)

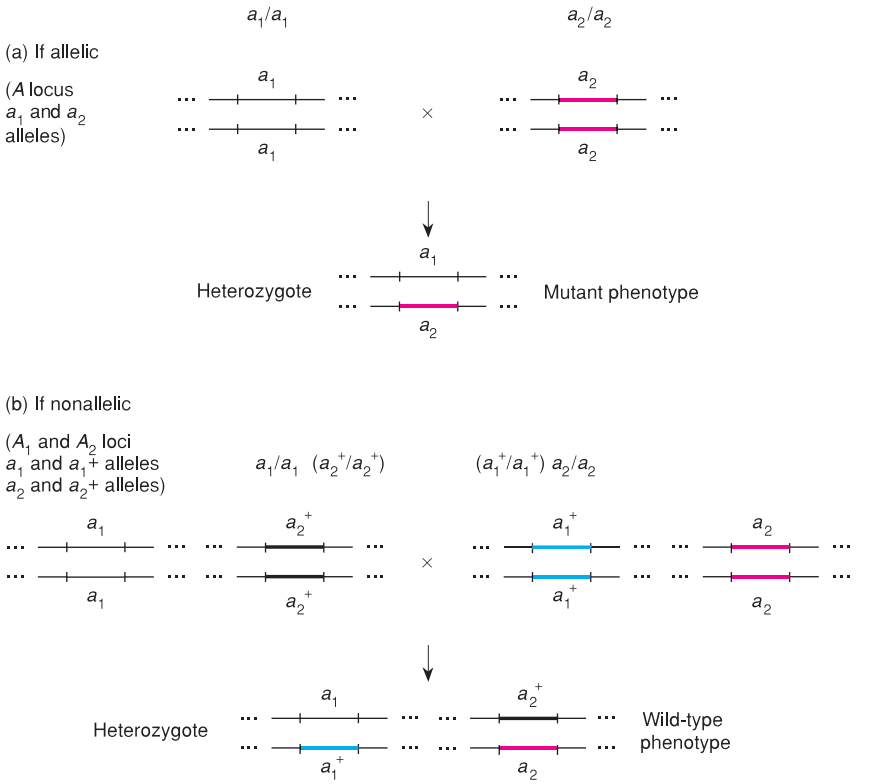


Figure 12.2 The complementation test defines allelism. Are two mutations (a_1, a_2) allelic if they affect the same trait? To find out, mutant homozygotes are crossed to form a heterozygote. (*a*) If the mutations are allelic, then both copies of the gene in the heterozygote are mutant, resulting in the mutant phenotype. (*b*) If the mutations are nonallelic, then there is a wild-type allele of each gene present in the heterozygote, resulting in the wild-type phenotype. (The two loci need not be on the same chromosome.)

structural change (involving the same base pairs) and are thus structural alleles. If a small amount of recombination occurs that generates wild-type offspring, then the two alleles are not mutations at the same point (fig. 12.4). Alleles that were functional but not structural were first termed **pseudoalleles** because it was believed that loci were made up of subloci. Fine-structure analysis led to the understanding that a locus is a length of genetic material divisible by recombination rather than a "bead on a string."

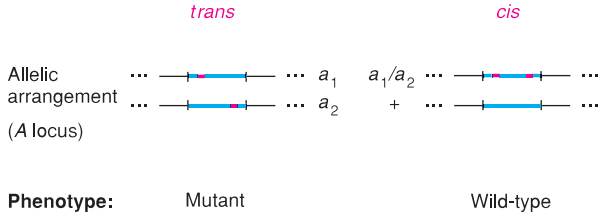


Figure 12.3 A heterozygote of two recessive mutations can have either the *trans* or *cis* arrangement. In the *trans* position, functional alleles produce a mutant phenotype. (*Red marks* represent mutant lesions.) In the *cis* position, functional alleles produce a wild-type phenotype. The *cis-trans* position effect thus reveals functional alleles.

Eye-color mutants of *Drosophila melanogaster* can be studied by complementational analysis. The white-eye locus has a series of alleles producing varying shades of red. This locus is sex linked, at about map position 3.0 on the X chromosome. (Several other eye-color loci on the X chromosome are not relevant to this cross—e.g., prune and ruby.) If an apricot-eyed female is mated with a white-eyed male, the female offspring are all heterozygous and have mutant light-colored eyes (fig. 12.5). Thus, apricot and white are functional alleles: they do not complement (table 12.2). To determine whether apricot and white are structural alleles, light-eyed females are crossed with white-eyed males, and the offspring are observed for the presence of wild-type or light-eyed males. Though their rate of appearance is less than 0.001%, this is significantly above the background mutation rate. The conclusion is that apricot and white are functional, but not structural, alleles.

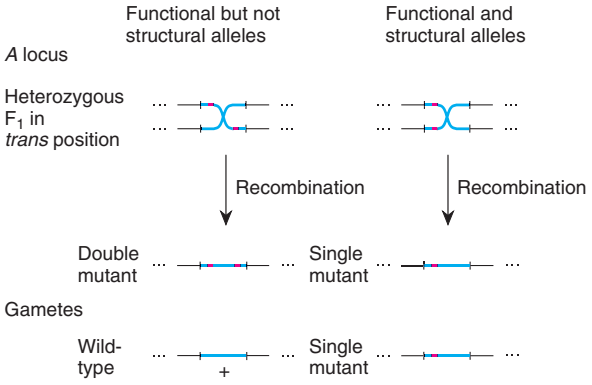


Figure 12.4 Functional alleles may or may not be structurally allelic. (*Red marks* represent mutant sites.) Functional alleles that are not also structural alleles can recombine between the mutant sites, resulting in occasional wild-type (and double mutant) offspring. Structural alleles (which are also always functional alleles) are defective at the same base pairs and cannot form either wild-type or double mutant offspring by recombination.

Fine-Structure Mapping

After Beadle and Tatum established in 1941 that a gene controls the production of an enzyme that then controls a step in a biochemical pathway, Benzer used analytical techniques to dissect the fine structure of the gene. Fine-structure mapping means examining the size and number of sites within a gene that are capable of mutation and recombination. In the late 1950s, when biochemical techniques were not yet available for DNA sequencing, Benzer used classical recombinational and mutational techniques with bacterial viruses to provide reasonable estimates on the details of fine structure and to give insight into the nature of the gene. He coined the terms **muton** for the smallest mutable site and **recon** for the smallest unit of recombination. It is now known that both muton and recon are a single base pair.

Before Benzer's work, genes were thought of as beads on a string. The very low rate of recombination between

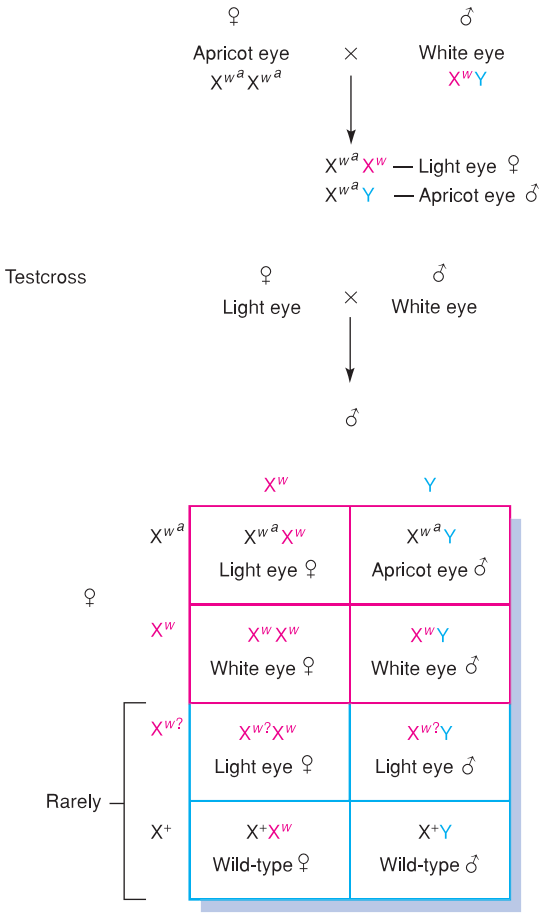


Figure 12.5 Crosses demonstrating that apricot and white eyes are functional, but not structural, alleles in *Drosophila*. Light-eyed females are heterozygous for both alleles. When testcrossed, they produce occasional offspring that are wild-type (X^+ allele) or light-eyed ($X^{w?}$ allele). This indicates a crossover between the two mutant sites (white and apricot) in the heterozygous females, producing, reciprocally, an allele with both mutational sites and the wild-type.

Table 12.2 Complementation Matrix of X-Linked *Drosophila* Eye-Color Mutants

	white	prune	apricot	buff	cherry	eosin	ruby
white (*w*)	−	+	−	−	−	−	+
prune (*pn*)		−	+	+	+	+	+
apricot (w^a)			−	−	−	−	+
buff (w^{bf})				−	−	−	+
cherry (w^{ch})					−	−	+
eosin (w^e)						−	+
ruby (*rb*)							−

Note: Plus sign indicates that female offspring are wild-type; minus sign indicates that they are mutant.

sites within a gene hampered the analysis of mutational sites within a gene by means of recombination. If two mutant genes are functional alleles (involving different sites on the same gene), a distinct probability exists that we will get both mutant sites (and both wild-type sites) on the same chromosome by recombination (see fig. 12.4); but, in view of the very short distances within a gene, this probability is very low. Although it certainly seemed desirable to map sites within the gene, the problem of finding an organism that would allow fine-structure analysis remained until Benzer decided to use phage T4.

***r* II Screening Techniques.** Benzer used the T4 bacteriophage because of the growth potential of phages, in which a generation takes about an hour and the increase in numbers per generation is about a hundredfold. Actually, any prokaryote or virus should suffice, but Benzer made use of other unique screening properties of the phage that made it possible to recognize one particular mutant in about a billion phages. Benzer used *r*II mutants of T4. These mutants produce large, smooth-edged plaques on *E. coli,* whereas the wild-type produces smaller plaques whose edges are not as smooth (see fig. 7.7).

The screening system that Benzer employed made use of the fact that *r*II mutants do not grow on *E. coli* strain K12, whereas the wild-type can. The normal host strain, *E. coli* B, allows growth of both the wild-type and *r*II mutants. Thus, various mutants can be crossed by mixed infection of *E. coli* B cells, and Benzer could screen for wild-type recombinants by plating the resultant progeny phages on *E. coli* K12 (fig. 12.6), on which only a wild-type recombinant produces a plaque. It is possible to detect about one recombinant in a billion phages, all in an afternoon's work. This ability to detect recombinants occurring at such a low level of frequency allowed Benzer to see recombinational events occurring very close together on the DNA, events that would normally occur at a frequency too low to detect in fruit flies or corn.

Benzer sought to map the number of sites subject to recombination and mutation within the *r*II region of T4. He began by isolating independently derived *r*II mutants and crossing them among themselves. The first thing he found was that the *r*II region was composed of two cistrons; almost all of the mutations belonged to one of two **complementation groups.** The *A*-cistron mutations would not complement each other but would complement the mutations of the *B* cistron. The exceptions were mutations that seemed to belong to both cistrons. These mutations were soon found to be deletions in which part of each cistron was missing (table 12.3).

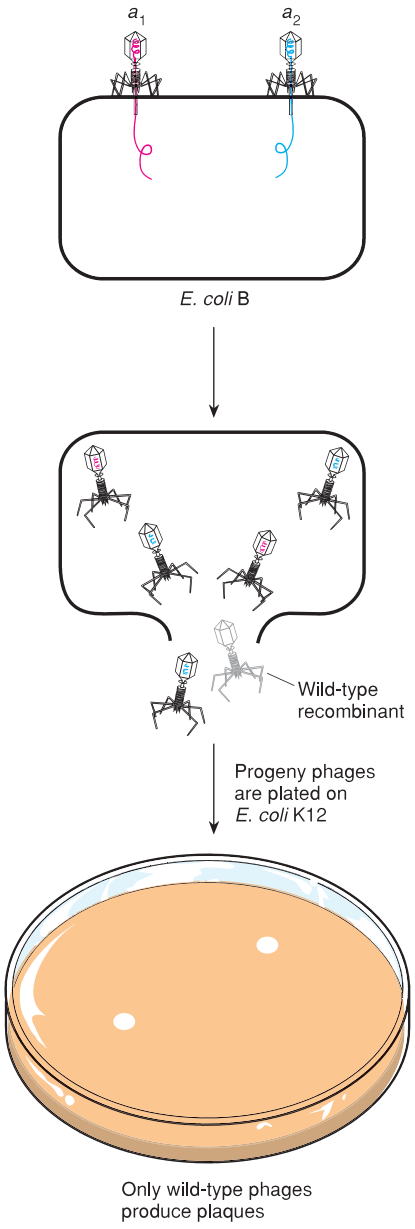


Figure 12.6 Using *E. coli* K12 and B strains to screen for recombination at the *r*II locus of phage T4. Two *r*II mutants are crossed by infecting the same B-strain bacteria with both phages. The offspring are plated on a lawn of K12 bacteria in which only wild-type phages can grow. The technique thus selects only wild-type recombinants.

Deletion Mapping. As the number of independently isolated mutations of the *A* and *B* cistrons increased, it became obvious that to make every possible pairwise cross would entail millions of crosses. To overcome this problem, Benzer isolated mutants that had partial or complete deletions of each cistron. Deletion mutations were easy to discover because they acted like structural alleles to alleles that were not themselves structurally allelic. In other words, if mutations *a, b,* and *c* are functional—but not

structural—alleles of each other, and mutation *d* is a structural allele to *a, b,* and *c,* then *d* must contain a deletion of the bases mutated in *a, b,* and *c.* Once a sequence of deletion mutations covering the *A* and *B* cistrons was isolated, a minimal number of crosses was required to localize a new mutation to a portion of one of the cistrons. A second series of smaller deletions within each region was then isolated, further localizing the mutation (fig. 12.7).

Next, each new mutant was crossed with each of the other mutants isolated in its subregion to localize the relative position of the new mutation. If the mutation was structurally allelic to a previously isolated mutation, it was scored as an independent isolation of the same mutation. If it was not a structural allele to any of the known mutations of the subregion, it was added as a new mutation point. The exact position of each new mutation within the region was determined by the relative frequency of recombination between it and the known mutations of this region (see chapter 7). Benzer eventually isolated about 350 mutations from eighty different subregions defined by deletion mutations. An abbreviated map is shown in figure 12.8.

What conclusions did Benzer draw from his work? First, he concluded that since all of the mutations in both *r*II cistrons can be ordered in a linear fashion, the original Watson-Crick model of DNA as a linear molecule was correct. Second, he concluded that reasonable inroads had been made toward saturating the map, localizing at least one mutation at every mutable site. Benzer reasoned that since many sites were represented by only one mutation, some sites must occur that were represented by zero mutations (i.e., not yet represented by a mutation). Since he had mapped about 350 sites, he calculated that there

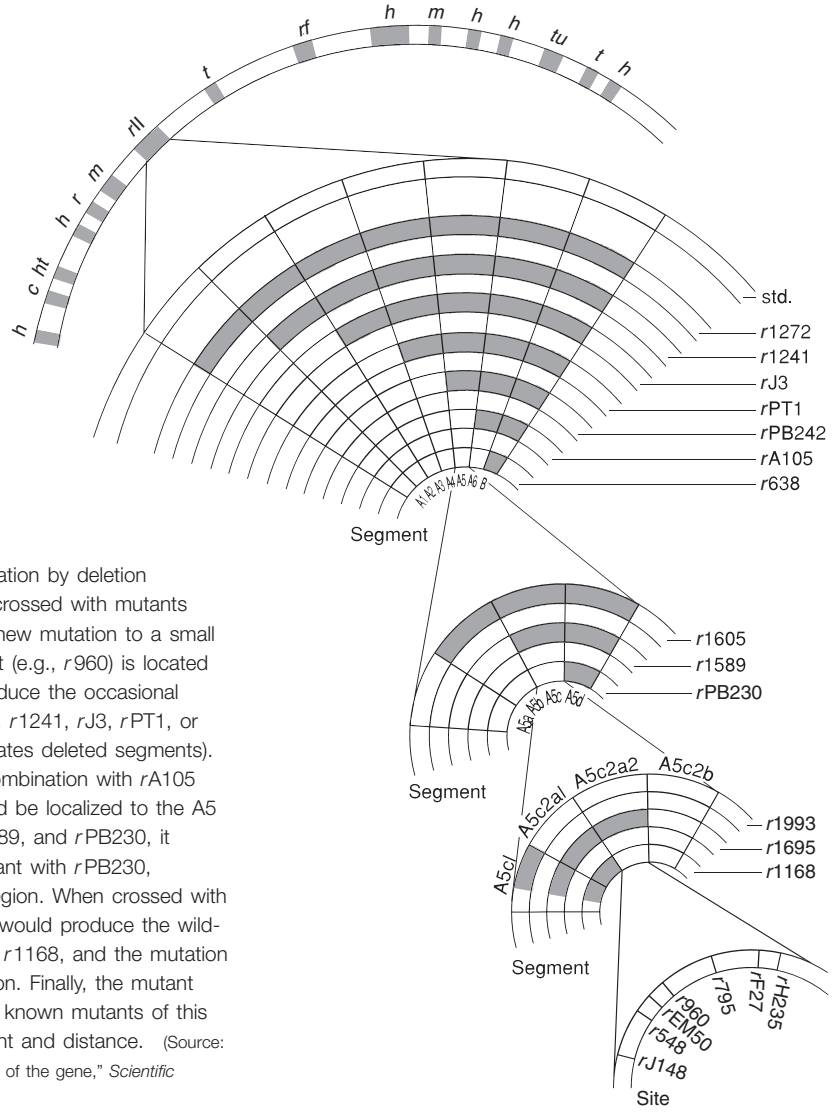


Figure 12.7 Localization of an *r*II mutation by deletion mapping. Newly isolated mutants are crossed with mutants with selected deletions to localize the new mutation to a small region of the cistron. If the new mutant (e.g., *r*960) is located in the A5c2a2 region, it would not produce the occasional wild-type by recombination with *r*1272, *r*1241, *r*J3, *r*PT1, or *r*PB242 (the *solid* part of the bar indicates deleted segments). It would produce the wild-type by recombination with *r*A105 and *r*638, and thus the mutation would be localized to the A5 region. When crossed with *r*1605, *r*1589, and *r*PB230, it would produce only the rare recombinant with *r*PB230, indicating the mutation is in the A5c region. When crossed with *r*1993, *r*1695, and *r*1168, the mutant would produce the wild-type by recombination with *r*1993 and *r*1168, and the mutation would be localized to the A5c2a2 region. Finally, the mutant would be crossed pairwise with all the known mutants of this region to determine relative arrangement and distance. (Source: Data from Seymour Benzer, "The fine structure of the gene," *Scientific American,* 206: 70–84, January 1962.)

Table 12.3 **Complementation Matrix of Ten *r*II Mutants**

	1	2	3	4	5	6	7	8	9	10
1	–	–	+	–	–	–	+	–	–	–
2		–	+	–	–	–	+	–	–	–
3			–	–	+	+	–	+	–	+
4				–	–	–	–	–	–	–
5					–	–	+	–	–	–
6						–	+	–	–	–
7							–	+	–	+
8								–	–	–
9									–	–
10										–

Note: Plus sign indicates complementation; minus sign indicates no complementation. The two cistrons are arbitrarily designated *A* and *B*. Mutants 4 and 9 must be deletions that cover parts of both cistrons. Alleles: *A* cistron: 1, 2, 4, 5, 6, 8, 9, 10; *B* cistron: 3, 4, 7, 9.

were at least another 100 sites still undetected by mutation. We now know that 450 sites is an underestimate. However, since the protein products of these cistrons were not isolated, there were no independent estimates of the number of nucleotides in these cistrons (number of amino acids times three nucleotides per codon). Thus, although Benzer had not saturated the map with mutations, he certainly had made respectable progress in dissecting the gene and demonstrating that it was not an indivisible unit, a "bead on a string."

Hot Spots. Benzer also looked into the lack of uniformity in the occurrence of mutations (note two major "hot spots" at B4 and A6c of fig. 12.8). Presuming that all base pairs are either AT or GC, this lack of uniformity was unexpected. Benzer suggested that spontaneous mutation is not just a function of the base pair itself, but is affected by the surrounding bases as well. This concept still holds.

To recapitulate, Benzer's work supports the model of the gene as a linear arrangement of DNA whose nucleotides are the smallest units of mutation. The link between any adjacent nucleotides can break in the recombinational process. The smallest functional unit, determined by a complementation test, is the cistron. Mutagenesis is not uniform throughout the cistron, but may depend on the particular arrangement of bases in a given region.

Intra-Allelic Complementation

Benzer warned that certainty is elusive in the complementation test because sometimes two mutations of the same functional unit (cistron) can result in partial activity. The

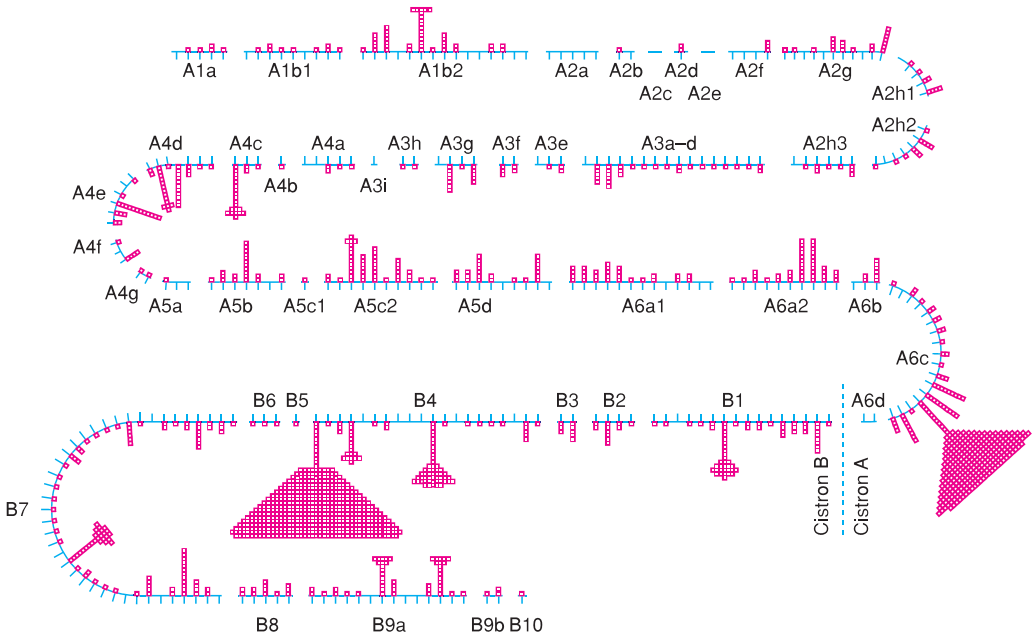


Figure 12.8 Abbreviated map of spontaneous mutations of the *A* and *B* cistrons of the *r*II region of T4. Each square represents one independently isolated mutation. Note the "hot spots" at A6c and B4. (From Seymour Benzer, "On the topography of the genetic fine structure", *Proceedings of the National Academy of Sciences USA* 47:403–15, 1961. Reprinted by permission.)

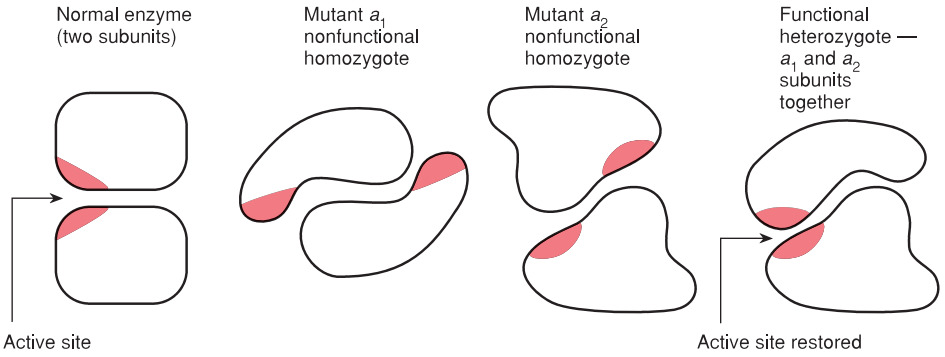


Figure 12.9 Intra-allelic complementation. With certain mutations, it is possible to get enzymatic activity in a heterozygote for two nonfunctional alleles, if the two polypeptides form a functional enzyme. (Active site is shown in *color.*)

problem can be traced to the interactions of subunits at the polypeptide level. Some proteins are made up of subunits, and it is possible that certain mutant combinations produce subunits that interact to restore the enzymatic function of the protein (fig. 12.9). This phenomenon is known as **intra-allelic complementation.** With this in mind, geneticists routinely use the complementation test to determine functional relationships among mutations.

Colinearity

Next we look at the colinearity of the gene and the polypeptide. Benzer's work established that the gene was a linear entity, as Watson and Crick had proposed. However, Benzer could not demonstrate the colinearity of the gene and its protein product. To do this, it is necessary to show that for every mutational change in the DNA, a corresponding change takes place in the protein product of the gene. Colinearity would be established by showing that nucleotide and amino acid changes occurred in a linear fashion and in the same order in the protein and in the cistron.

Ideally, Benzer himself might have solved the colinearity issue. He was halfway there, with his 350 or so isolated mutations of phage T4. However, Benzer did not have a protein product to analyze; no mutant protein had been isolated from *r*II mutants. In the midst of competition to find just the right system, Charles Yanofsky of Stanford University and his colleagues emerged in the mid-1960s with the required proof, showing that the order of a polypeptide's amino acids corresponded to the nucleotide sequence in the gene that specified it. Yanofsky's success rested with his choice of an amenable system, one using the enzymes from a biochemical pathway.

Yanofsky did his research on the tryptophan biosynthetic pathway in *E. coli.* The last enzyme in the pathway, tryptophan synthetase, catalyzes the reaction of indole-3-glycerol-phosphate plus serine to tryptophan and 3-phosphoglyceraldehyde. The enzyme itself is made of four subunits specified by two separate cistrons, with each polypeptide present twice.

Yanofsky and his colleagues concentrated on the *A* subunit. They mapped *A*-cistron mutations with transduction (see chapter 7) using the transducing phage P1. They first tested each new mutant against a series of deletion mutants to establish the region where the mutation was. Then they crossed mutants for a particular region among themselves to establish relative positions and distances.

The protein products of the bacterial genes were isolated using electrophoresis and chromatography to establish the fingerprint patterns of the proteins (see chapter 11). Assuming a single mutation, a comparison of the mutant and the wild-type fingerprints would show a difference of just one polypeptide spot (fig. 12.10), avoiding the need to sequence the entire protein. The mutant amino acid was

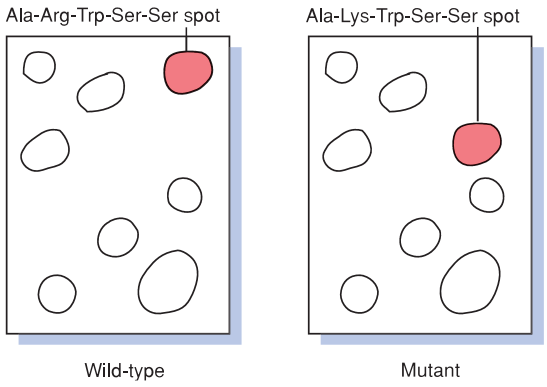


Figure 12.10 Difference in "fingerprints" between mutant and wild-type polypeptide digests. The single spot that differs in the mutant can be isolated and sequenced, eliminating the need to sequence the whole protein.

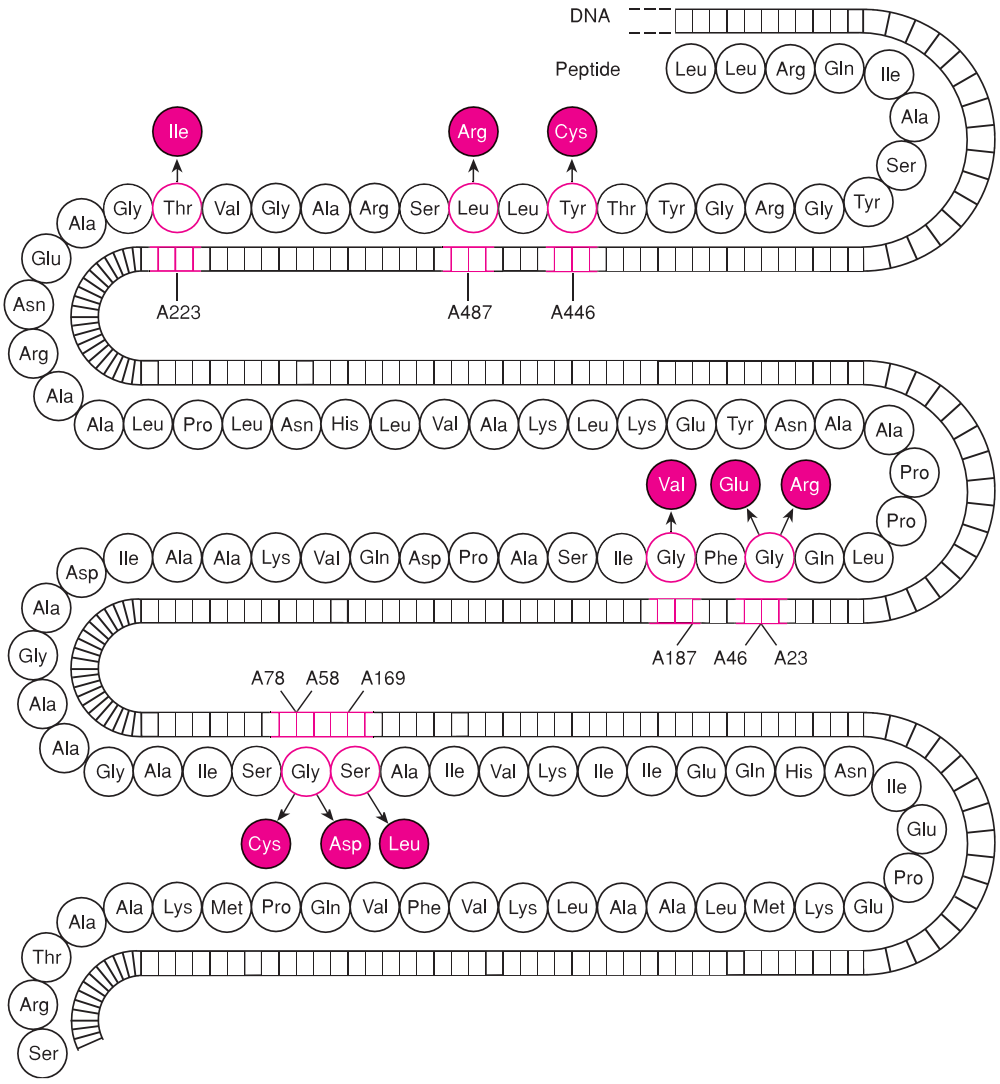


Figure 12.11 Amino acid sequence of the carboxyl terminal end of the tryptophan synthetase *A* protein and its DNA. Mutations are shown on the DNA (e.g., A446), as are the changed amino acids of those mutations (in *color*). DNA and protein changes are colinear.

identified by analysis of just this one spot. Figure 12.11 shows the details of nucleotide and amino acid changes for nine of the mutations in this 267-amino acid protein.

We can see from this figure that nine mutations in the linear *A* cistron of tryptophan synthetase are colinear with nine amino acid changes in the protein itself. In two cases, two mutations mapped so close as to be almost indistinguishable. In both cases, the two mutations proved to be in the same codon: the same amino acid position was altered in each (A23–A46, A58–A78). Thus, exactly as predicted and expected, colinearity exists between gene and protein. Brenner and his colleagues, using head-protein mutants of phage T4, independently confirmed this work at the same time.

Spontaneous Versus Induced Mutation

H. J. Muller won the Nobel Prize for demonstrating that X rays can cause mutations. This work was published in 1927 in a paper entitled "Artificial Transmutation of the Gene." At about the same time, L. J. Stadler induced mutations in barley with X rays. The basic impetus for their work was the fact that mutations occur so infrequently that genetic research was hampered by the inability to

Hermann J. Muller (1890–1967). (Courtesy of National Academy of Sciences.)

Lewis J. Stadler (1896–1954). (*Genetics,* 41, 1956: frontispiece.)

obtain mutants. Muller exposed flies to varying doses of X rays and then observed their progeny. He came to several conclusions. First, X rays greatly increased the occurrence of mutations. Second, the inheritance patterns of X-ray-induced mutations and the resulting phenotypes of organisms were similar to those that resulted from natural, or "spontaneous," mutations.

Mutation Rates

The **mutation rate** is the number of mutations that arise per cell division in bacteria and single-celled organisms, or the number of mutations that arise per gamete in higher organisms. Mutation rates vary tremendously depending upon the length of genetic material, the kind of mutation, and other factors. Luria and Delbrück, for example, found that in *E. coli* the mutation rate per cell division of Ton^s to Ton^r was 3×10^{-8}, whereas the mutation rate of the wild-type to the histidine-requiring phenotype (His^+ to His^-) was 2×10^{-6}. The rate of **reversion** (return of the mutant to the wild-type) was 7.5×10^{-9}. The mutation and reversion rates differ because many different mutations can cause the His phenotype, whereas reversion requires specific, and hence less probable, changes to correct the His phenotype back to the wild-type. The lethal mutation rate in *Drosophila* is about 1×10^{-2} per gamete for the total genome. This number is relatively large because, as with His, many different mutations produce the same phenotype (lethality, in this case).

Point Mutations

The mutations of primary concern in this chapter are **point mutations,** which consist of single changes in the nucleotide sequence. (In chapter 8 we discussed chromosomal mutations, changes in the number and visible structures of chromosomes.) If the change is a replacement of some kind, then a new codon is created. In many cases, this new codon, upon translation, results in a new amino acid. As discussed in chapter 11, one of the outcomes of redundancy in the genetic code is partial protection of the cell from the effects of mutation; common amino acids have the most codons, similar amino acids have similar codons, and the wobble position of the codon is the least important position in translation. However, when base changes result in new amino acids, new proteins appear. These new proteins can alter the morphology or physiology of the organism and result in phenotypic novelty or lethality.

Frameshift Mutation

A point mutation may consist of replacement, addition, or deletion of a base (fig. 12.12). Point mutations that add

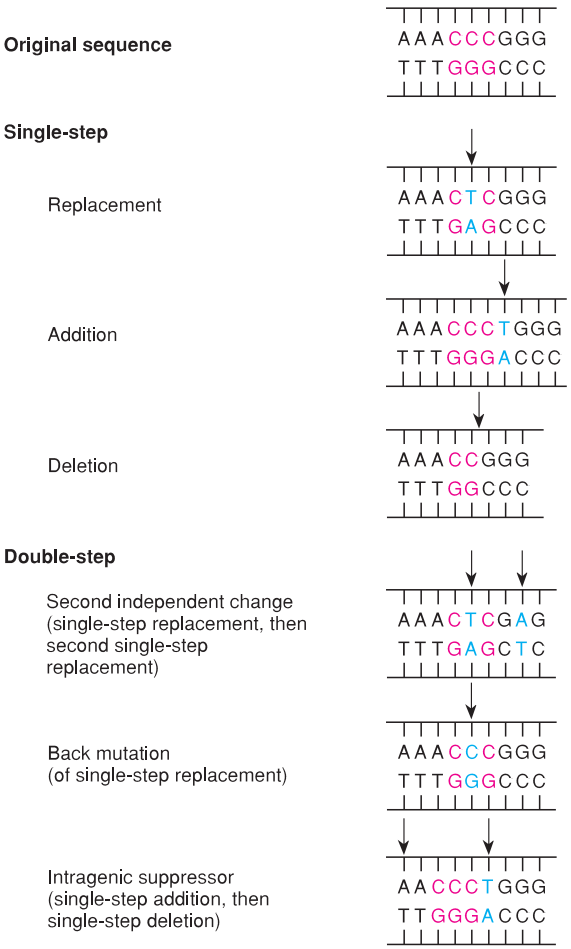


Figure 12.12 Types of DNA point mutations. Single-step changes are replacements, additions, or deletions. A second point mutation in the same gene can result either in a double mutation, reversion to the original, or intragenic suppression. In this case, intragenic suppression is illustrated by the addition of one base followed by the nearby deletion of a different base.

or subtract a base are, potentially, the most devastating in their effects on the cell or organism because they change the reading frame of a gene from the site of mutation onward (fig. 12.13). A frameshift mutation causes two problems. First, all the codons from the frameshift on will be different and thus yield (most probably) a useless protein. Second, stop-signal information will be misread. One of the new codons may be a nonsense codon, which causes translation to stop prematurely. Or, if the translation apparatus reaches the original nonsense codon, it is no longer recognized as such because it is in a different reading frame, and therefore, the translation process continues beyond the end of the gene.

Back Mutation and Suppression

A second point mutation in the same gene can have one of three possible effects (see fig. 12.12). First, the mutation can result in either another mutant codon or in one codon that has experienced two changes. Second, if the change is at the same site, the original sequence can be returned, an effect known as **back mutation:** the gene then becomes a revertant, with its original function restored. Third, **intragenic suppression** can take place. Intragenic suppression occurs when a second mutation in the same gene masks the occurrence of the original mutation without actually restoring the original sequence. The new sequence is a double mutation that appears to have the original (unmutated) phenotype. In figure 12.12, a T addition is followed by an A deletion that substitutes the AACCCT sequence for the original AAACCC. These sequences, when transcribed (UUGGGA, UUUGGG), are codons for leucine-glycine and phenylalanine-glycine, respectively. Intragenic suppression occurs whether the new codons are for different amino acids or the same amino acids, as long as the phenotype of the organism is reverted approximately to the original. Suppressed mutations can be distinguished from true back mutations either by subtle differences in phenotype, by genetic crosses, by changes in the amino acid sequence of a protein, or by DNA sequencing.

Conditional Lethality

A class of mutants that has been very useful to geneticists is the conditional-lethal mutant, a mutant that is lethal under one set of circumstances but not under another set. **Nutritional-requirement mutants** are good examples (see chapter 7). **Temperature-sensitive mutants** are conditional-lethal mutants that have made it possible for geneticists to work with genes that control vital functions of the cell, such as DNA synthesis. Many temperature-sensitive mutants are completely normal at 25° C but cannot synthesize DNA at 42° C. Presumably, temperature-sensitive mutations result in enzymes with amino acid substitutions that cause protein denaturation to occur at temperatures above normal. Thus, the enzyme has normal function at 25° C, the **permissive temperature,** but is nonfunctional at 42° C, the **restrictive temperature.**

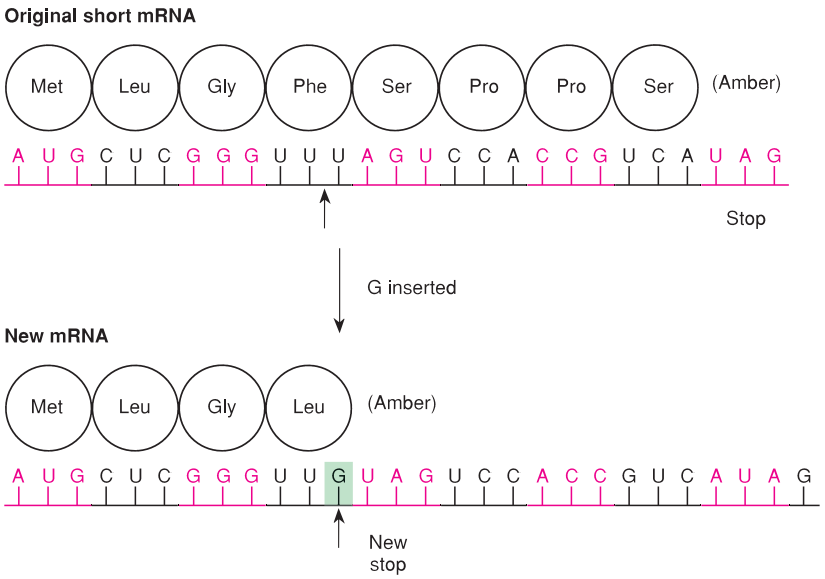


Figure 12.13 Possible effects of a frameshift mutation. The insertion of a single base results in the creation of a new stop sequence (*amber*). The result will be premature termination of translation.

The interesting thing about most conditional-lethal mutants of *E. coli* that cannot synthesize DNA at the restrictive temperature is that they have a completely normal DNA polymerase I. From this information, we infer that polymerase I is not the enzyme *E. coli* normally uses for DNA replication. When an organism with a conditional mutation of polymerase I was isolated, it was able to replicate its DNA normally, but unable to repair damage to the DNA. This led to the conclusion that polymerase I is primarily involved in repair rather than replication of DNA. Conditional-lethal mutants thus allow genetic analysis on genes otherwise impossible to study.

Spontaneous Mutagenesis

Watson and Crick originally suggested that mutation could occur spontaneously during DNA replication if pairing errors occurred. If a base of the DNA underwent a proton shift into one of its rare tautomeric forms **(tautomeric shift)** during the replication process, an inappropriate pairing of bases would occur. Normally, adenine and cytosine are in the amino (NH_2) form. Their tautomeric shifts are to the imino (NH) form. Similarly, guanine and thymine go from a keto (C=O) form to an enol (COH) form (fig. 12.14). Table 12.4 shows the new

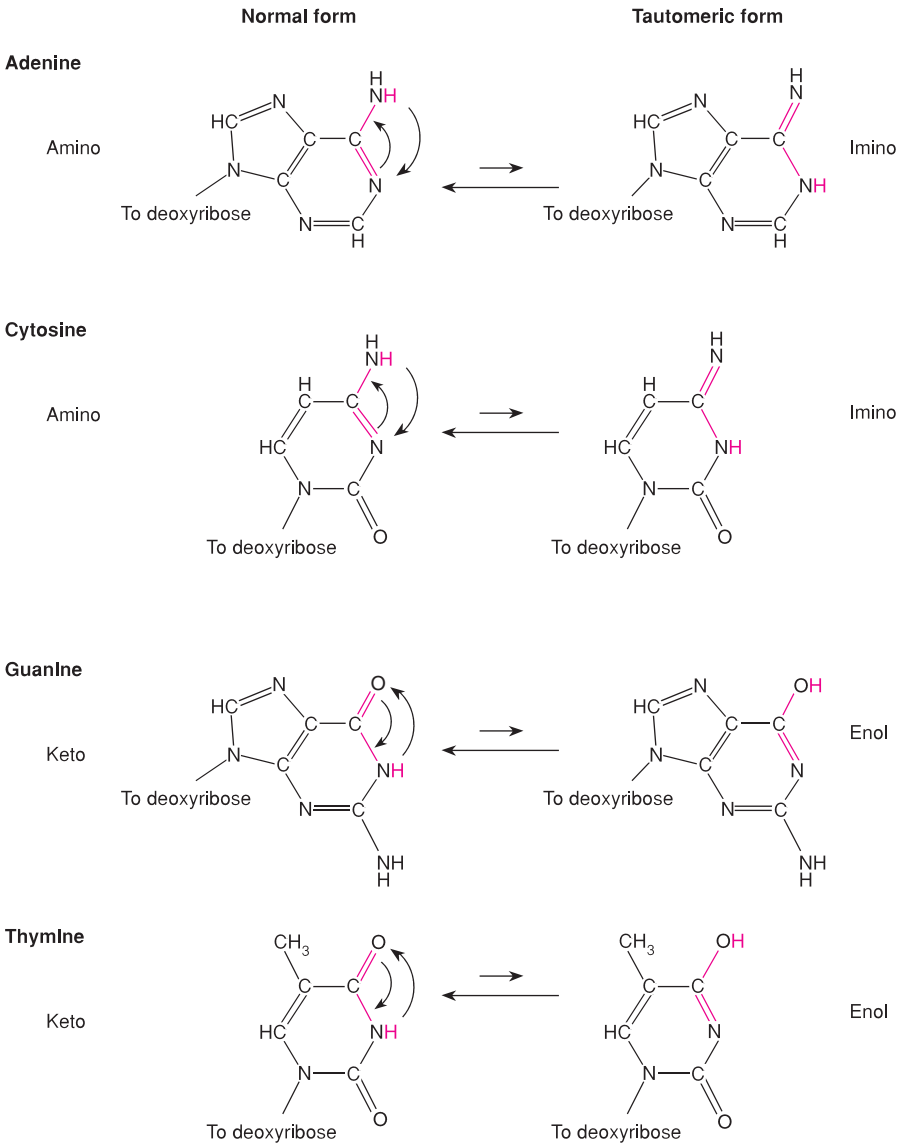


Figure 12.14 Normal and tautomeric forms of DNA bases. Adenine and cytosine can exist in the amino, or the rare imino, forms; guanine and thymine can exist in the keto, or rare enol, forms.

Table 12.4 **Pairing Relationships of DNA Bases in the Normal and Tautomeric Forms**

Base	In Normal State Pairs with	In Tautomeric State Pairs with
A	T	C
T	A	G
G	C	T
C	G	A

base pairings that would occur following tautomeric shifts of the DNA bases. Figure 12.15 illustrates the molecular structure of one of these tautomeric pairings.

During DNA replication, a tautomeric shift in either the incoming base (*substrate transition*) or the base already in the strand (*template transition*) results in mispairing. The mispairing will be permanent and result in a new base pair after an additional round of DNA replication. The original strand is unchanged (fig. 12.16).

In the example in figure 12.16, the replacement of one base pair maintains the same purine-pyrimidine relationship: AT is replaced by GC and GC by AT. In both examples, a purine-pyrimidine combination is replaced by a purine-pyrimidine combination. (Or, more specifically, a purine replaces another purine: guanine replaces adenine in the first example and adenine replaces guanine in the second.) The mutation is referred to as a **transition mutation:** a purine (or pyrimidine) replaces another purine (or pyrimidine) through a transitional state involving a tautomeric shift. When a purine replaces a pyrimidine or vice versa, it is referred to as a **transversion mutation.**

Transversions may arise by a combination of two events, a tautomerization and a base rotation. (We saw base rotations in the formation of Z DNA in chapter 9.) For example, an AT base pair can be converted to a TA base pair (a transversion) by an intermediate AA pairing

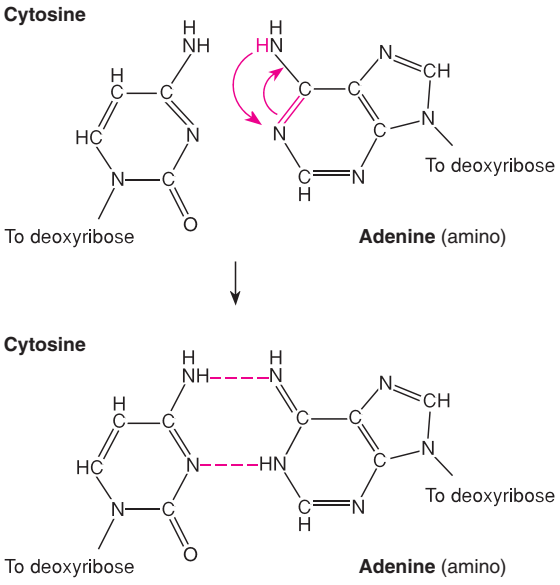


Figure 12.15 Tautomeric forms of adenine. In the common amino form, adenine does not base-pair with cytosine; in the tautomeric imino form, it can.

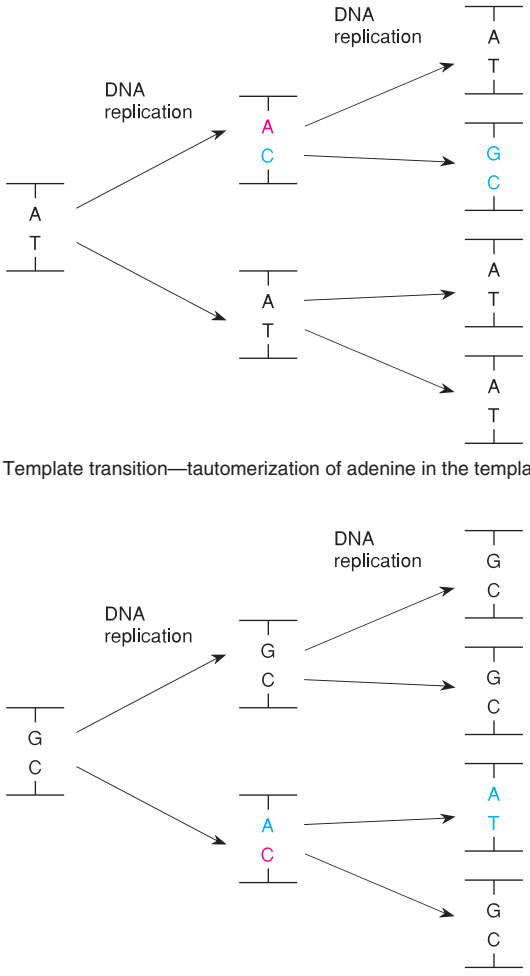


Figure 12.16 Tautomeric shifts result in transition mutations. The tautomerization can occur in the template base or in the substrate base. Tautomeric shifts are shown in *red;* the resulting transition in *blue.* The transition shows up after a second generation of DNA replication.

(fig. 12.17). Adenine can pair with adenine if one of the bases undergoes a tautomeric shift while the other rotates about its base-sugar (glycosidic) bond (fig. 12.18). The normal configuration of the base is referred to as the *anti* configuration; the rotated form is the *syn* configuration. Since we now believe that as many as 10% of bases may be in the *syn* configuration at any moment, the transversion mutagenesis rate should be about 10% of the transition mutagenesis rate, a value not inconsistent with current information.

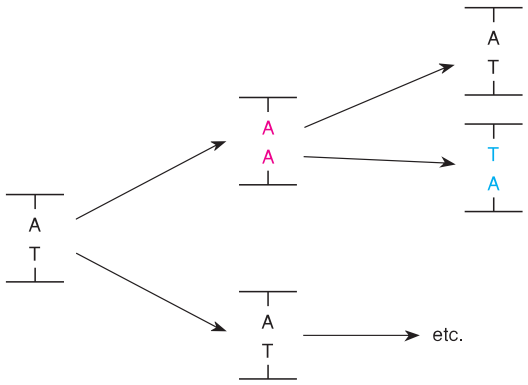


Figure 12.17 A model for transversion mutagenesis. An AT base pair can be converted to a TA base pair (a transversion) by way of an intermediate AA base pair. One of the *red* bases is in the rare tautomeric form, while the other is in the *syn* configuration. After a second round of DNA replication, one DNA duplex will have a transversion at that point (*blue*).

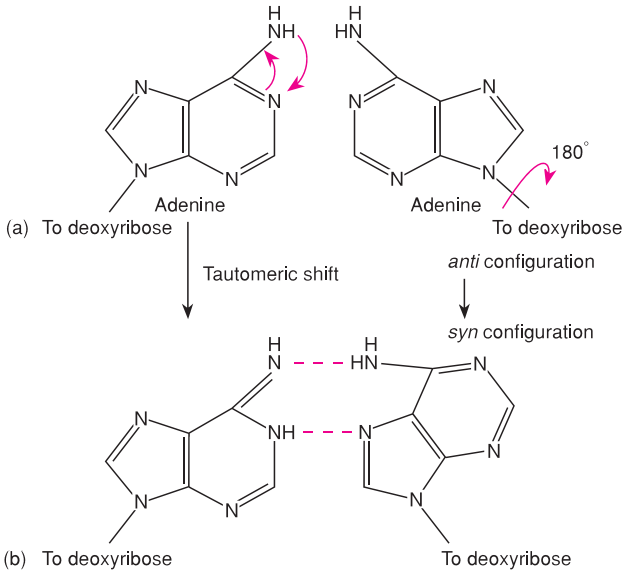


Figure 12.18 Transversion mutagenesis. An AA base pair can form if one base undergoes a tautomeric shift while the other rotates about its glycosidic (sugar) bond. In (*a*), both bases are in their normal configurations; no hydrogen bonding occurs. In (*b*), hydrogen bonds are possible.

Some base-pair mutations can have serious results. If guanine undergoes an oxidation to 8-oxoguanine (fig. 12.19), it pairs with adenine. A GC base pair is converted to a TA base pair through an 8-oxoguanine-adenine intermediate. This transversion has been found to be common in cancers.

Since 1953, when Watson and Crick first described the structure of DNA, tautomerization has been accepted as the obvious source of most transition mutations. However, recent structural data has cast some doubt on this assumption. X-ray crystallography and nuclear magnetic resonance (NMR) studies indicate that both bases in transition mismatches may be in their normal forms. Other mechanisms, similar to wobble base pairing (see chapter 11), may be responsible for most transition mutations. These studies also indicate that some transversions result from direct purine-purine or pyrimidine-pyrimidine base pairing during DNA synthesis. Much work needs to be done to clarify the nature of spontaneous mutagenesis.

Chemical Mutagenesis

Muller demonstrated that X rays can cause mutation. Certain chemical and temperature treatments can also cause mutation. Determining the mode of action of various chemical mutagens has provided insight into the mutational process as well as the process of carcinogenesis (box 12.1). In addition, knowing how chemical mutagens act has allowed geneticists to produce large numbers of specific mutations at will (box 12.2).

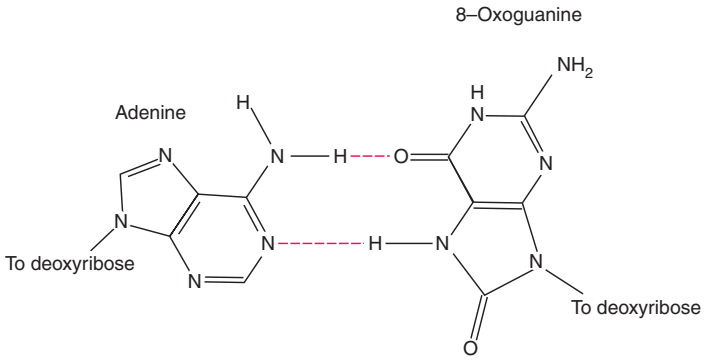


Figure 12.19 The structure of 8-oxoguanine, which base pairs with adenine, converting a GC base pair to a TA base pair—a transversion. (After C. Mol, et al., "DNA repair mechanisms for the recognition and removal of damaged DNA bases," *Annual Review of Biophysics and Biomolecular Structure,* 28:101–28,1999, figure 6.)

Transitions

Transitions are routinely produced by base analogues. Two of the most widely used base analogues are the pyrimidine analogue 5-bromouracil (5BU) and the purine analogue 2-aminopurine (2AP; fig. 12.20). The mutagenic mechanisms of the two are similar. The 5-bromouracil is incorporated into DNA in place of thymine; it acts just like thymine in DNA replication and, since it doesn't alter the hydrogen bonding, should induce no mutation. However, it seems that the bromine atom causes 5-bromouracil to tautomerize more readily than thymine does. Thus, 5-bromouracil goes from the keto form (fig. 12.20) to the enol form more readily than thymine. Transitions frequently result when the enol form of 5-bromouracil pairs with guanine.

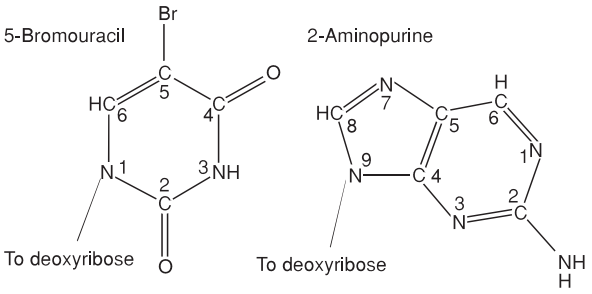


Figure 12.20 Structure of the base analogues 5-bromouracil (5BU) and 2-aminopurine (2AP).

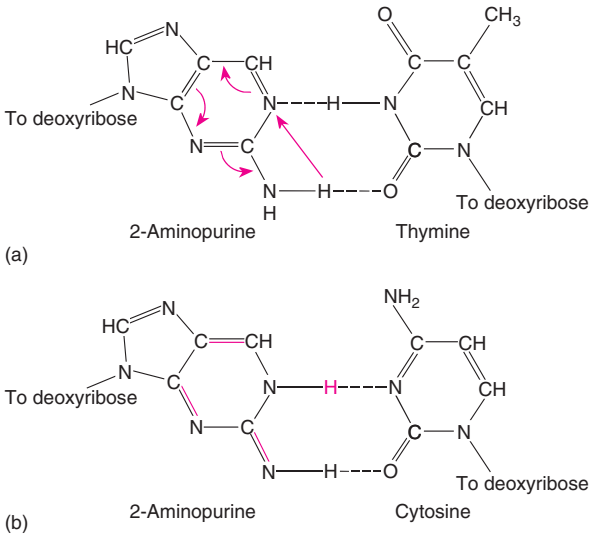


Figure 12.21 Two possible base pairs with 2-aminopurine. (*a*) In the normal state, 2-aminopurine acts like adenine and pairs with thymine. (*b*) In the rare state, 2-aminopurine acts like guanine and forms complementary base pairs with cytosine.

The 2-aminopurine is mutagenic by virtue of the fact that it can, like adenine, form two hydrogen bonds with thymine. When in the rare state, it can pair with cytosine (fig. 12.21). Thus, at times it replaces adenine, and at other times guanine. It promotes transition mutations.

Nitrous acid (HNO_2) also readily produces transitions by replacing amino groups on nucleotides with keto groups ($—NH_2$ to $=O$). The result is that cytosine is converted to uracil, adenine to hypoxanthine, and guanine to xanthine. As figure 12.22 shows, transition

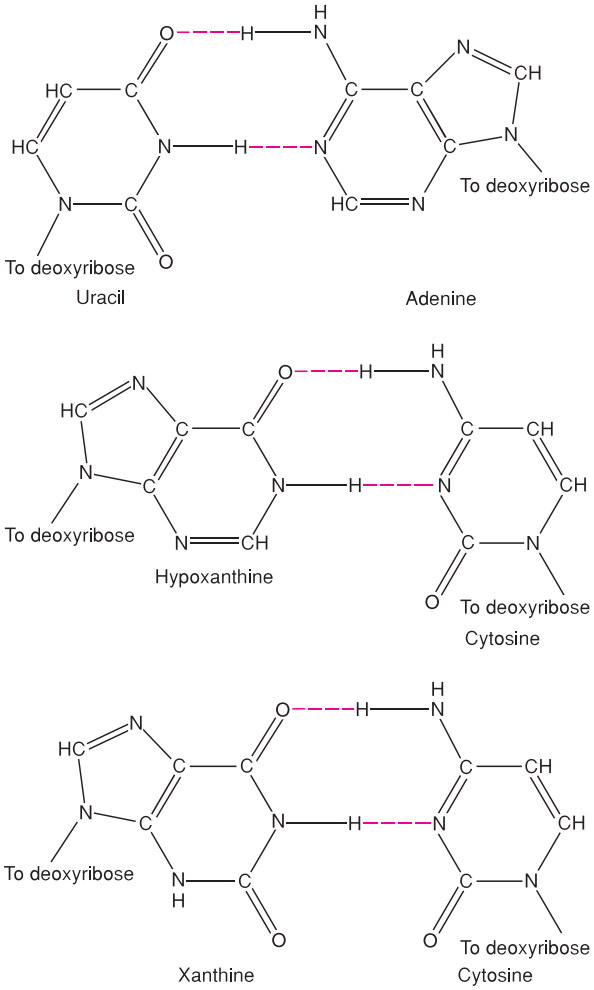


Figure 12.22 Nitrous acid converts cytosine to uracil, adenine to hypoxanthine, and guanine to xanthine. Uracil pairs with adenine, whereas its progenitor, cytosine, normally pairs with thymine; hypoxanthine pairs with cytosine, whereas its progenitor, adenine, normally pairs with thymine; and xanthine pairs with cytosine—the same base that guanine, its progenitor, pairs with. Thus, only the first two pairings result in transition mutations.

BOX 12.1

Biomedical Applications

The Ames Test for Carcinogens

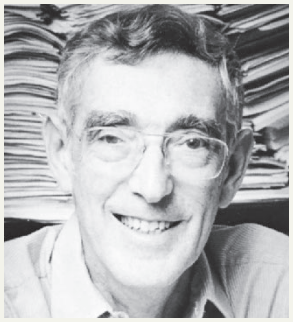
Which chemicals cause cancer in human beings? It is difficult to determine whether any substance is a carcinogen. Tests for carcinogenicity usually involve administering the substance in question to laboratory rats or mice to determine whether the substance actually causes cancer in these animals. Even these tests, however, are not absolute predictors of cancer in people. Tests of this nature are very expensive ($1 million to $2 million each) and time-consuming (three to four years). Since more than fifty thousand different chemical compounds are used in industry, with thousands more added each year, the challenge of making the working environment as well as the general environment safe seems overwhelming. We can, however, make a preliminary determination about the cancer-causing properties of any substance very quickly because of the relationship between mutagenicity and carcinogenicity. Many substances in the environment that can cause cancer also cause mutations. Both kinds of effects are related to DNA damage.

Bruce Ames, at the University of California at Berkeley, developed a routine screening test for mutagenicity. Substances that prove positive in this test are suspected of being carcinogens and would have to be tested further to determine their potential to cause cancer in mammals.

Ames worked with a strain of *Salmonella typhimurium* that requires histidine to grow. This strain will not grow on minimal medium. However, the strain will grow if a mutagen is added to the medium, causing the defective gene in the histidine pathway to revert to the wild-type. (Mutagens inducing gross chromosomal damage, such as deletions or inversions, will not be detected.) Under normal circumstances, there is a background mutation rate; a certain number of *Salmonella* cells revert spontaneously, and therefore a certain number of colonies will grow on the minimal medium. A mutagen, however, increases the number of colonies that can grow on minimal medium. This procedure is, therefore, a rapid, inexpensive, and easy test for mutagenicity.

To improve this test's ability to detect carcinogens, Ames added a supplement of rat liver extract to the medium. It is known that, although many substances are themselves not carcinogens, the breakdown of these substances in the liver creates substances that are carcinogenic. Rat liver enzymes act on a substance the same way human livers do, converting a noncarcinogenic primary substance into a possible carcinogen. The liver enzymes can also make a mutagen nonmutagenic.

Other short-term tests are in use that effectively duplicate the Ames test. These include tests for mutagenicity in mouse lymphoma cells and two tests in Chinese hamster ovary cells: a test for chromosomal aberrations and a test for sister-chromatid exchanges. None of these tests surpasses the Ames test, which has scored better than 90% correct when tested with hundreds of known carcinogens. Thousands of other substances have been subjected to this test; many have proven to be mutagenic. These substances are usually withdrawn from the workplace or home environment. From time to time, we read that a certain substance is believed to be carcinogenic and is being removed from grocery store shelves. Examples have included hair dyes, food preservatives, food-coloring agents, and artificial sweeteners. Many of these first were suspected after they failed the Ames test.



Bruce Ames (1928–).
(Courtesy Dr. Bruce Ames.)

BOX 12.2

Experimental Methods

In Vitro Site-Directed Mutagenesis

One way of studying the way that proteins work is to change the sequence of amino acids in the protein. For example, if a scientist were working on the active site of a particular enzyme, he or she could learn how the enzyme modifies its substrates by changing one or a few amino acids. Changes could be made, for example, in order to study the role of shape or charge on the functioning of the enzyme. Advances in recombinant DNA techniques have made it possible for a research scientist to create exactly the changes he or she wants in a protein.

To begin with, the gene for the protein or enzyme must be cloned so that it can be manipulated (see chapter 13). Once cloned, deletions are easy to create with restriction endonucleases (described in chapter 13). If a particular endonuclease cuts the gene in two places, the intervening segment can be spliced out (fig. 1*a*; see chapter 13). If the endonuclease cuts only once, exonucleases can digest the ends of the cut, extending the deletion away from the cut in both directions (fig. 1*b*). Insertions can be created by either cutting the gene and repairing the single-stranded ends (fig. 1*c*) or by creating an oligonucleotide (a linker) with the desired sequence and inserting the linker at the site of an endonuclease cut (fig. 1*d*).

Far more impressive, however, is the ability to change a single specific codon in order to replace any amino acid in the protein with any other amino acid. The process involves directed mutagenesis using artificially created oligonucleotides.

Basically, a short sequence of DNA (an oligonucleotide) is synthesized complementary to a region of the cloned gene, but with a change in one or more bases of a codon to specify a different amino acid. That oligonucleotide is then hybridized with the single-stranded form of the clone (fig. 2). Although one or more

continued

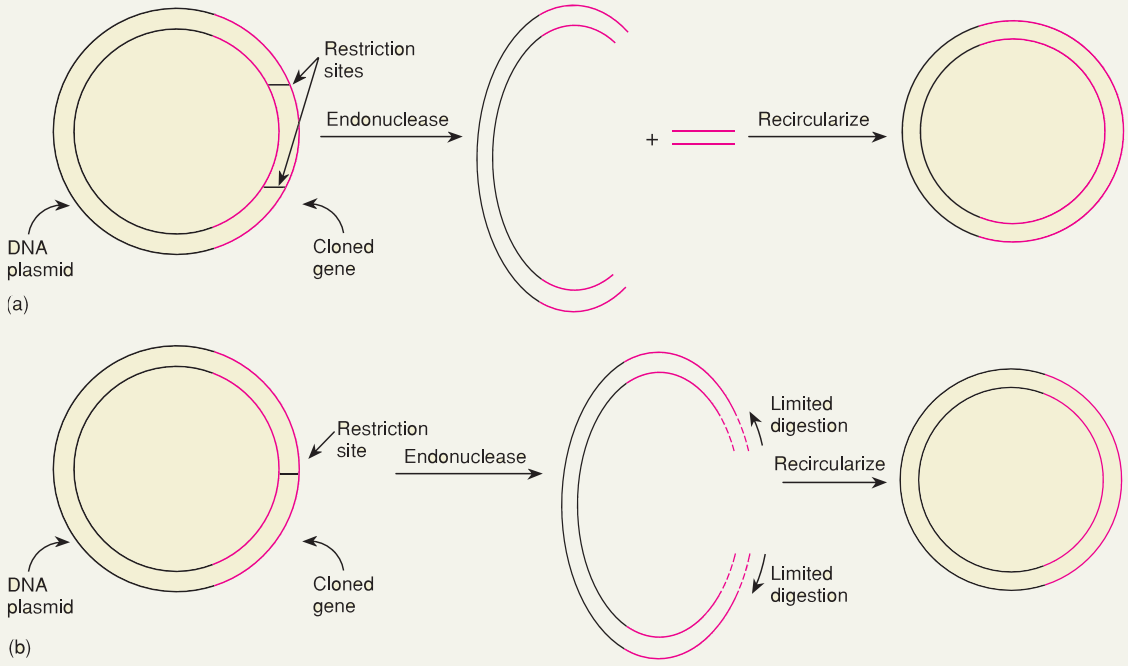
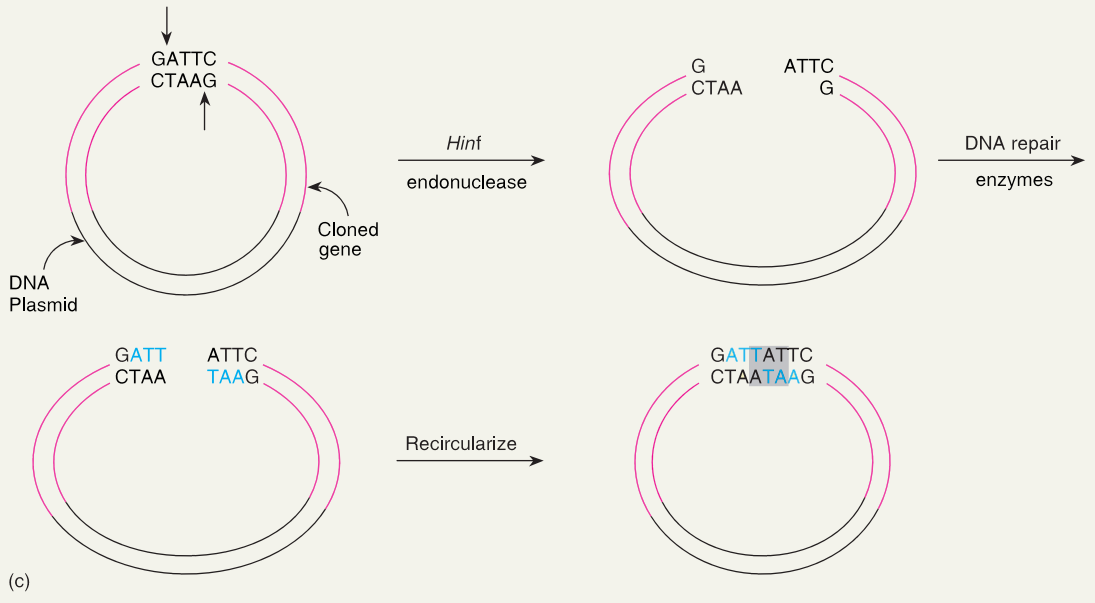


Figure 1 A cloned gene can be mutated in several ways. (*a*) If a restriction endonuclease has two sites in the gene, the intermediate piece can be spliced out. (*b*) If the endonuclease has only one site, the gene can be opened at that site, and limited digestion by exonucleases will delete part of the gene.

continued

BOX 12.2 CONTINUED



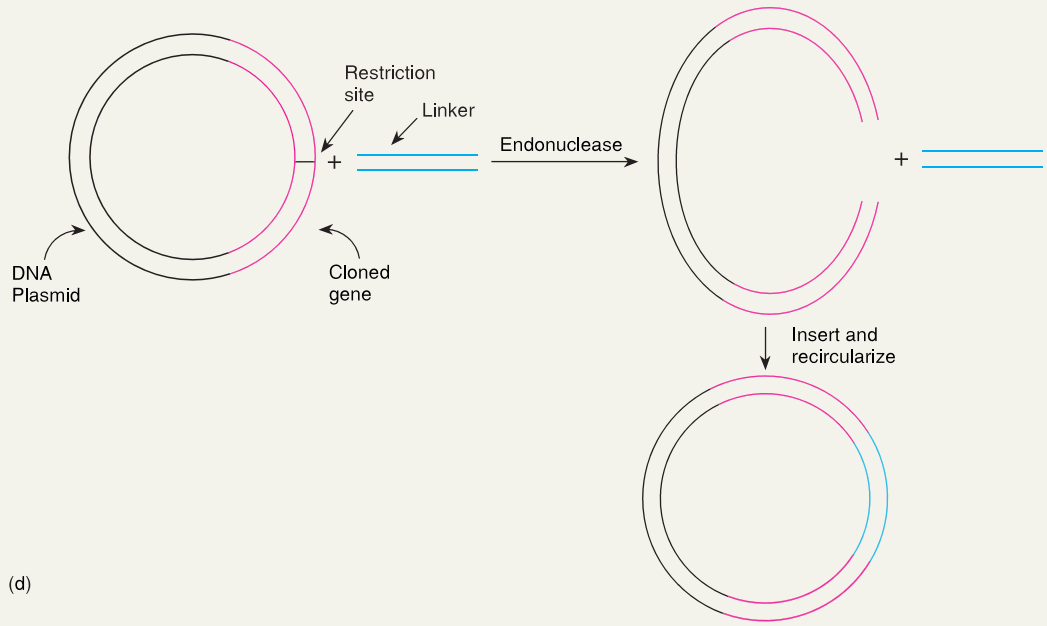


Figure 1—*continued* (*c*) If an endonuclease has an offset region between its splice points of three or six nucleotides (one or two codons), that length can be inserted by repairing the single-stranded ends after cutting by the endonuclease. The resulting blunt ends can be spliced together. (Note that actually an ATT region has been converted to an ATTATT region. If reading codons along the DNA, the actual insertion is of a TAT codon.) (*d*) A linker of any length (usually the length of a specific number of codons) can be inserted at a restriction site.

bases will not match, hybridization can usually be facilitated by adjusting the pH or ionic strength of the solution. The hybridized oligonucleotide is then used as a primer for DNA replication; the whole plasmid is replicated, resulting in hybrid DNA. In subsequent DNA replications of the hybrid, both the original gene and the mutated DNA will be produced. The latter can be isolated by appropriate selection methods; it is a plasmid with a cloned gene that has the exact mutation the researcher wanted. Using techniques of this type, geneticists have made many advancements in understanding exactly how various components of an enzyme contribute to its function.

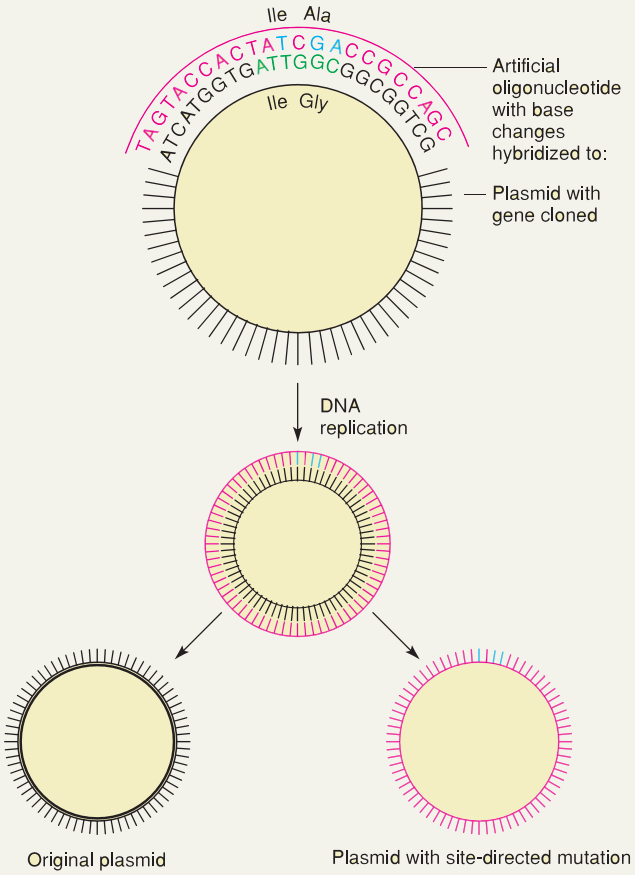


Figure 2 Site-directed mutagenesis can involve any nucleotide(s). In this case, an inserted gene with an Ile-Gly sequence is converted, at the direction of the investigator, to an Ile-Ala sequence. A single-stranded form of the plasmid is isolated. A synthetically prepared oligonucleotide (twenty-three bases in this example) is added. It can be made to hybridize at the complementary site despite differing by three bases. Then DNA replication is carried out using the oligonucleotide configuration as a primer. After the strands of the duplex are separated, the investigators can isolate the original plasmid as well as the mutated plasmid. (Note that the investigators changed two codons, although they changed only one amino acid, because they also wanted to introduce an Alu site at that point for future studies.) (From J. E. Villafranca, et al., "Directed mutagenesis of dihydrofolate reductase," *Science* 222:782–88. Copyright 1983 by the AAAS.)

mutation results from two of the changes. Uracil pairs with adenine instead of guanine, thus leading to a UA base pair in place of a CG base pair; hypoxanthine (H) pairs with cytosine instead of thymine, the original base paired with adenine. Thus, in this case, an HC base pair replaces an AT base pair. Both of these base pairs (UA and HC) are transition mutations. Xanthine, however, pairs with cytosine just as guanine does. Thus, the replacement of guanine with xanthine does not cause changes in base pairing.

Like nitrous acid, heat can also deaminate cytosine to form uracil and thus bring about transitions (CG to TA). Apparently, heat can also bring about transversions by an unknown mechanism.

Transversions

Ethyl methane sulfonate ($CH_3SO_3CH_2CH_3$) and ethyl ethane sulfonate ($CH_3CH_2SO_3CH_2CH_3$) are agents that cause the removal of purine rings from DNA. The multistep process begins with the ethylation of a purine ring and ends with the hydrolysis of the glycosidic (purine-deoxyribose) bond, causing the loss of the base. These sites where this happens are referred to as AP (apurinic-apyrimidinic) sites. If the AP site is not repaired, any of the four DNA bases could be inserted into the new strand opposite the gap (fig. 12.23). If thymine is placed in the newly formed strand, then the original base pair is restored; insertion of cytosine results in a transition mutation; insertion of either adenine or guanine results in a transversion mutation. Of course, the gap is still there, and it continues to generate new mutations each generation until it is repaired. During DNA replication in *E. coli,* the polymerase tends to place adenine opposite the gap more frequently than it places other bases.

Insertions and Deletions

The molecules of the acridine dyes, such as proflavin and acridine orange (fig. 12.24), are flat. Presumably, they initiate mutation by inserting into the DNA double helix, causing the helix to buckle in the region of insertion, possibly leading to base additions and deletions during DNA replication. Crick and Brenner used acridine-induced mutations to demonstrate both that the genetic

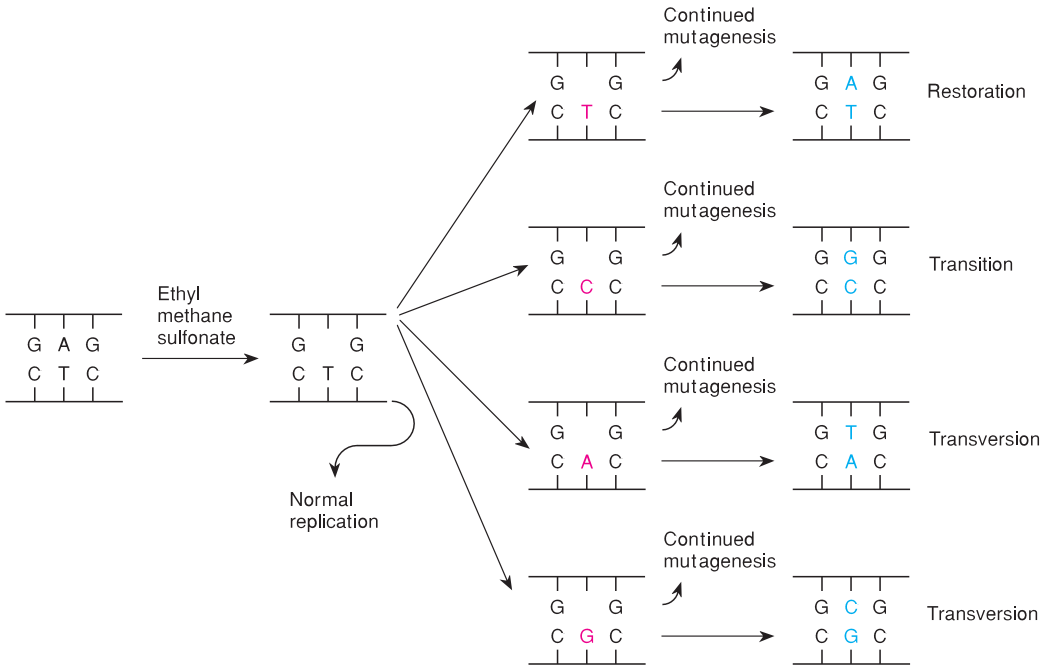


Figure 12.23 Four possible outcomes after treatment of DNA with an alkylating agent, which removes the purine—adenine in this example. The bases shown in *red* are the four bases that DNA polymerase may insert opposite the gap. After another round of DNA replication, the gap remains to generate further mutations. The inserted base forms a base pair (*blue*), which can be a restoration or a transition or transversion mutation.

code was read from a fixed point and that it was triplet (chapter 11).

Misalignment Mutagenesis

Additions and deletions in DNA can also come about by misalignment of a template strand and the newly formed (progeny) strand in a region containing a repeated sequence. For example, in figure 12.25 we expect the progeny strand to contain six adjacent adenines because the template strand contains six adjacent thymines. Misalignment of the progeny strand results in seven consecutive adenines: six thymines replicated, plus one already replicated but misaligned. Misalignment of the template strand results in five consecutive adenines because one thymine is not available in the template. Regions with long runs of a particular base may be very mutation prone. They may explain the "hot spots" observed by Benzer (see fig. 12.8) and others.

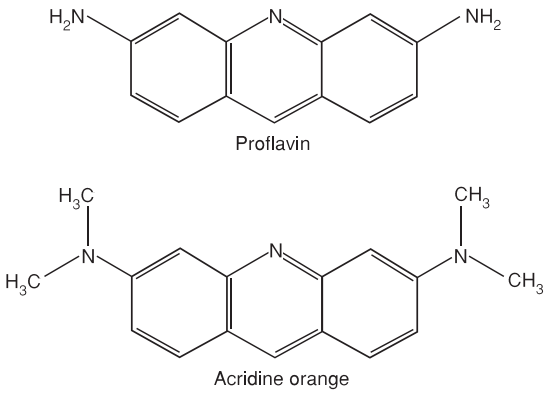


Figure 12.24 Structure of two acridine dyes: proflavin and acridine orange.

Intergenic Suppression

When a critical mutation occurs in a codon, several routes can still lead to survival of the individual; simple reversion and intragenic suppression are two that we have already considered. A third route is through **intergenic suppression**—restoration of the function of a mutated gene by changes in a different gene, called a **suppressor gene.** Suppressor genes are usually transfer RNA genes. When mutated, intergenic suppressors change the way in which a codon is read.

Suppressor genes can restore proper reading to nonsense, missense, and frameshift mutations. **Nonsense mutations** convert a codon that originally specified an amino acid into one of the three nonsense

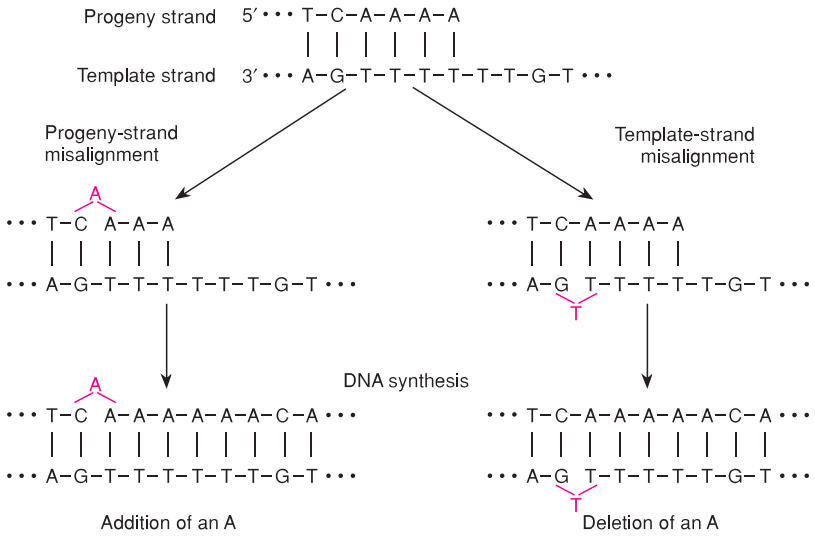
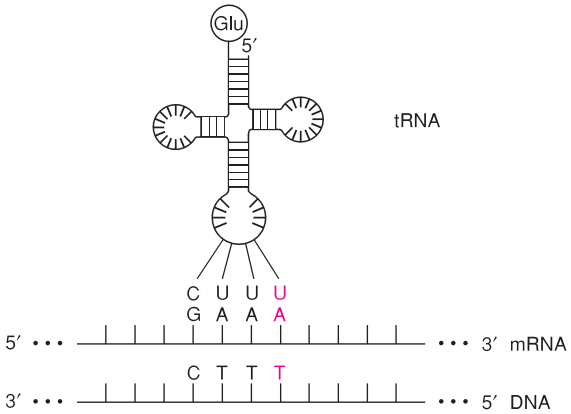


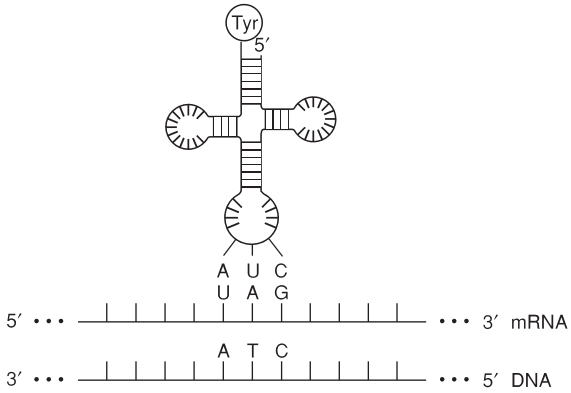
Figure 12.25 Misalignment of a template or progeny strand during DNA synthesis. If the progeny strand is misaligned after DNA replication has begun, the resulting progeny strand will have an additional base. If the template strand is misaligned during DNA replication, the resulting progeny strand will have a deleted base. These changes will show up after another round of DNA replication. (From J.W. Drake, B.W. Glickman, and L.S. Ripley, "Updating the Theory of Mutation," *American Scientist,* 71:621–630, 1983. Reprinted with permission of American Scientist, magazine of Sigma Xi, The Scientific Research Society.)

codons. **Missense mutations** change a codon so that it specifies a different amino acid. Frameshift mutations, by additions or deletions of nucleotides, cause an alteration in the reading frame of codons. A frameshift mutation, caused by the insertion of a single base, can be suppressed by a transfer RNA that has an added base in its anticodon (fig. 12.26*a*). It reads four bases as a codon and thus restores the original reading frame.

The transfer RNA produced by a nonsense suppressor gene reads the nonsense codon as if it were a codon for an amino acid; an amino acid is placed into the protein, and reading of the messenger RNA continues. At least three suppressors of the mutant amber codon (UAG) are known in *E. coli.* One suppressor puts tyrosine, one puts glutamine, and one puts serine into the protein chain at the point of an amber codon. Normally, tyrosine transfer RNA has the anticodon 3′-AUG-5′. The suppressor transfer RNA that reads amber as a tyrosine codon has the anticodon 3′-AUC-5′, which is complementary to amber. Hence, a mutated tyrosine transfer RNA reads amber as a tyrosine codon (fig. 12.26*b*).



(a) Frameshift suppression



(b) Nonsense suppression

Figure 12.26 Frameshift and nonsense suppression by mutant transfer RNAs. In (*a*), a thymine has been inserted into DNA, resulting in a frameshift. However, a transfer RNA with four bases in the codon region reads the inserted base as part of the previous codon in the messenger RNA. The frameshift thus does not occur. In (*b*), an *amber* mutation (UAG), which normally results in chain termination, is read as tyrosine by a mutant tyrosine transfer RNA that has the anticodon sequence complementary to the *amber* codon.

If the amber nonsense codon is no longer read as a stop signal, then won't all the genes terminating in the amber codon continue to be translated beyond their ends, causing the cell to die? In the tyrosine case, two genes for tyrosine transfer RNA were found; one contributes the major fraction of the transfer RNAs, and the other, the minor fraction. It is the minor-fraction gene that mutates to act as the suppressor. Thus, most messenger RNAs are translated normally, and most amber mutations result in premature termination, although a sufficient number are translated (suppressed) to ensure the viability of the mutant cell. In general, intergenic suppressor mutants would be eliminated quickly in nature because they are inefficient—the cells are not healthy. In the laboratory, we can provide special conditions that allow them to be grown and studied.

Mutator and Antimutator Mutations

Whereas intergenic suppressors represent mutations that "restore" the normal phenotype, mostly through mutation of transfer RNA loci, **mutator** and **antimutator mutations** cause an increase or decrease in the overall mutation rate of the cell. They are frequently mutations of DNA polymerase, which, as you remember, not only polymerizes DNA nucleotides 5′ → 3′ complementary to the template strand, but also checks to be sure that the correct base was put in (they proofread). If, in the proofreading process, the polymerase discovers an error, it can correct this error with its 3′ → 5′ exonuclease activity. Mutator and antimutator mutations sometimes involve changes in the polymerase's proofreading ability (exonuclease activity).

Phage T4 has its own DNA polymerase with known mutator and antimutator mutants. Mutator mutants are very poor proofreaders (they have low exonuclease-to-polymerase ratios), and thus they introduce mutations throughout the phage genome. Antimutator mutants, however, have exceptionally efficient proofreading ability (high exonuclease-to-polymerase ratios) and, therefore, result in a very low mutation rate for the whole genome (box 12.3).

BOX 12.3

Experimental Methods

Adaptive Mutation

Although we discuss evolution in detail at the end of the book, note here that we view mutation as a random process, not one that occurs because a cell "needs" a particular mutation. For example, Luria and Delbrück's work showed that the mutation in *E. coli* for resistance to phage T1 occurred randomly before exposure to the phage, not because the cells would benefit from the mutation.

Because of the entrenched dogma that mutations occur through random processes, the scientific community was startled when, in 1988, John Cairns—a highly respected senior scientist—and colleagues reported a new observation—**adaptive,** or **directed, mutation,** occurring when the cell needed it. Their system was the *lacZ* gene in *E. coli.* Cells that could not use lactose as an energy source ($lacZ^-$) were plated on a medium in which lactose was the sole energy source. Some mutants already there of course produced colonies ($lacZ^+$). The expectation was that no new mutations would occur over time because the $lacZ^-$ cells would have either died or stopped metabolizing. Unexpectedly, Cairns and colleagues found that more and more colonies appeared over time, coming from cells that had mutated to $lacZ^+$. As a control, they looked for revertants of other genes not involved with lactose metabolism. These mutations did not occur in a directed manner.

Scientists were extremely skeptical of this work for two reasons. First, it seemed to fly in the face of our common understanding of the mutational process. Second, there were no obvious explanations for how this could occur. Numerous articles were published refuting the notion of adaptive mutations and suggesting other explanations for the results. These explanations included artifacts of miscounting cells to mutants that were extremely slow growing, but there all the time.

In the past several years, other scientists have found at least a half dozen similar results in other organisms and other genes. The debate is ongoing, but work published in mid-1994 seems to have recast it into the realm of methods of mutagenesis rather than non-Darwinian processes. Several scientists found that the "directed" mutations seem to be of a certain type, mainly single nucleotide deletions within runs of the same nucleotide: for example, a deletion of a C in a CCCC sequence. This type of error happens during DNA replication and could be the result of a repair deficiency. That is, under extreme duress, the cells may be going into a "hypermutational" mode, or selection may favor hypermutable genotypes in which repair mechanisms are shut down in order to intentionally create lots of errors in the DNA. Any errors that do not alleviate the problem result in cell death, a death that was inevitable anyway; however, some errors will correct the problem ($lacZ^-$ to $lacZ^+$). Those cells will survive. One recent study indicated that a subpopulation of about 0.06% of the population was hypermutable, with a mutation rate about 200 times that of the normal cells. That group of cells could account for the adaptive mutations. Other scientists found that when the locus of importance was on a plasmid, the adaptive mutation could occur by increased replication of the plasmid. (Currently, the favored term is *adaptive mutation,* rather than *directed mutation,* indicating that the mutations are useful to the cell, not that some unknown process directs them.)

This controversy, although not necessarily resolved, has actually brought out the best in the scientific method: A skeptical scientific community tried its best to refute an unreasonable observation. Other scientists then repeated the observation and extended it, making it more worthy of further study. Finally, further work has given us reasonable mechanisms that not only require no rejection of the original concept of random mutation, but actually give us hypotheses to test further.

DNA REPAIR 

Radiation, chemical mutagens, heat, enzymatic errors, and spontaneous decay constantly damage DNA. For example, it is estimated that several thousand DNA bases are lost each day in every mammalian cell due to spontaneous decay. Some types of DNA damage interfere with DNA replication and transcription. In the long evolutionary challenge to minimize mutation, cells have evolved numerous mechanisms to repair damaged or incorrectly replicated DNA. Many enzymes, acting alone or in concert with other enzymes, repair DNA. Repair systems are generally placed in four broad categories: damage reversal, excision repair, double-strand break repair, and

postreplicative repair. Enzymes that process repair steps have been conserved during evolution. That is, enzymes found in *E. coli* have homologues in yeast, fruit flies, and human beings. However, eukaryotic systems are almost always more complex.

Damage Reversal

Ultraviolet (UV) light causes linkage, or **dimerization,** of adjacent pyrimidines in DNA (fig. 12.27). Although cytosine-cytosine and cytosine-thymine dimers are occasionally produced, the principal products of UV irradiation are thymine-thymine dimers. These can be repaired in several different ways. The simplest is to reverse the dimerization process and restore the original unlinked thymines.

In *E. coli,* an enzyme called DNA photolyase, the product of the *phr* gene (for **photoreactivation**), binds to dimerized thymines. When light shines on the cell, the enzyme breaks the dimer bonds with light energy. The enzyme then falls free of the DNA. This enzyme thus reverses the UV-induced dimerization. Another example of an enzyme that performs direct DNA repair is O^6-mGua DNA methyltransferase, which removes the methyl groups from O^6-methylguanine, the major product of DNA-methylating agents (fig. 12.28). Although other repair mechanisms seem to be present in all organisms, photoreactivation is not; it is apparently absent in human beings.

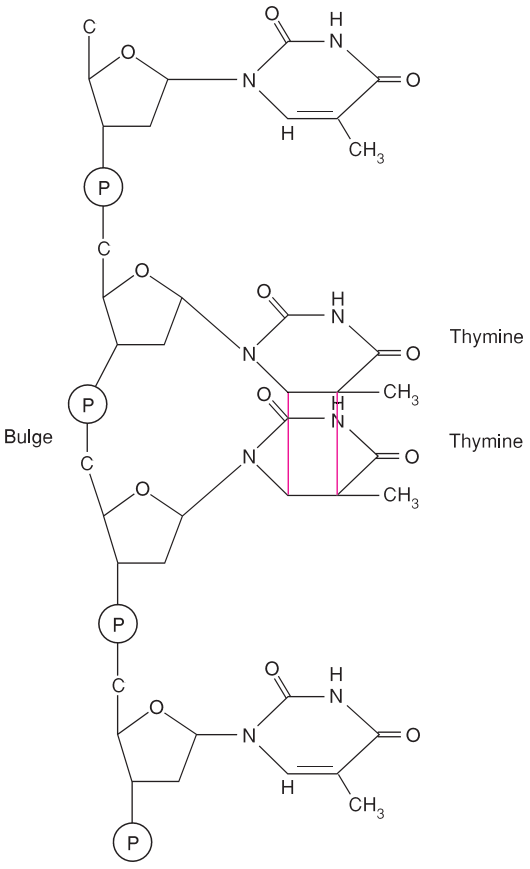


Figure 12.27 UV-induced dimerization of adjacent thymines in DNA. The *red lines* represent the dimer bonds in the adjacent thymines.

Excision Repair

Excision repair refers to the general mechanism of DNA repair that works by removing the damaged portion of a DNA molecule. Various enzymes can sense damage or distortion in the DNA double helix. During excision repair, bases and nucleotides are removed from the damaged strand. The gap is then patched using complementarity with the remaining strand. We can broadly categorize these systems as base excision repair and nucleotide excision repair, which includes mismatch repair. We will discuss only the major repair pathways; others exist. Presumably, redundancy in repair has been selected for because of the critical need to keep DNA intact and relatively mutation free.

Base Excision Repair

A base can be removed from a nucleotide within DNA in several ways: by direct action of an agent such as radiation, by spontaneous hydrolysis, by an attack of oxygen free radicals, or by **DNA glycosylases,** enzymes that sense damaged bases and remove them. Currently, at least five DNA glycosylases are known. For example, uracil-DNA glycosylase, the product of the *ung* gene in

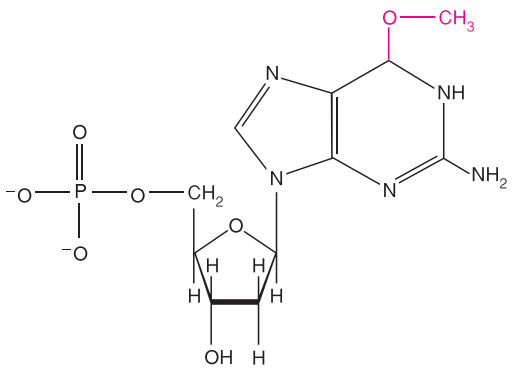


Figure 12.28 The structure of O^6-methylguanine. The *red color* shows the modification of guanine, in which the normal configuration is a double-bonded oxygen (keto form).

E. coli, recognizes uracil within DNA and cleaves it out at the base-sugar (glycosidic) bond. The resulting site is called an AP (apurinic-apyrimidinic) site, because of the lack of a purine or pyrimidine at the site (see fig. 12.23). An **AP endonuclease** then senses the minor distortion of the DNA double helix and initiaties excision of the single AP nucleotide in a process known as **base excision repair.** The AP endonuclease nicks the DNA at the 5′ side of the base-free AP site. A DNA polymerase then inserts a nucleotide at the AP site; an exonuclease, lyase, or phosphodiesterase enzyme then removes the base-free nucleotide. (Lyases are enzymes that can break C-C, C-O, and C-N bonds.) DNA ligase then closes the nick (fig. 12.29). The replacement of just one base occurs 80–90% of the time. In the remaining 10–20% of cases, several nucleotides may be removed, depending probably on which DNA polymerase (I or III) first repairs the site (fig. 12.29). In mammals, DNA polymerase β performs two roles in base excision repair: It both inserts a new base where the AP site was and also eliminates the AP nucleotide residue by exonuclease activity.

One question that concerned scientists was how the glycosylases gain access to the inappropriate or damaged bases within the double helix. Recently, it has been demonstrated that these enzymes remove the inappropriate or damaged bases by first flipping them out of the interior of the double helix in a process called **base flipping.** For example, the enzyme in human beings that recognizes 8-oxoguanine in DNA (see fig. 12.19), 8-oxoguanine DNA glycosylase, flips the base out to excise it. Base flipping seems to be a common mechanism in repair enzymes that need access to bases within the double helix (fig. 12.30).

Nucleotide Excision Repair

Whereas base excision repair is initiated by glycosylases and usually involves the replacement of only one nucleotide residue, **nucleotide excision repair** is initiated by enzymes that sense distortions in the DNA backbone and replace a short stretch of nucleotides. For example, six enzymes in *E. coli* excise a short stretch of DNA containing thymine dimers if the dimerization is not reversed by photoreactivation. Two copies of the protein product of the *uvrA* gene (for ultraviolet light—UV—repair) combine with one copy of the product of the *uvrB* gene to form a $UvrA_2UvrB$ complex that moves along the DNA, looking for damage (fig. 12.31). (The complex has 5′ to 3′ helicase activity.) When the complex finds damage such as a thymine dimer, with moderate to large distortion of the DNA double helix, the $UvrA_2$ dimer dissociates, leaving the UvrB subunit alone. This causes the DNA to bend and attracts the protein product of the *uvrC* gene, UvrC. The UvrB subunit first

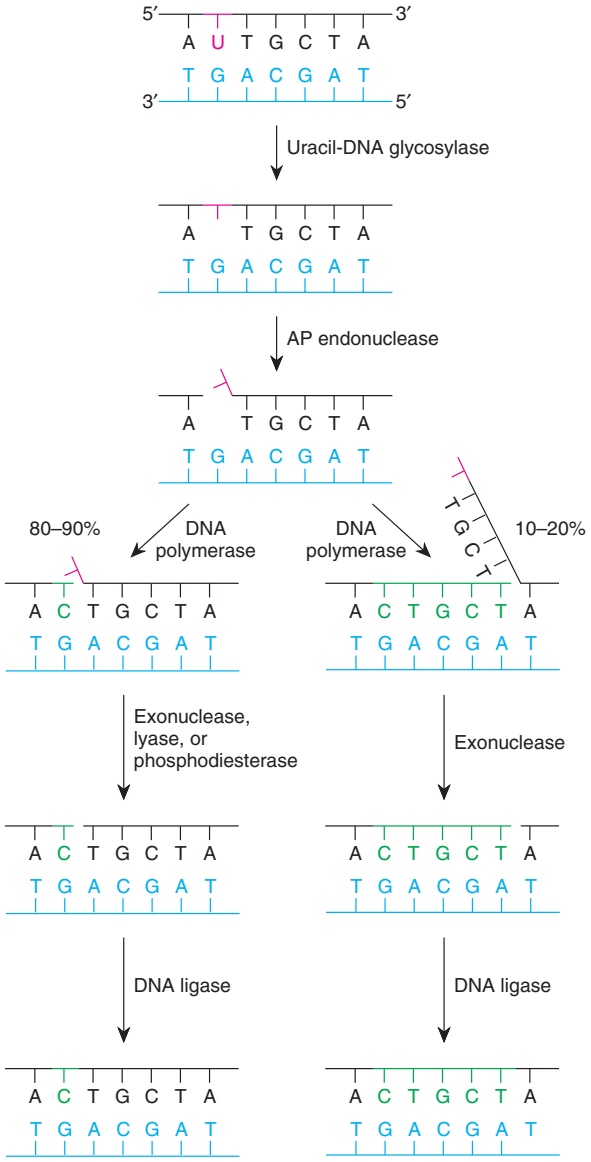
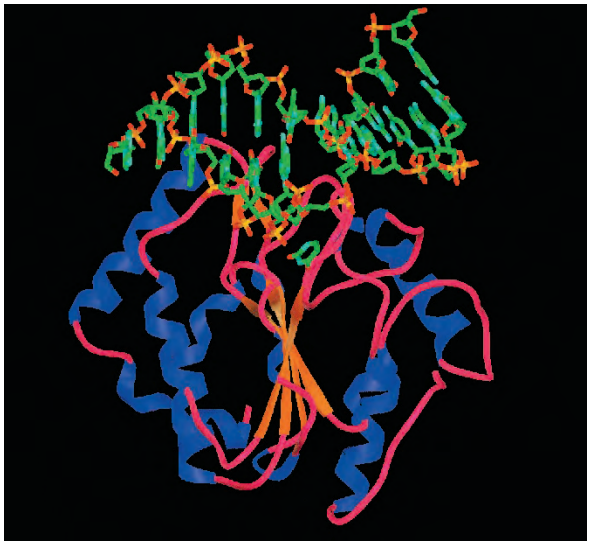
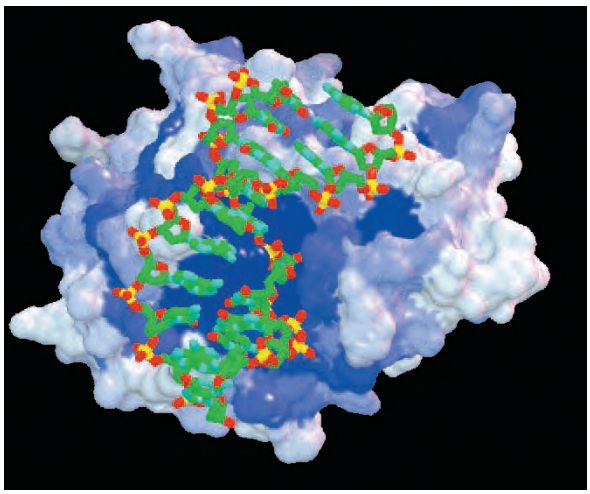


Figure 12.29 Mechanism of base excision repair. In this case, a uracil-DNA glycosylase enzyme removes a uracil (*red*) from DNA. An AP (apurinic-apyrimidinic) endonuclease nicks the DNA on the 5′ side of the base-free site. Between 80 and 90% of the time, a DNA polymerase will replace the single nucleotide (*green*); an exonuclease, lyase, or phosphodiesterase will remove the base-free nucleotide. The final nick is sealed with DNA ligase. Between 10 and 20% of the time, the DNA polymerase will extend polymerization beyond the single nucleotide. In those cases, an exonuclease and DNA ligase finish the repair. (From T. Lindahl, "The Croonian Lecture, 1996: Endogenous Damage to DNA," *Philosophical Transactions of the Royal Society of London,* B351, pp. 1529–1538, figure 6, 1996.)



(a)



(b)

Figure 12.30 Two views of the human enzyme, uracil-DNA glycosylase, bound to DNA that has a uracil present. In (*a*), the uracil-containing nucleotide residue has been flipped out and the uracil cleaved; in (*b*), the uracil-containing residue has been flipped out, but the uracil has not yet been cleaved. In both, DNA is a *green* stick figure with *red* oxygen, *yellow* phosphorus, and *blue* nitrogen atoms. In (*a*), the enzyme is shown as a ribbon diagram; note the cleaved uracil bound to the ribbon. In (*b*), the enzyme is shown as a molecular surface; the uracil-containing nucleotide is flipped out by the *purple* knob just left of and below center of the structure with the flipped-out residue to the right. (From S. Parikh, C. Mol, and J. Tainer, "Base excision repair enzyme family portrait" in *Structure*, 1997, 5:1543–1550, fig 1 a&b, p. 1544. Courtesy of J.A. Tainer, The Scripps Research Institute.)

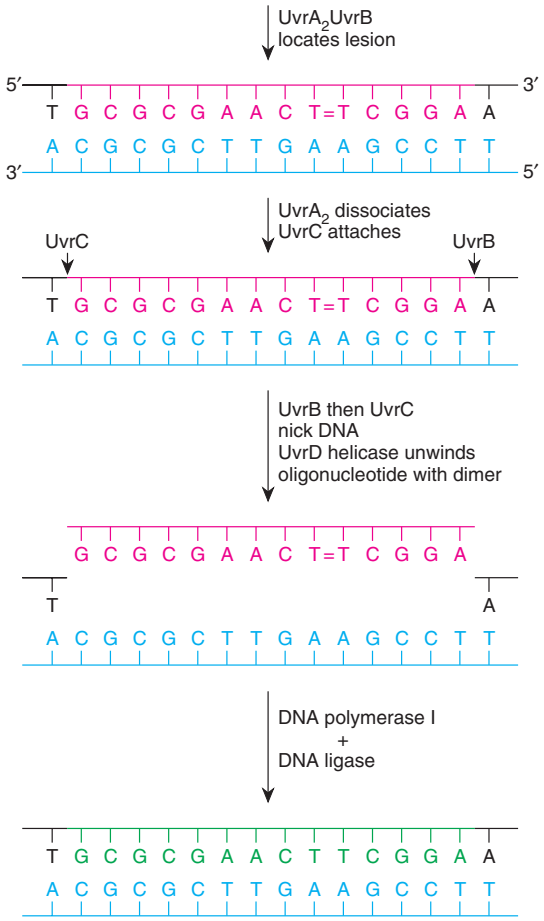


Figure 12.31 Nucleotide excision repair. A lesion in DNA (a thymine dimer) is located by a protein made of two copies of UvrA and one of UvrB. Then, the UvrA subunits detach, and UvrC attaches on the 5′ side of the lesion. UvrB nicks the DNA on the 3′ side and UvrC on the 5′ side of the lesion; UvrD helicase unwinds the oligonucleotide containing the lesion (*red*). DNA polymerase I and DNA ligase then repair the patch (*green*).

nicks (hydrolyzes) the DNA four to five nucleotides on the 3′ side of the lesion; next, the UvrC subunit nicks the DNA eight nucleotides on the 5′ side of the lesion. (The three components, UvrA, UvrB, and UvrC, are together called the *ABC excinuclease,* for excision endonuclease.) The enzyme helicase II, the product of the *uvrD* gene, then removes the twelve- to thirteen-base oligonucleotide as well as UvrC. DNA polymerase I fills in the gap and, in the process, evicts the UvrB, and DNA ligase closes the remaining nick (fig. 12.31). This is another relatively simple system designed to detect helix distortions and repair them.

Like base excision repair, nucleotide excision repair is present in all organisms. In yeast, approximately twelve genes are involved, many in what is called the RAD3 group. In human beings, twenty-five proteins are involved; they remove twenty-seven to twenty-nine nucleotides, as compared to twelve to thirteen in *E. coli.*

Transcription and nucleotide excision repair are linked in eukaryotes. Transcription factor TFIIH (see chapter 11) is involved in repair of UV damage; it has helicase activity and is found in both processes. Since it has been shown that genes that are actively being transcribed are preferentially repaired, we can now envision a model in which transcription, when blocked by a DNA lesion like a thymine dimer, signals the formation of a repair complex, using TFIIH in both processes. In prokaryotes, RNA polymerase dissociates from the DNA in this circumstance, losing the nascent transcript. This would be inefficient in eukaryotes, whose genes are much longer and more expensive to transcribe; for example, the human dystrophin gene, defective in the disease Duchenne muscular dystrophy, is 2.4 million bases long and takes almost eight hours to transcribe. We believe that eukaryotic RNA polymerase II backs up when stalled at a DNA lesion and continues after the lesion is repaired, without losing the transcript. Much active research is going on in this area.

In human beings, the autosomal recessive trait *xeroderma pigmentosum* is caused by an inability to repair thymine dimerization induced by UV light. Persons with this trait freckle heavily when exposed to the UV rays of the sun, and they have a high incidence of skin cancer. There are seven complementation groups (loci *XPA-XPG*) whose protein products are involved in the first steps of nucleotide excision repair and whose defects cause xeroderma pigmentosum in human beings. One of them, *XPD,* is a component of TFIIH.

Excision repair triggered by mismatches is referred to as **mismatch repair,** which encompasses about 99% of all DNA repairs. As DNA polymerase replicates DNA, some errors are made that the proofreading polymerase does not correct. For example, a template G can be paired with a T rather than a C in the progeny strand. The GT base pair does not fit correctly in the DNA duplex. The mismatch repair system, which follows behind the replicating fork, recognizes this problem. This system, whose members in *E. coli* are specified by the *mutH, mutL, mutS,* and *mutU* genes, is responsible for the removal of the incorrect base by an excision repair process. (The genes are called *mut* for mutator because mutations of these genes cause high levels of spontaneous mutation in the cells. The *mutU* gene is also known as *uvrD.*) The mismatch repair enzymes initiate the removal of the incorrect base by nicking the DNA strand on one side of the mismatch.

You might wonder how the mismatch repair system recognizes the progeny, rather than the template, base as the wrong one. After all, in a mismatch, there are no defective bases—theoretically, either partner could be the "wrong" base. In *E. coli,* the answer lies in the methylation state of the DNA. DNA methylase, the product of the *dam* locus, methylates 5′-GATC-3′ sequences, which are relatively common in the DNA of *E. coli,* at the adenine residue. Since the mismatch repair enzymes follow the replication fork of the DNA, they usually reach the site of mismatch before the methylase does. Template strands will be methylated, whereas progeny strands, being newly synthesized, will not be. Thus, the methylation state of the DNA cues the mismatch repair enzymes to eliminate the progeny-strand base for repair. After the methylase passes by, both strands of the DNA are methylated, and the methylation cue is gone.

In figure 12.32, we present one model of mismatch repair. The MutS protein, in the form of a homodimer—two

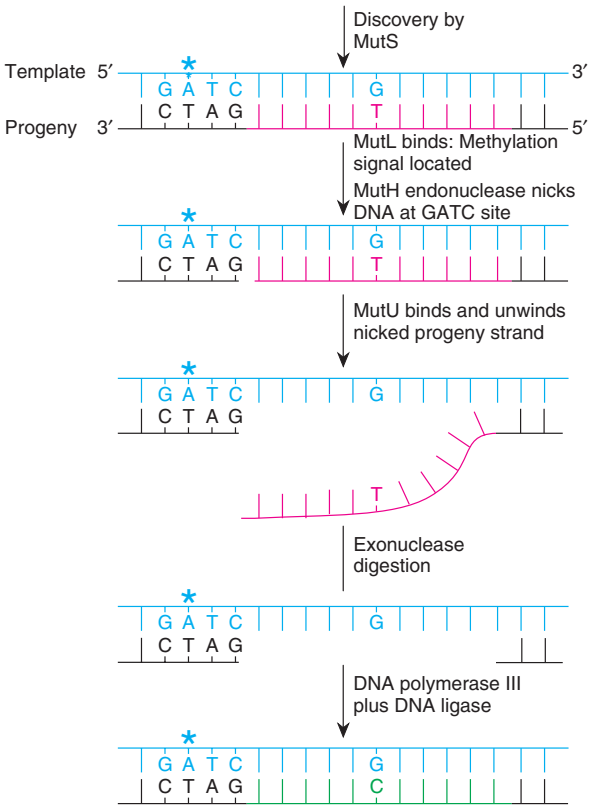


Figure 12.32 Mismatch repair. The MutS protein discovers mismatches; MutL binds and the MutH endonuclease nicks the progeny strand at the 3′-CTAG-5′ sequence. MutU helicase unwinds the nicked oligonucleotide with the mismatch (*red*). Exonuclease digestion, followed by DNA polymerase III and DNA ligase repair, completes the operation.

copies of the same protein—finds the mismatch. MutL, also in the form of a homodimer, then binds, and together they find the methylation signal. They also activate the endonuclease MutH, which then nicks the unmethylated strand at the 3′-CTAG-5′ recognition site, which can be one thousand to two thousand bases away from the mismatch. At the recognition site, the MutS-MutL tetramer loads the helicase MutU (UvrD), which then unwinds the nicked strand. Any one of at least four different exonucleases then attacks the unwound oligonucleotide. DNA polymerase III then repairs the gap, and DNA ligase seals it. This sequence of events highlights a common theme in DNA repair: Once a lesion is found, the damaged DNA has some protein bound to it until the repair is finished.

Our understanding of DNA damage and repair helps provide an answer to an evolutionary question—Why does DNA have thymine while RNA has uracil? If we live in an RNA world, in which RNA evolved first, why don't DNA and RNA both contain uracil? One answer is that a common damage to cytosine, spontaneous deamination, results in uracil. If uracil were a normal base in DNA, the conversion of cytosine to uracil by deamination would not leave any clue to a mismatch repair system that a mutation had occurred. Thus, thymine replaces uracil in DNA, since thymine is not confused with any other normal base in DNA by common spontaneous changes. In fact, cytosine, guanine, adenine, and thymine are not converted simply to any other of the bases in DNA. Hence, changes of these bases leave clues for the repair systems.

Double-Strand Break Repair

Some damage to DNA, such as that caused by ionizing radiation, is capable of breaking both strands of the double helix. When that happens, the cell uses one of two mechanisms to repair the broken ends: It can simply bring the ends back together (a process called *nonhomologous end joining*), or it can use a mechanism that relies on the nucleotide sequences of a homologous piece of DNA, such as a sister chromatid or a homologous chromosome. That method is called *homology-directed recombination.*

In nonhomologous end joining, a protein called Ku, a heterodimer of Ku70 and Ku80, binds to broken chromosomal ends. It then recruits a protein kinase (PK_{CS}); their interaction and the interaction with other proteins is stabilized by a scaffold protein called XRCC4 (for X-ray cross complementation group 4). The complex directs the annealing of the broken ends by DNA ligase IV. No particular sequence information is used, and if more than two broken ends are present, incorrect attachments can take place (e.g., translocations). The second method, homology-directed recombination, involves a second piece of DNA homologous to the broken piece. The method is very similar to our current model of DNA recombination and is discussed in the section entitled "Recombination" later in the chapter.

Postreplicative Repair

When DNA polymerase III encounters certain damage in *E. coli,* such as thymine dimers, it cannot proceed. Instead, the polymerase stops DNA synthesis and, leaving a gap, skips down the DNA to resume replication as far as eight hundred or more bases away. If allowed to remain, this gap will result in deficient and broken DNA. Since part of one strand is absent and the other has damage, there appears to be no viable template for replicating new DNA. However, the cell has two mechanisms to repair this gap: one uses polymerases that can replicate these lesions, and the other is a repair process that uses homologous DNA.

Originally, several proteins were known to facilitate the replication of DNA with lesions; they were believed to interact with the polymerase to make it capable of using damaged DNA as a template. We now know that these proteins are, in fact, polymerases that have the ability to replicate damaged DNA. In *E. coli,* polymerase V can copy damaged DNA. In yeast, polymerases η and ζ, also called REV3/7 and RAD30 polymerases, respectively, can also copy damaged DNA. Some of these polymerases are relatively error free; polymerase V and polymerases η put adenine-containing nucleotides opposite dimerized thymines. However, polymerases ζ and the *E. coli* polymerase IV, which also appears during times of damage, are error prone in their replicative roles. One possible reason for this is that the error-prone polymerases developed by evolutionary processes: They create mutations at a time when the cell might need variability. That is, DNA damage can occur when the environment is stressful for the cell; variability might help the cell survive. As we will see later, the cell can sense DNA damage and act appropriately.

In addition to using repair polymerases, the cell can use a second repair mechanism to replicate damaged DNA when the polymerase leaves a gap. A replication fork creates two DNA duplexes. Thus, an undamaged copy of the region with the lesion exists on the other daughter duplex. A group of enzymes, with one specified by the *recA* locus having central importance, repairs the gap. Since the repair takes place at a gap created by the failure of DNA replication, the process is called **postreplicative repair.** The *recA* locus was originally discovered and named in the recombination process. In fact, postreplicative repair is sometimes called recombinational repair, and it shares many enzymes with recombination.

The RecA Protein

The RecA protein has two major properties. First, it coats single-stranded DNA (fig. 12.33) and causes that coated, single-stranded DNA to invade double-stranded DNA (fig. 12.34). By invasion, we mean that the single-stranded DNA attempts to form complementary base pairs with the antiparallel strand of the double-stranded DNA while displacing the other strand of that double helix. A mechanism for this activity, assuming two sites on the enzyme, appears in figure 12.35. RecA continues to move the single-stranded DNA along the double-stranded DNA until a region of homology is found. The second major property of the RecA protein is that, when stimulated by the presence of single-stranded DNA, it causes autocatalysis of another repressor, called LexA, and thus initiates several sequences of reactions.

The RecA protein is responsible for filling a postreplicative gap in newly replicated DNA with a strand from the undamaged sister duplex. Gap-filling

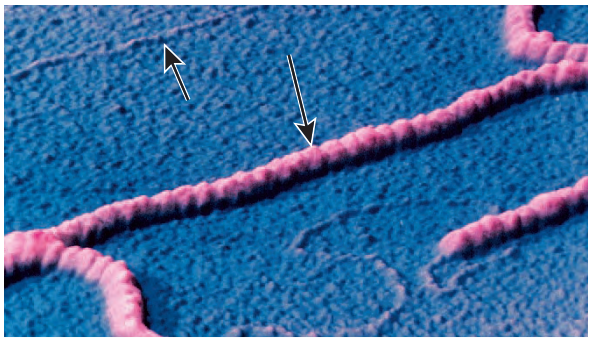


Figure 12.33 Scanning tunneling microscope picture of single-stranded DNA coated with RecA protein (*large arrow*). The *small arrow* indicates uncoated, double-stranded DNA. (In fig. 12.35, we show how the very large, coated DNA can invade the very small, uncoated DNA.) (© Science VU-IBMRL/Visuals Unlimited.)

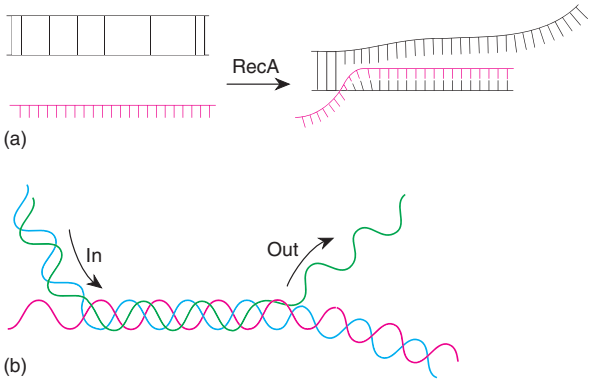


Figure 12.34 One property of the RecA protein causes single-stranded DNA to invade double-stranded DNA and to move along it until a region of complementarity is found. (*a*) Diagrammatic representation of the invasion of the RecA-coated single-stranded DNA. (*b*) A more realistic diagram of the same event. (Reproduced, with permission, from the *Annual Review of Biochemistry,* Volume 61, © 1992 by Annual Reviews, Inc.)

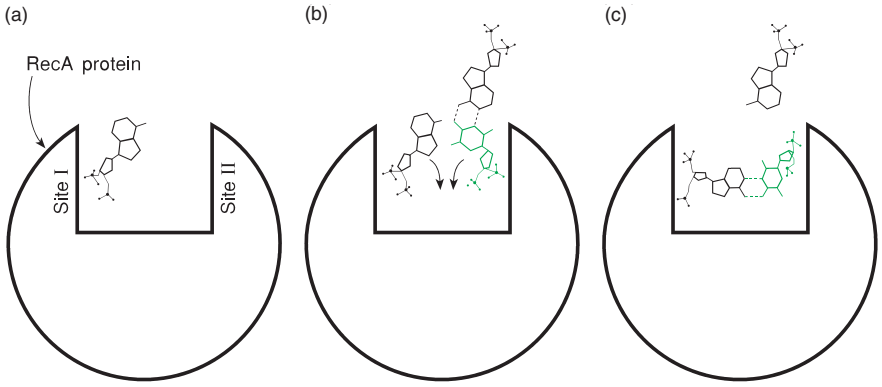


Figure 12.35 A model of how the RecA protein can cause single-stranded DNA to invade a double-stranded molecule. (*a*) Axial view of one nucleotide (with two phosphate groups) of single-stranded DNA attached at site I in this cross-sectional diagram of the RecA protein. The protein is about 60% larger than actually shown. (*b*) Duplex DNA is bound at site II of RecA. (*c*) RecA protein rotates the bases so that the single-stranded DNA forms a complementary base pair with one strand of the duplex, leaving the other strand of the duplex unpaired (see fig. 12.34). (Reprinted with permission from P. Howard-Flanders, et al., "Role of RecA protein spiral filaments in genetic recombination," *Nature,* 309:215–20. Copyright © 1984 Macmillan Magazines, Limited.)

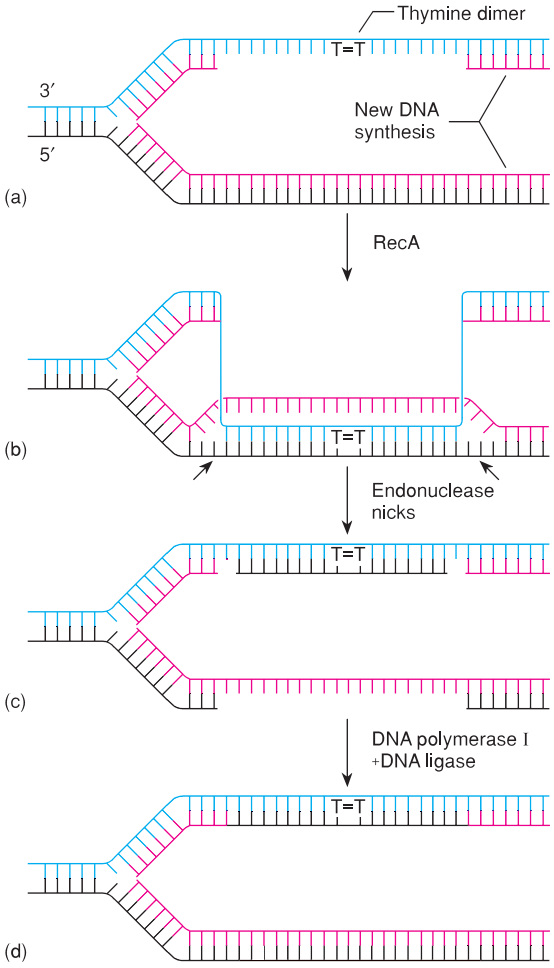


Figure 12.36 RecA-dependent postreplicative DNA repair. DNA polymerase III skips past a thymine dimer during DNA replication (*a*). With the help of RecA, the single strand with the thymine dimer invades the normal sister duplex (*b*). An endonuclease nicks the new duplex at either side of the thymine dimer site, freeing the new duplex with the thymine dimer and leaving the sister duplex single-stranded (*c*). Repair enzymes then create two intact daughter duplexes (*d*).

processes then complete both strands. In figure 12.36*a,* we see a replication fork with a gap in the progeny strand in the region of a thymine dimer. The RecA protein is responsible for the damaged single strand invading the sister duplex (fig. 12.36*b*). Endonuclease activity then frees the double helix containing the thymine dimer (fig. 12.36*c*). DNA polymerase I and DNA ligase return both daughter helices to the intact state (fig. 12.36*d*). The thymine dimer still exists, but now its duplex is intact, and another cell cycle is available for photoreactivation or excision repair to remove the dimer.

The SOS Response

Postreplicative repair is part of a cell reaction called the **SOS response.** When an *E. coli* cell is exposed to excessive quantities of UV light, other mutagens, or agents that damage DNA (such as alkylating or cross-linking agents), or when DNA replication is inhibited, gaps are created in the DNA. In the presence of this single-stranded DNA, the RecA protein interacts with the LexA protein, the product of the *lexA* gene. The LexA protein normally represses about eighteen genes, including itself. The other genes include *recA, uvrA, uvrB, and uvrD;* two genes that inhibit cell division, *sulA* and *sulB;* and several others. Each of these genes has a consensus sequence in its promoter called the **SOS box:** 5′-CTGX_{10}CAG-3′ (where X_{10} refers to any ten bases). The LexA protein normally binds at the SOS box, limiting the transcription of these genes. When single-stranded DNA activates RecA, RecA interacts with the LexA protein to trigger the autocatalytic properties of LexA (fig. 12.37). Transcription then follows from all the genes having an SOS box. The two inhibitors of cell division, the products of the *sulA* and *sulB* genes, presumably increase the amount of time the cell has to repair the damage before the next round of DNA replication.

Eventually, the DNA damage is repaired. There is no single-stranded DNA to activate RecA, and, therefore, LexA is no longer destroyed. LexA again represses the suite of proteins involved in the SOS response, and the SOS response is over. Table 12.5 summarizes some of the enzymes and proteins involved in DNA repair.

As we will mention in chapter 14, λ prophage can be induced into vegetative growth by UV light. This is another effect of the SOS response. RecA not only causes the LexA protein to be inactivated, but also directly inactivates the λ repressor, the product of the λ *cI* gene. From an evolutionary point of view, it makes sense for phage λ to have evolved a repressor protein that the RecA protein inactivates. As a prophage, λ is dependent on the survival of the host cell. When that survival might be in jeopardy, the prophage would be at an advantage if it could sense the danger and make copies of itself that could leave the host. One of these times might be when the host has suffered a lot of DNA damage. The SOS response is a signal to a prophage that the cell has received that damage. Hence, the prophage is induced when RecA acts as a protease; the λ repressor is destroyed, the cro protein becomes dominant, and vegetative growth follows. From an evolutionary perspective, the *E. coli* cell has not created an enzyme (RecA) that seeks out the λ repressor for the benefit of λ. Rather, the λ repressor has evolved for its own advantage to be sensitive to RecA.

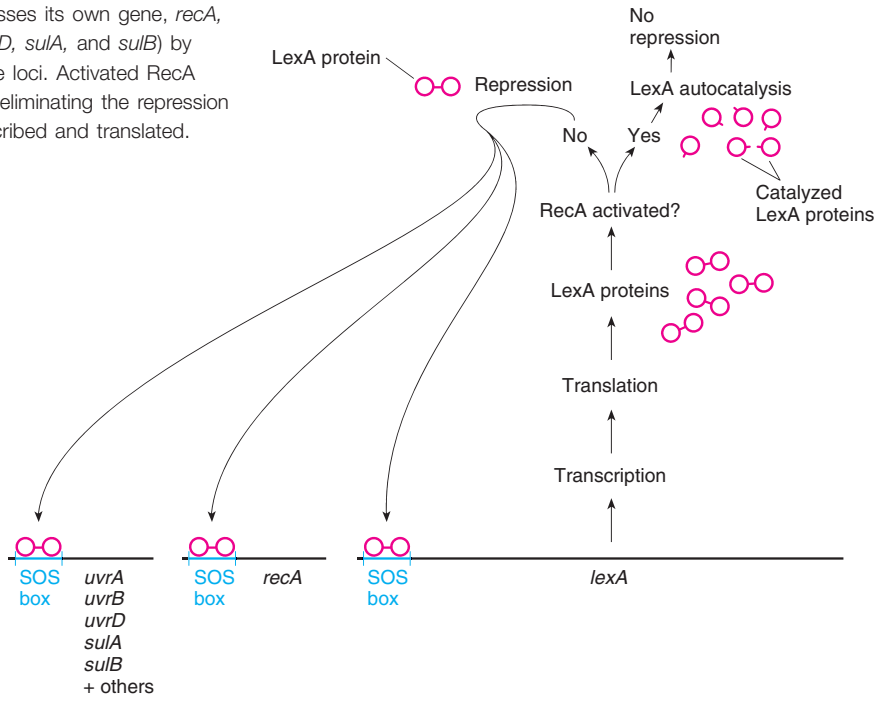


Figure 12.37 The LexA protein represses its own gene, *recA,* and several other loci (*uvrA, uvrB, uvrD, sulA,* and *sulB*) by binding at the SOS box in each of the loci. Activated RecA protein causes autocatalysis of LexA, eliminating the repression of all these loci, which are then transcribed and translated.

RECOMBINATION

Although recombination, the nonparental arrangement of alleles in progeny, can come about both by independent assortment and crossing over, we are concerned here with recombination due to crossing over between homologous pieces of DNA **(homologous recombination).** We briefly discuss transpositional recombination in chapter 14 and site-specific recombination (e.g., λ integration) in chapters 7, 14, and 16.

Recombination is a **breakage-and-reunion** process. Homologous parts of chromosomes come into apposition and are then reconnected in a crosswise fashion (see fig. 6.4). This general model fits what we know about the concordance of recombination and repair: Both involve breakage of the DNA and a small amount of repair synthesis, and both involve some of the same enzymes.

Double-Strand Break Model of Recombination

In 1964, R. Holliday suggested a model of homologous recombination that involved simultaneous breaks in one strand each of the two double helices that were to cross over. In 1983, J. Szostak and colleagues put forth a different model, initiated by a double-strand break in one of the double helices. At first, this model was not considered seriously because a double-strand break was thought too dangerous a DNA lesion for cellular enzymes to create. However, we now know that the double-strand break model is generally correct, and we refer to the **Holliday junction** for an intermediate stage in the process. The model depends on DNA complementarity between the recombining molecules and is thus a model of great precision.

We begin with two double helices lined up as they would be, for example, in a meiotic tetrad, ready to undergo recombination (fig. 12.38*a*). The first step of the process is a double-stranded break in one of the double helices. In eukaryotes, the protein Spo11 accomplishes

R. Holliday (1932–). (Courtesy of James L. German, III, M.D.)

Table 12.5 **Some of the Enzymes and Proteins Involved in DNA Repair in *E. coli,* Not Including DNA Polymerase I and III, DNA Ligase, and Single-Strand Binding Proteins**

Enzyme	Gene	Action
Damage Reversal		
DNA photolyase	*phr*	Undimerizes thymine dimers
DNA methyltransferase	*ada*	Demethylates guanines in DNA
Base Excision Repair		
Uracil-DNA glycosylase	*ung*	Removes uracils from DNA
Endonuclease IV	*nfo*	Nicks AP sites on the 5′ side
Exonuclease, lyase, or phosphodiesterase	*several*	Removes base-free nucleotide
Nucleotide Excision Repair		
UvrA	*uvrA*	With UvrB, locates thymine dimers and other distortions
UvrB	*uvrB*	Nicks DNA on the 3′ side of the lesion
UvrC	*uvrC*	Nicks DNA on the 5′ side of the lesion
UvrD (helicase II)	*uvrD*	Unwinds oligonucleotide
Mismatch Repair		
MutH	*mutH*	Nicks DNA at recognition sequence
MutL	*mutL*	Recognizes mismatch
MutS	*mutS*	Binds at mismatch
MutU (UvrD)	*mutU*	Unwinds oligonucleotide
Exonucleases	*recJ, xseA, sbcB*	Degrades unwound oligonucleotide
DNA methylase	*dam*	Methylates 5′-GATC-3′ DNA sequences
Double-Strand Break Repair		
Ku	*Ku70, Ku80*	Binds to broken chromosomal ends
PK_{CS}	PK_{CS}	Protein kinase
DNA ligase IV	*LIG4*	Ligates broken ends of DNA
XRCC4	*XRCC4*	Stabilization protein
Postreplicative Repair		
Polymerase IV	*DinB*	DNA polymerase
Polymerase V	*UmuC, UmuD*	DNA polymerase
Polymerase η	*RAD30*	DNA polymerase
Polymerase ζ	*REV3, REV7*	DNA polymerase
RecA	*recA*	Single-stranded DNA invades double-stranded DNA; causes LexA to autocatalyze; protease
LexA	*lexA*	Represses SOS proteins
SulA, SulB	*sulA, sulB*	Inhibit cell division

this. The break is followed by 5′ → 3′ exonuclease activity to widen the gaps formed in the double helix and create 3′ single-stranded tails (fig. 12.38*b, c*). These tails are coated with RecA protein that then catalyzes the invasion of one of the single strands into the intact double helix in direct apposition (fig. 12.38*d*). Repair of single-stranded DNA by DNA polymerase I and DNA ligase then replaces sections of previously digested DNA (fig. 12.38*e*). At this point, there is no "lost" genetic material; however the two double helices are interlocked and need to be freed of each other. Before that happens, however, **branch migration** can take place, a process in which the crossover point can slide down the duplexes (fig. 12.38*f*). In *E. coli,* the RuvAB complex, the product of the *ruvA* and *ruvB* genes that together form an ATP-dependent motor, moves the junction point

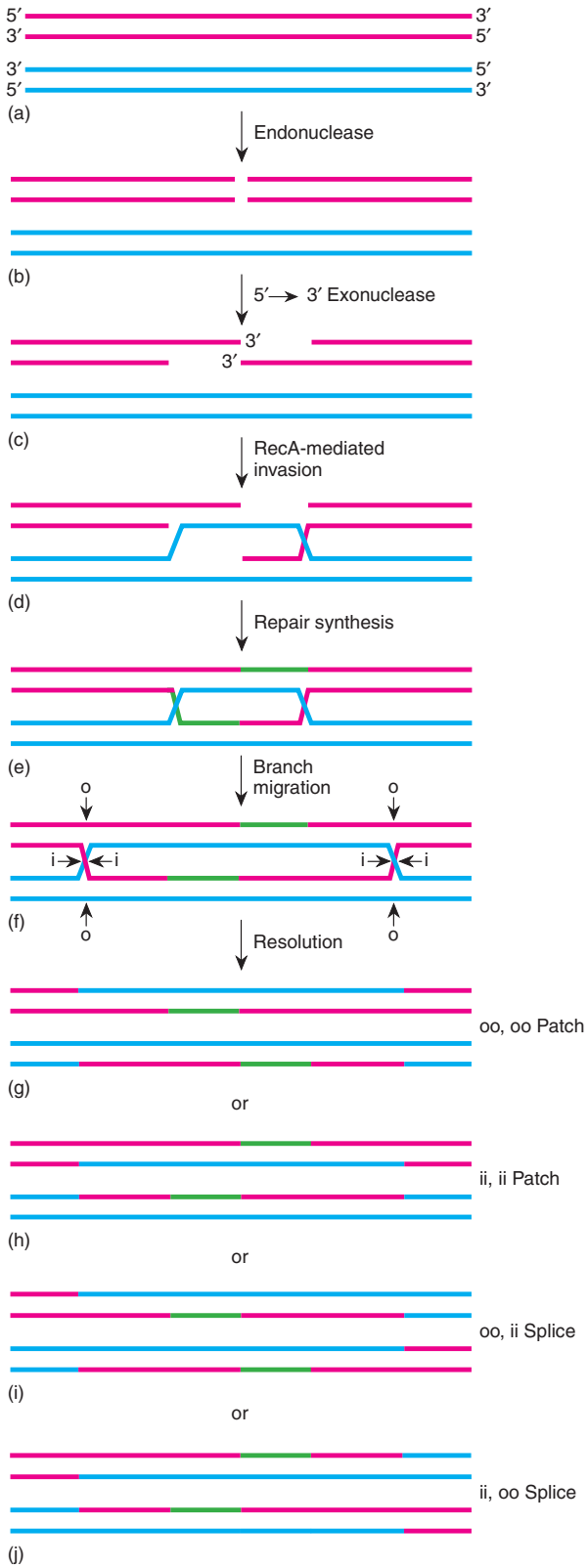


Figure 12.38 The double-strand break model of genetic recombination. Two homologous duplexes (*a: red, blue*) of the four present in a meiotic tetrad are shown. An endonuclease creates a double-stranded break in one of the duplexes (*b*). A 5′→ 3′ exonuclease then digests away from the break in both directions, creating 3′ tails (*c*). RecA-mediated invasion occurs in the second duplex (*d*), followed by repair synthesis to close all gaps (*e*). Branch migration then takes place (*f*). Each of the Holliday junctions is then resolved independently, either by nicks in the two outer strands (*o*) or the two inner strands (*i*). Therefore, four resolution structures are possible (*g–j*). In patches, the ends of each duplex are the same as the original, indicating that there may not be recombination for loci flanking the point of crossover. In splices, the ends of each duplex have recombined, indicating that flanking loci may have crossed over. (Reprinted from *Cell*, Vol. 87, Frank Stahl, "Meiotic Recombination in Yeast: Coronation of the Double-Stranded-Break Repair Model," pp. 965–968, Copyright © 1996, with permission from Elsevier Science.)

(fig. 12.39). As the junction points move, they create heteroduplex DNA, places where the two strands of each double helix come from different original helices. These stretches have the potential to produce mismatches where the two chromatids differed originally. To resolve the cross-linked duplexes, a second cut at each junction is required.

Each of the two crossover points is a Holliday junction. If we open these junctions, we can see that each can be resolved in two different ways. (RuvC endonuclease, the protein product of the *ruvC* gene, resolves the Holliday junctions in *E. coli;* see fig. 12.39. RuvC cuts the Holliday junction at the consensus sequence 5′[A or T]TT[G or C]-3′. The cut is on the 3′ side of the two thymines.) Since there are two Holliday junctions per crossover, there are four potential combinations, as shown in figure 12.38*g–j.* Some of these combinations produce patches, where no recombination takes place among loci to the sides of the hybrid piece. Other combinations produce splices, where reciprocal recombination of loci takes place at the ends. The Holliday junctions can be seen in the electron microscope (fig. 12.40). Note that homology-directed recombination to repair double-strand breaks is similar to the process shown in figure 12.38.

Bacterial Recombination

In bacterial recombination, a linear molecule recombines with a circular molecule (see fig. 7.15). Usually, invading DNA originates in the linear molecule. The RecBCD protein, whose subunits are the products of the *recB, recC,* and *recD* loci, initiates the first steps in forming an invading linear DNA molecule. RecBCD is a helicase, an exonuclease, and an endonuclease.

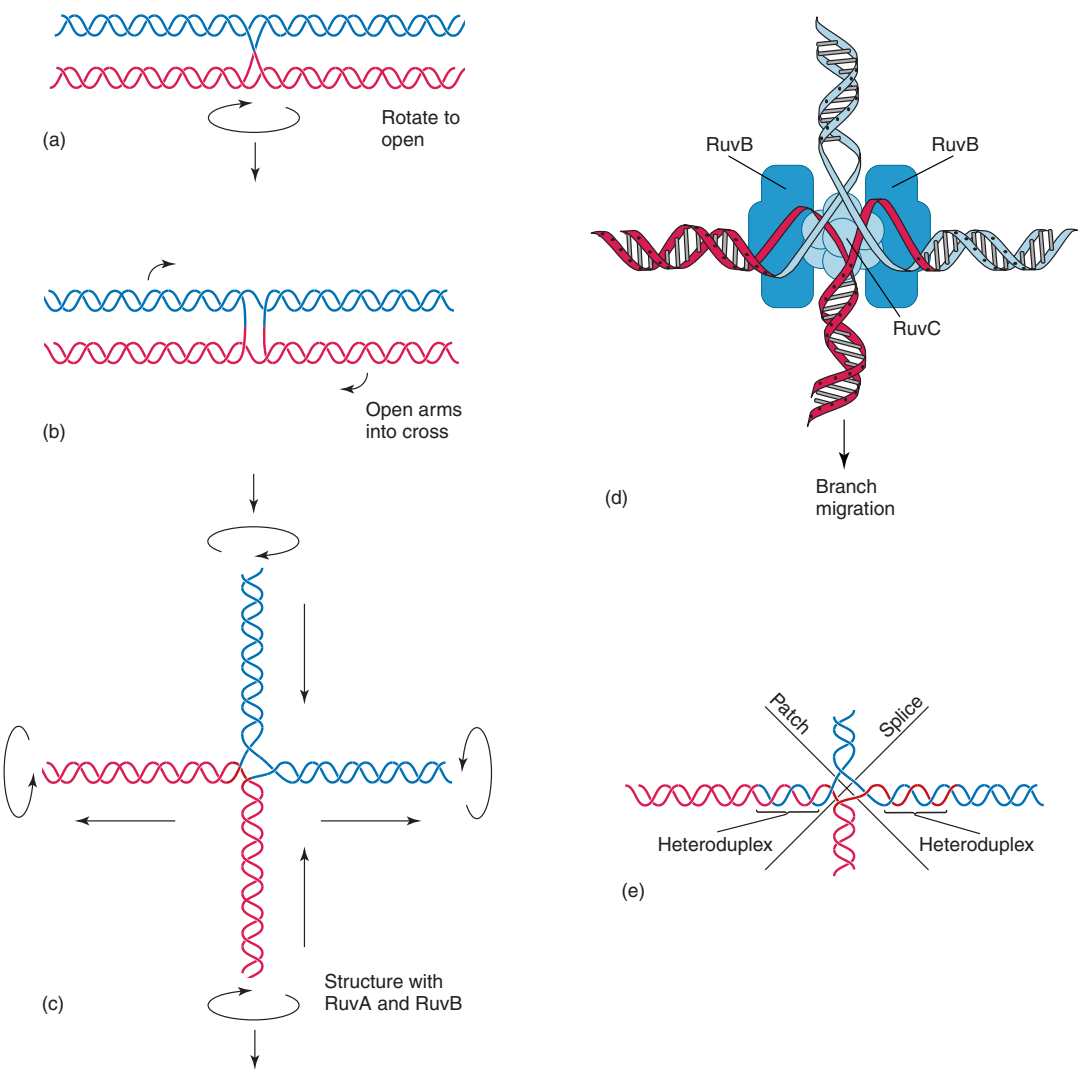


Figure 12.39 Branch migration at a Holliday junction. (*a*) Two double helices (*red, blue*) are connected by a crossover. (*b*) The structure opens when one of the double helices rotates. (*c*) We further clarify the structure by separating the arms into an open cross, showing the direction in which the arms move during branch migration (*arrows:* the equivalent of pulling out on the left and right arms, drawing in the top and bottom arms). (*d*) A more realistic drawing with the RuvA and RuvB proteins, indicating one of the RuvA tetramers behind the center of the cross. A second tetramer (not shown) is located above the cross center, forming a RuvA sandwich of the cross center. The RuvB hexamers are shown on either side of the cross. (*e*) RuvC can resolve the cross to form either a splice or a patch, depending on which cut is made. (Reprinted with permission from *Nature,* Vol. 374, C. Parsons, et al., "Structure of a Multisubunit Complex that Promotes DNA Branch Migration." Copyright © 1995 Macmillan Magazines Limited.)

The RecBCD protein enters a DNA double helix from one end and travels along it in an ATP-dependent process. As it travels along the DNA, it acts as a 3′ → 5′ exonuclease, degrading one strand of the linear double helix (fig. 12.41). This process continues until RecBCD comes to a **chi site,** the sequence 5′-GCTGGTGG-3′, which appears about a thousand times on the *E. coli* chromosome. RecBCD's recognition of that sequence attenuates its 3′ → 5′ exonuclease activity and enhances its 5′ → 3′ exonuclease activity, begun after an endonucleolytic cleavage. From that point on, RecBCD creates a 3′ overhang or tail. That tail is coated by RecA and then

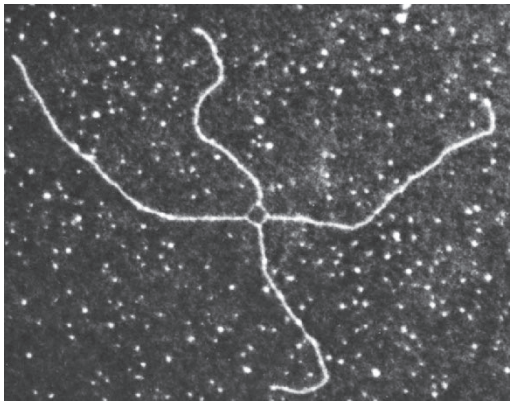


Figure 12.40 A Holliday intermediate structure, equivalent to the structure seen in figure 12.39*c*. Each arm is about 1 micron long. (H. Potter and D. Dressler. "DNA recombination: In vivo and in vitro studies," *Cold Spring Harbor Symposium on Quantitative Biology,* Volume XLIII, 1979, pp. 969–85.)

invades the circular bacterial chromosome to initiate a crossover event (fig. 12.41). After this pairing, the unpaired segments of the double helix of the bacteria and the exogenote are both degraded. Finally, DNA ligase seals the circular double helix. The resulting hybrid DNA will then be open to mismatch repair that can restore either original base pairs or base pairs from the invading DNA (see below).

Hybrid DNA

The result of bacterial recombination or meiotic recombination with branch migration is a length of hybrid DNA. This **hybrid DNA,** also called **heterozygous DNA** or **heteroduplex DNA,** has one of two fates, if we assume a difference in base sequences in the two strands. Either the heteroduplex can separate unchanged at the next cell division, or the cell's mismatch repair system can repair it (fig. 12.42). Without appropriate methylation cues, the mismatch repair system can convert the CA base pair to either a CG or a TA base pair. If TA were the original bacterial base pair, conversion to CG would be a successful recombination, whereas return of the CA to TA would be restoration rather than recombination.

Recombination in yeast, or any other eukaryote, generates two heteroduplexes. The repair process can cause **gene conversion** (fig. 12.43), the alteration of progeny ratios indicating that one allele was converted to another, a phenomenon seen in up to 10% of yeast asci. The mismatched AC will be changed to an AT or a GC base pair; the mismatched TG base pair will be changed to TA or CG. The result of the repair, as shown at the bottom of figure 12.43, can be gene conversion in which an ex-

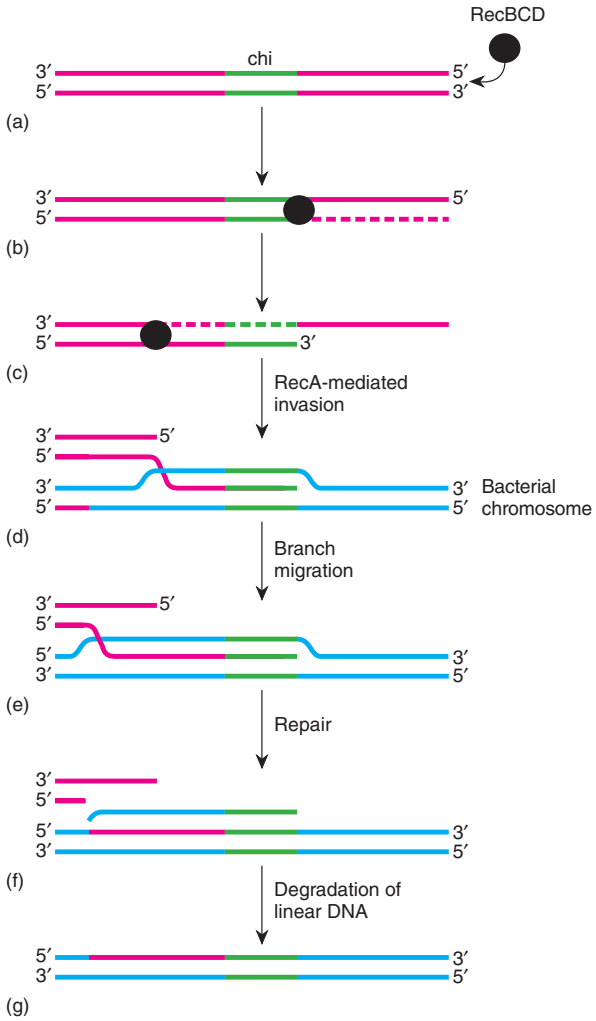


Figure 12.41 RecBCD enters a linear DNA double helix (*red*) at one end and travels along it, digesting the 3′ strand. When the protein encounters a chi site (*green*), it cuts the other strand and begins acting as a 3′ → 5′ exonuclease, creating a 3′ overhang (*b, c*). The 3′ overhang can then invade a double helix mediated by RecA. Repair and degradation of the linear DNA results in hybrid DNA in the bacterial chromosome, which can be fixed by the mismatch repair system.

pected ratio of 2:2 ($a^- a^- a^+ a^+$) is converted to a 3:1 ratio ($a^- a^- a^- a^+$) or a 1:3 ratio ($a^- a^+ a^+ a^+$). If the heteroduplexes are not repaired, then a single cell generates both kinds of offspring after one round of DNA replication. Thus, the colony from the cell will be half wild-type (a^+) and half mutant (a^-). We see this phenomenon only in an Ascomycete fungus, such as yeast, in which all the products of a single meiosis remain together.

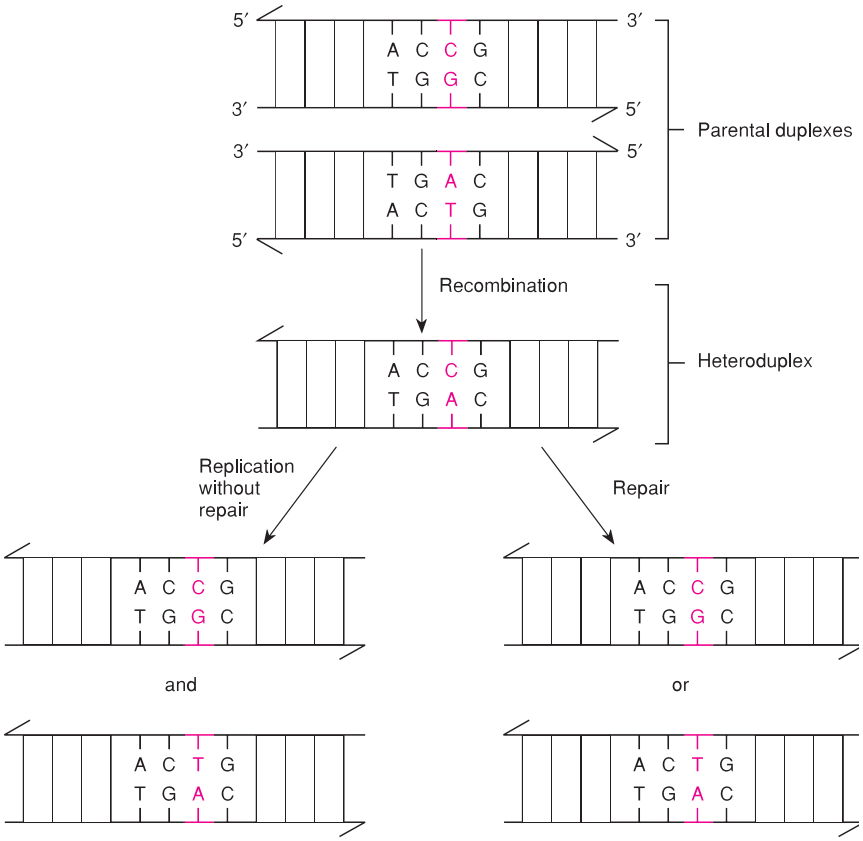


Figure 12.42 Fate of a heteroduplex DNA. *Recombination* results in heteroduplex DNA with mismatched bases. *Replication without repair* produces two different daughter molecules. *Repair* converts the mismatched base pair to one or the other normal base pair.

SUMMARY

STUDY OBJECTIVE 1: To look at the nature of mutation in prokaryotes 316–317

In 1943, Luria and Delbrück demonstrated that bacterial changes are true mutations similar to mutations in higher organisms. They showed that a high variability occurs in the number of mutants in small cultures as compared with the number of mutants in repeated subsamples of a large culture. Mutations thus occur spontaneously; they are caused by mutagens, which include chemicals and radiation. This chapter is concerned primarily with point mutations rather than changes in whole chromosomes or chromosomal parts.

STUDY OBJECTIVE 2: To analyze functional and structural allelism and examine the mapping of mutant sites within a gene 317–324

Allelism is defined by the *cis-trans* complementation test. Complementation implies independent loci, or nonallelic genes. Lack of complementation implies allelism. Functional alleles that differ from each other at the same nucleotides are also called structural alleles. Benzer used T4 phages to do fine-structure studies using complementation testing and deletion mapping.

STUDY OBJECTIVE 3: To verify the colinearity of gene and protein 324–325

Yanofsky demonstrated colinearity of the gene and protein. He had the advantage of working with a gene whose protein product was known.

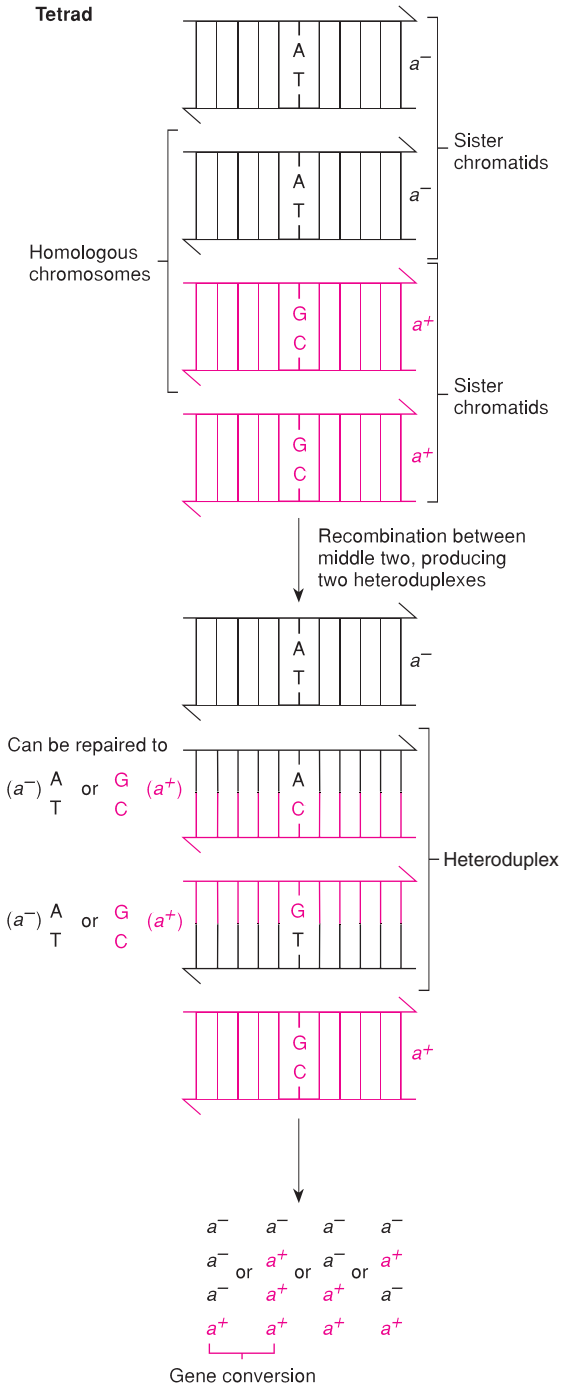


Figure 12.43 Recombination and repair can cause gene conversion. During recombination, heteroduplex DNA is formed, containing mismatched base pairs. Without methylation cues, repair enzymes convert the mismatch to a complementary base pair, in a random fashion—that is, an AC base pair can be converted to either an AT (a^- allele) or a GC (a^+ allele) base pair. Two of the four possible repair choices create 3:1 ratios of alleles rather than the expected 2:2 ratios in the offspring. The 3:1 ratio represents gene conversion.

STUDY OBJECTIVE 4: To study mutagenesis 325–338

After a mutation, the normal phenotype, or an approximation of it, can be restored either by back mutation or suppression. Intragenic suppression occurs when a second mutation within the same gene causes a return of normal or nearly normal function. Intergenic suppression occurs when a second mutation happens, usually in a transfer RNA gene, that counteracts the original mutation. Nonsense, missense, and frameshift mutations can all be suppressed.

Spontaneous mutation probably occurs primarily because of tautomerization of the bases of DNA. If a base is in the rare form during DNA replication, it can form unusual or mutant base pairings. We describe the mechanisms of action of the most common mutagens.

STUDY OBJECTIVE 5: To investigate the processes of DNA repair and recombination 339–352

DNA repair processes can be divided into four categories: damage reversal, excision repair, double-strand break repair, and postreplicative repair. Photoreactivation is an example of damage reversal. Thymine dimers are undimerized by a photolyase enzyme in the presence of light energy. Excision repair removes a damaged section of a DNA strand. Repair enzymes fill in the gap. Excision repair can be divided into three types. In base excision repair, bases are removed by environmental causes or by glycosylases that sense damaged bases. AP endonuclease and an exonuclease, phosphodiesterase, or lyase then removes the base-free nucleotide. Some enzymes use base flipping to gain access to nucleotides in the double helix.

In nucleotide excision repair, enzymes in the Uvr system remove a patch containing the lesion, usually a thymine dimer. In mismatch repair, enzymes of the Mut system use methylation cues to remove a progeny patch containing the mismatch.

Double-strand break repair relies on one of two mechanisms. In nonhomologous end joining, the cell simply brings the broken ends back together. In homology-directed recombination, the cell repairs the broken ends using a recombinational mechanism.

Postreplicative repair fills in gaps left by DNA polymerase III. Some polymerase enzymes can use lesions as templates. Otherwise, the RecA protein is central to the process. A single strand from the undamaged duplex is

used to fill the gap in the damaged duplex. Single-stranded DNA induces the SOS response, which temporarily eliminates LexA-mediated repression.

Recombination in eukaryotes begins with a double-stranded break in one double helix, followed by invasion of one of the ends into the other double helix. Repair, ligation, and branch migration follow. The crossover points, called Holliday junctions, need to be resolved, resulting in patches and splices. In *E. coli,* the RecBCD protein invades linear DNA, creating tails for invasion of the circular bacterial chromosome. Recombination results in heteroduplex DNA, which, if repaired, can lead to gene conversion. Thus, a battery of enzymes within the cell can modify DNA. These enzymes serve in DNA replication, repair, and recombination.

SOLVED PROBLEMS

PROBLEM 1: An investigator isolates two recessive wing mutants of *Drosophila melanogaster.* The flies differ in wing vein pattern. Are the mutations that cause these variants allelic?

Answer: To verify allelism, the investigator must create a heterozygote of the two mutations by either mating the flies, if they are of opposite sexes, or breeding each mutant into a separate stock for new matings. If the heterozygotes are of the wild-type, then the mutations are not allelic. If the heterozygote has a mutant phenotype, then we presume the mutations are functional alleles. (Allelism should be verified in females to be sure that the locus is not on the X chromosome, since males have only one.) If the mutations are functional alleles, it is possible to determine whether they are also structural alleles by looking for wild-type offspring of the heterozygote. If they occur at a rate higher than the background mutation rate, the alleles are not structural alleles. If wild-type offspring occur at the mutation rate, the alleles are presumably structural.

PROBLEM 2: What is the difference between mismatch repair and AP repair?

Answer: Both processes are similar in that they entail removal of an incorrect base in a DNA double helix by an excision process followed by a repair process. The processes differ in the event that triggers them. Mismatch repair is triggered by a base pair that does not occupy the correct space in the double helix—that is, by a non-Watson and Crick pairing (not AT or GC). AP repair is triggered by enzymes that recognize a missing base.

PROBLEM 3: What role does the RecBCD protein play in recombination?

Answer: For recombination to take place in *E. coli,* a single strand of DNA from the exogenote must insinuate itself into the chromosomal double helix with the help of the RecA protein. It is the RecBCD protein that creates the single-stranded DNA. It does so by traveling down the double helix, creating a single-strand tail in its wake. At chi sites, it switches the activity of the enzyme from a $3' \rightarrow 5'$ exonuclease to a $5' \rightarrow 3'$ exonuclease, creating a $3'$ tail. RecA can then act on this single-stranded tail to initiate recombination.

EXERCISES AND PROBLEMS*

MUTATION

1. Construct a data set that Luria and Delbrück might have obtained that would prove the mutation theory wrong.
2. What types of enzymatic functions are best studied using temperature-sensitive mutations?
3. Seven arginine-requiring mutants of *E. coli* were independently isolated. All pairwise matings were done (by transduction) to determine the number of loci (complementation groups) involved. If a (+) in the following figure indicates growth and a (−) no growth on minimal medium, how many complementation groups are involved? Why is only "half" a table given? Must the upper left to lower right diagonal be all (−)?

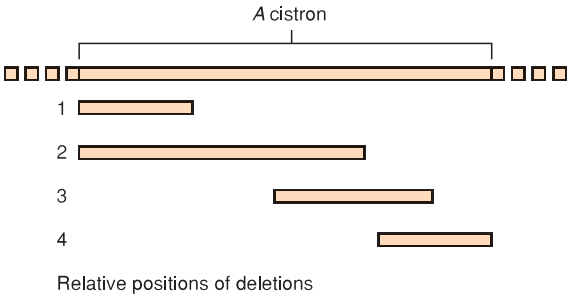
	1	2	3	4	5	6	7
1	–	+	+	+	+	–	–
2		–	+	+	–	+	+
3			–	–	+	+	+
4				–	+	+	+
5					–	+	+
6						–	–
7							–

4. Several *r*II mutations (M to S) have been localized to the *A* cistron because of their failure to complement with a known deletion in that cistron. The phages carrying these mutations are then mated pairwise with the following series of subregion deletions. The mating is done on *E. coli* B and plated out on *E. coli*

*Answers to selected exercises and problems are on page A-13.

K12 (see the figure). A (+) shows the presence of plaques on K12, whereas a (−) shows an absence of growth. Look at the deletion map of the area and localize each of the *r*II *A* mutations on this map.

	Deletion			
Mutant	**1**	**2**	**3**	**4**
M	+	+	–	+
N	+	–	–	+
O	+	+	+	–
P	+	+	–	–
Q	+	–	+	+
R	–	–	+	+
S	–	–	–	+



5. A *Drosophila* worker isolates four eye-color forms of the fly: wild-type, white, carmine, and ruby. (The worker does not know that white, carmine, and ruby are three separate loci on the X chromosome.) What crosses could the researcher make to determine allelic relations of the genes? What results would be expected? A new mutant, eosin, is isolated. What crosses should be carried out to determine that eosin is an allele of white?

6. Define structural and functional alleles. What is the *cis* part of a *cis-trans* complementation test?

7. Did Benzer and Yanofsky work with genes that had intervening sequences? What relevance might introns have to their work?

8. How can intra-allelic complementation result in incorrect conclusions about allelism?

9. *E. coli* bacteria of strain K12 are lysogenic for phage λ. Why won't *r*II mutants of phage T4 grow in these bacteria?

10. Diagram the tautomeric base pairings in DNA. What base pair replacements occur because of the shifts?

11. What is the difference between a substrate and a template transition mutation?

12. Describe two mechanisms for transversion mutagenesis.

13. 5-bromouracil, 2-aminopurine, proflavin, ethyl ethane sulfonate, and nitrous acid are chemical mutagens. What does each do?

14. A point mutation occurs in a particular gene. Describe the types of mutational events that can restore a functional protein, including intergenic events. Consider missense, nonsense, and frameshift mutations.

15. Why does misalignment result in addition or deletion of bases?

16. What are the differences and similarities between intergenic and intragenic suppression?

17. Eight independent mutants of *E. coli* requiring tryptophan (*trp*) are isolated. Complementation tests are performed on all pairwise combinations. Based on the results shown, determine how many genes you have identified and which mutants are in which genes (+ = complementation; − = no complementation).

	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**
1	–	+	–	–	+	+	+	–
2		–	+	+	–	+	+	+
3			–	–	+	–	–	–
4				–	+	+	+	–
5					–	+	+	+
6						–	–	+
7							–	+
8								–

18. Complementation tests are usually done with recessive mutations for if the mutations were dominant, all progeny would be mutant, regardless of whether the genes are allelic. Suppose you have isolated in a diploid species two independent dominant mutations that each confer resistance to the drug cycloheximide. Call these mutations *Chx-1* and *Chx-2*. What crosses can you perform to determine whether the mutations are allelic? Your crosses should allow you to determine whether the mutations are allelic, nonallelic and unlinked, or nonallelic and linked.

19. A series of overlapping deletions in phage T4 are isolated. All pairwise crosses are performed, and the progeny scored for wild-type recombinants. In the following table, + = wild-type progeny recovered; − = no wild-type progeny recovered.

	1	**2**	**3**	**4**	**5**
1	–	+	–	–	–
2		–	+	+	–
3			–	+	+
4				–	–
5					–

a. Draw a deletion map of these mutations.
b. A point mutation, 6, is isolated and crossed with all of the deletion strains. Wild-type recombinants are recovered only with strains 2 and 3. What is the location of the point mutation?

20. Hydroxylamine is a chemical that causes exclusively C → T transition mutations. Can hydroxylamine reverse nonsense mutations? Explain.

21. A nonsense suppressor is isolated and shown to involve a tyrosine transfer RNA. When this mutant transfer RNA is sequenced, the anticodon turns out to be normal, but a mutation is found in the dihydrouridine loop. What does this finding suggest about how a transfer RNA interacts with the messenger RNA?

22. Devise selection-enrichment procedures for isolating the following kinds of mutants:
a. extra-large bacterial cells
b. nonmotile ciliated protozoans

23. Two chemically induced mutants, *x* and *y,* are treated with the following mutagens to see if revertants can be produced: 2-amino purine (2AP), 5-bromouracil (5BU), acridine dye (AC), hydroxylamine (HA), and ethylmethanesulfonate (EMS). In the following table, + = revertants and − = no revertants. For each mutation, determine the probable base change that occurred to change the wild-type to the mutant.

	Chemical				
Mutant	**2AP**	**5BU**	**AC**	**HA**	**EMS**
x	−	+	−	+	+
y	+	−	−	−	−

24. What situation will lead to a false positive in a complementation test—or, in other words, indicate two genes when, in fact, the mutations are in the same gene?

25. What situation will lead to false negatives in a complementation test—or, in other words, indicate mutations are in the same gene when, in fact, they are in different genes?

26. Suppose you repeat the Luria-Delbrück fluctuation test, but this time you look for *lac* colonies. Your "individual cultures" produce the following numbers of lac colonies: 20, 25, 22, 18, 24, 19, 17, 25, 26, and 18. Subsamples from the bulk culture give results identical to these. What can you conclude?

27. You have isolated a new histidine auxotroph, and, despite all efforts, you cannot produce any revertants. What probably happened to produce the original mutant?

DNA REPAIR

28. UV light causes thymine dimerization. Describe the mechanisms, in order of efficiency, that can repair the damage. Name the enzymes involved.

29. What types of damage do excision repair endonucleases recognize?

30. What are the functions of the RecA protein? How is it involved in phage λ induction? (*See also* RECOMBINATION)

RECOMBINATION

31. Diagram, in careful detail, a recombination by way of the double-strand break model. What enzymes are required at each step?

32. What are the different enzymes involved in reciprocal and nonreciprocal recombination?

CRITICAL THINKING QUESTIONS

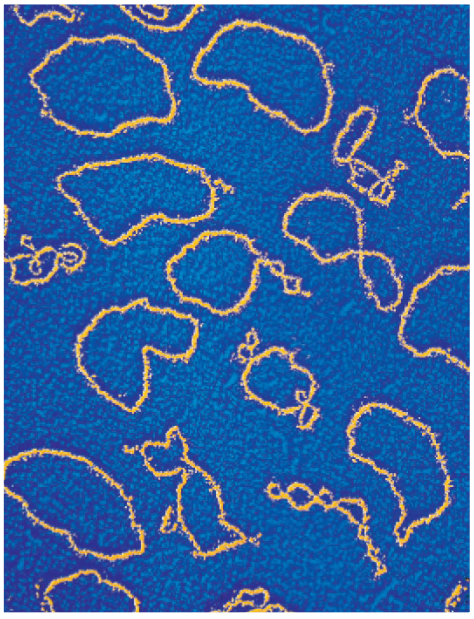
1. Charles Yanofsky demonstrated colinearity of the gene and its protein product. What are the alternatives to colinearity?

2. Comment on the statement that DNA is a molecule designed for replication and repair.

Suggested Readings for chapter 12 are on page B-10.

13

GENOMICS, BIOTECHNOLOGY, AND RECOMBINANT DNA



Artificially colored transmission electron micrograph of DNA plasmids from the bacterium *Escherichia coli*. These plasmids are used in genetic engineering.

(© Dr. Gopal Murti/SPL/Photo Rsearchers, Inc.)

STUDY OBJECTIVES

1. To look at the techniques of gene cloning 358
2. To examine the techniques of creating restriction maps 377
3. To study the methods of DNA sequencing 383
4. To look at the goals and methods of the Human Genome Project 390
5. To look at the practical benefits and human issues of genetic engineering 397

STUDY OUTLINE