DNA REPLICATION MECHANISMS

All organisms must duplicate their DNA with extraordinary accuracy before each cell division. In this section, we explore how an elaborate "replication machine" achieves this accuracy, while duplicating DNA at rates as high as 1000 nucleotides per second.

Base-Pairing Underlies DNA Replication and DNA Repair

As introduced in Chapter 1, *DNA templating* is the mechanism the cell uses to copy the nucleotide sequence of one DNA strand into a complementary DNA sequence (**Figure 5–2**). This process entails the recognition of each nucleotide in the DNA *template strand* by a free (unpolymerized) complementary nucleotide, and it requires the separation of the two strands of the DNA helix. This separation exposes the hydrogen-bond donor and acceptor groups on each DNA base for base-pairing with the appropriate incoming free nucleotide, aligning it for its enzyme-catalyzed polymerization into a new DNA chain.

The first nucleotide-polymerizing enzyme, **DNA polymerase**, was discovered in 1957. The free nucleotides that serve as substrates for this enzyme were found to be deoxyribonucleoside triphosphates, and their polymerization into DNA required a single-stranded DNA template. **Figure 5–3** and **Figure 5–4** illustrate the stepwise mechanism of this reaction.

The DNA Replication Fork Is Asymmetrical

During DNA replication inside a cell, each of the two original DNA strands serves as a template for the formation of an entire new strand. Because each of the two daughters of a dividing cell inherits a new DNA double helix containing one original and one new strand (**Figure 5–5**), the DNA double helix is said to be replicated "semiconservatively" by DNA polymerase. How is this feat accomplished?

Analyses carried out in the early 1960s on whole replicating chromosomes revealed a localized region of replication that moves progressively along the parental DNA double helix. Because of its Y-shaped structure, this active region is called a **replication fork** (**Figure 5–6**). At the replication fork, a multienzyme complex that contains the DNA polymerase synthesizes the DNA of both new daughter strands.

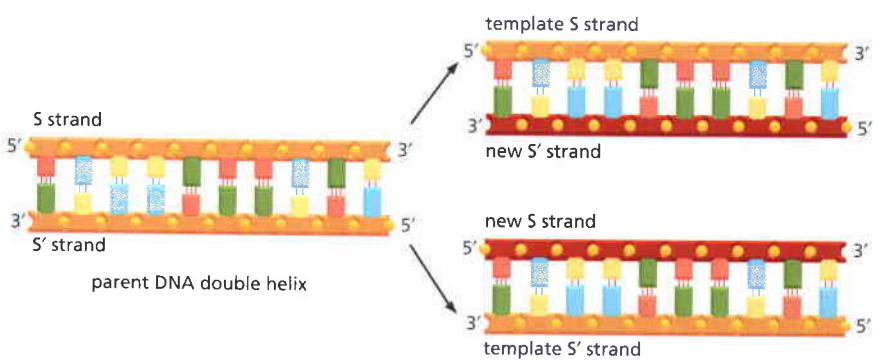

Figure 5–2 The DNA double helix acts as a template for its own duplication. Because the nucleotide A will pair successfully only with T, and G only with C, each strand of DNA can serve as a template to specify the sequence of nucleotides in its complementary strand by DNA base-pairing. In this way, a double-helical DNA molecule can be copied precisely.

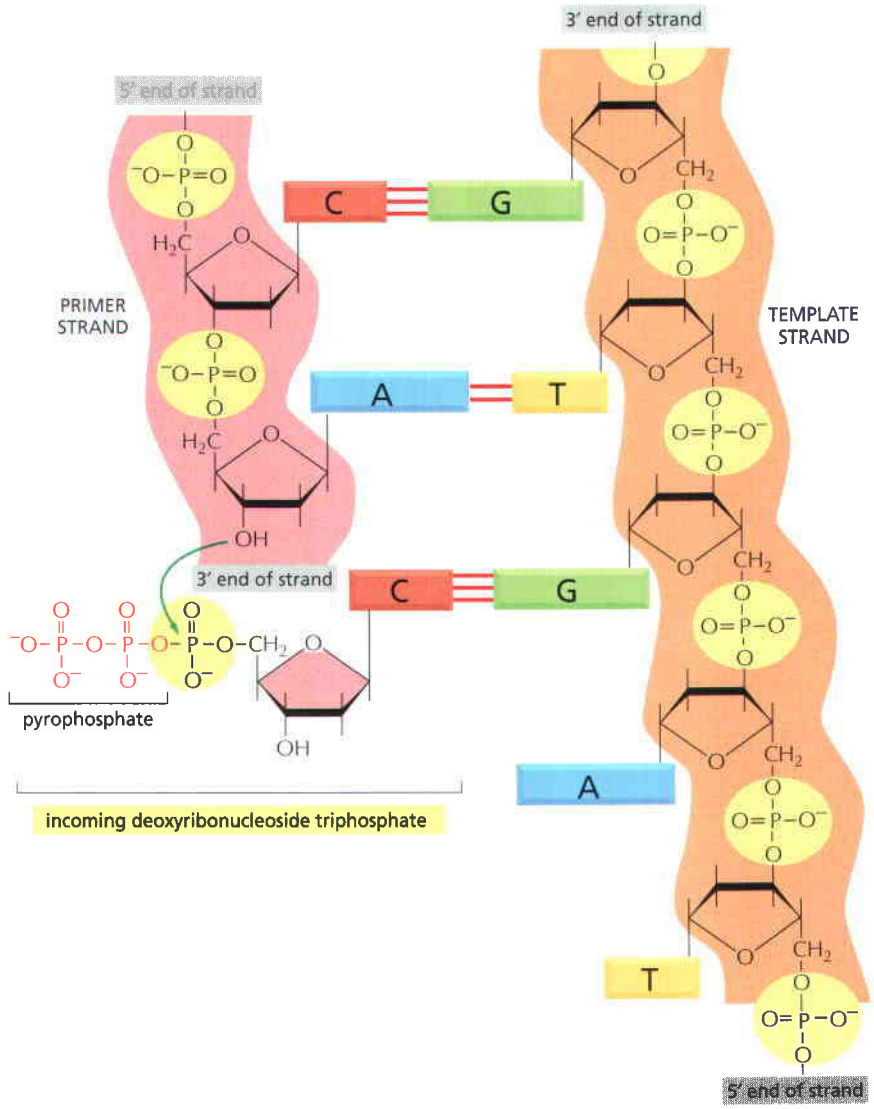

Figure 5–3 The chemistry of DNA synthesis. The addition of a deoxyribonucleotide to the 3′ end of a polynucleotide chain (the *primer strand*) is the fundamental reaction by which DNA is synthesized. As shown, base-pairing between an incoming deoxyribonucleoside triphosphate and an existing strand of DNA (the *template strand*) guides the formation of the new strand of DNA and causes it to have a complementary nucleotide sequence.

Initially, the simplest mechanism of DNA replication seemed to be the continuous growth of both new strands, nucleotide by nucleotide, at the replication fork as it moves from one end of a DNA molecule to the other. But because of the antiparallel orientation of the two DNA strands in the DNA double helix (see Figure 5–2), this mechanism would require one daughter strand to polymerize in the 5′-to-3′ direction and the other in the 3′-to-5′ direction. Such a replication fork would require two distinct types of DNA polymerase enzymes. However, all of the many DNA polymerases that have been discovered can synthesize only in the 5′-to-3′ direction.

How, then, can a DNA strand grow in the 3′-to-5′ direction? The answer was first suggested by the results of an experiment performed in the late 1960s. Researchers added highly radioactive ^{3}H-thymidine to dividing bacteria for a few seconds, so that only the most recently replicated DNA—that just behind the replication fork—became radiolabeled. This experiment revealed the transient existence of pieces of DNA that were 1000–2000 nucleotides long, now

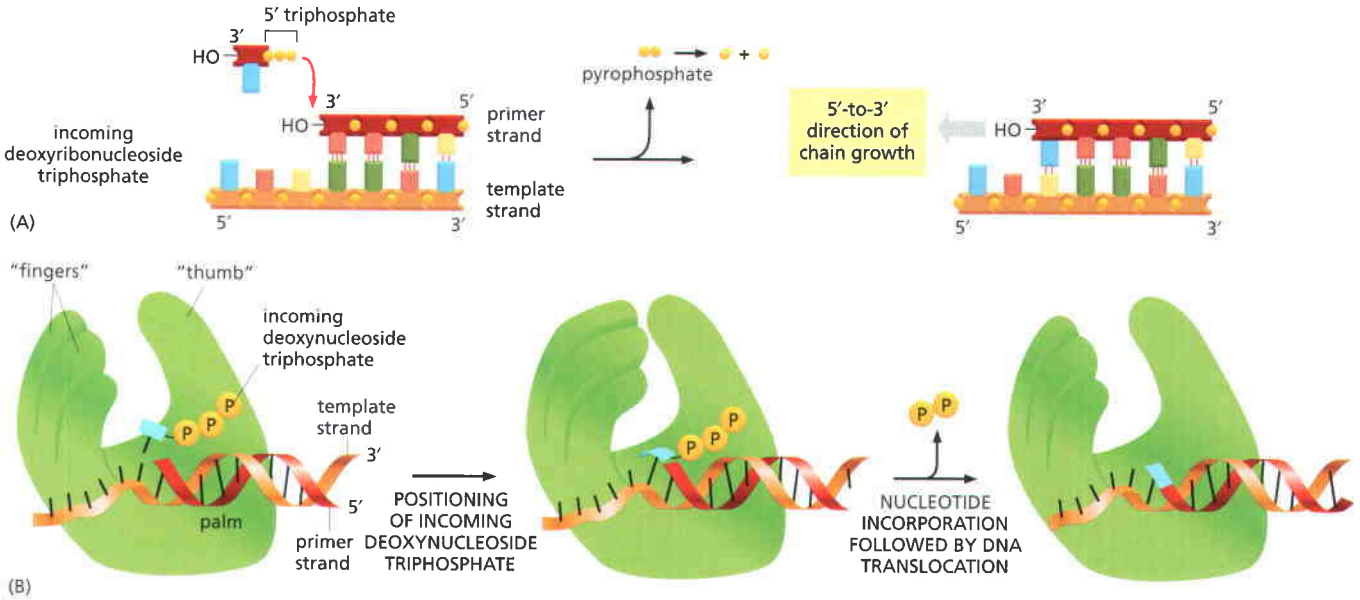

Figure 5–4 DNA synthesis catalyzed by DNA polymerase. (A) As indicated, DNA polymerase catalyzes the stepwise addition of a deoxyribonucleotide to the 3′-OH end of a polynucleotide chain, the *primer strand* that is paired to a second *template strand*. The newly synthesized DNA strand therefore polymerizes in the 5′-to-3′ direction as shown in the previous figure. Because each incoming deoxyribonucleoside triphosphate must pair with the template strand to be recognized by the DNA polymerase, this strand determines which of the four possible deoxyribonucleotides (A, C, G, or T) will be added. The reaction is driven by a large, favorable free-energy change, caused by the release of pyrophosphate and its subsequent hydrolysis to two molecules of inorganic phosphate. (B) The shape of a DNA polymerase molecule, as determined by x-ray crystallography. Roughly speaking DNA polymerases resemble a right hand in which the palm, fingers, and thumb grasp the DNA and form the active site. In the sequence shown, the correct positioning of an incoming deoxynucleoside triphosphate causes the fingers of the polymerase to tighten, thereby initiating the nucleotide addition reaction. Dissociation of pyrophosphate causes release of the fingers and translocation of the DNA by one nucleotide so the active site of the polymerase is ready to receive the next deoxynucleoside triphosphate.

commonly known as *Okazaki fragments*, at the growing replication fork. (Similar replication intermediates were later found in eucaryotes, where they are only 100–200 nucleotides long.) The Okazaki fragments were shown to be polymerized only in the 5′-to-3′ chain direction and to be joined together after their synthesis to create long DNA chains.

A replication fork therefore has an asymmetric structure (**Figure 5–7**). The DNA daughter strand that is synthesized continuously is known as the **leading strand**. Its synthesis slightly precedes the synthesis of the daughter strand that is synthesized discontinuously, known as the **lagging strand**. For the lagging strand, the direction of nucleotide polymerization is opposite to the overall direction of DNA chain growth. The synthesis of this strand by a discontinuous "backstitching" mechanism means that DNA replication requires only the 5′-to-3′ type of DNA polymerase.

The High Fidelity of DNA Replication Requires Several Proofreading Mechanisms

As discussed above, the fidelity of copying DNA during replication is such that only about 1 mistake occurs for every 10^9 nucleotides copied. This fidelity is much higher than one would expect from the accuracy of complementary base-

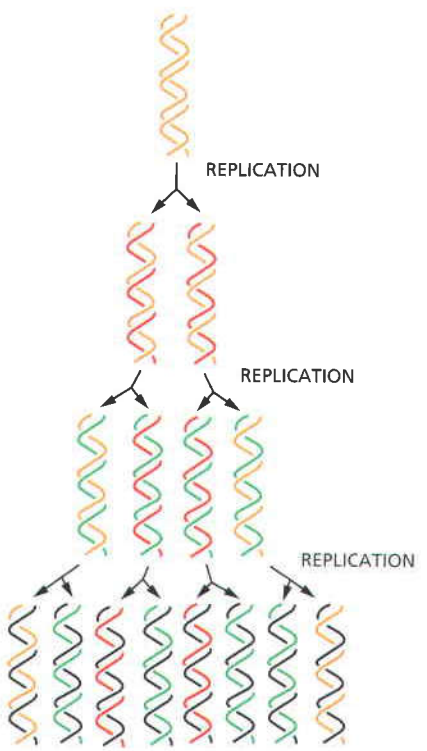

Figure 5–5 The semiconservative nature of DNA replication. In a round of replication, each of the two strands of DNA is used as a template for the formation of a complementary DNA strand. The original strands therefore remain intact through many cell generations.

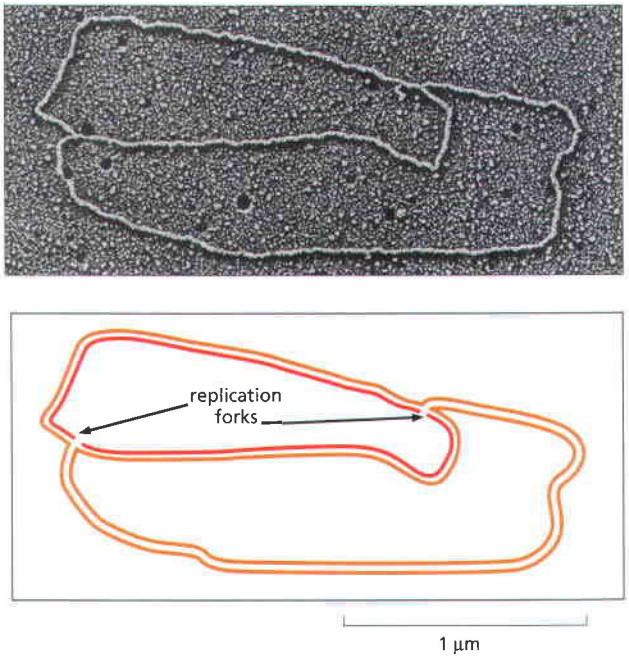

Figure 5–6 Two replication forks moving in opposite directions on a circular chromosome. An active zone of DNA replication moves progressively along a replicating DNA molecule, creating a Y-shaped DNA structure known as a replication fork: the two arms of each Y are the two daughter DNA molecules, and the stem of the Y is the parental DNA helix. In this diagram, parental strands are *orange;* newly synthesized strands are *red*. (Micrograph courtesy of Jerome Vinograd.)

pairing. The standard complementary base pairs (see Figure 4–4) are not the only ones possible. For example, with small changes in helix geometry, two hydrogen bonds can form between G and T in DNA. In addition, rare tautomeric forms of the four DNA bases occur transiently in ratios of 1 part to 10^4 or 10^5. These forms mispair without a change in helix geometry: the rare tautomeric form of C pairs with A instead of G, for example.

If the DNA polymerase did nothing special when a mispairing occurred between an incoming deoxyribonucleoside triphosphate and the DNA template, the wrong nucleotide would often be incorporated into the new DNA chain, producing frequent mutations. The high fidelity of DNA replication, however, depends not only on the initial base-pairing but also on several "proofreading" mechanisms that act sequentially to correct any initial mispairings that might have occurred.

DNA polymerase performs the first proofreading step just before a new nucleotide is added to the growing chain. Our knowledge of this mechanism comes from studies of several different DNA polymerases, including one produced by a bacterial virus, T7, that replicates inside *E. coli*. The correct nucleotide has a higher affinity for the moving polymerase than does the incorrect nucleotide, because the correct pairing is more energetically favorable. Moreover, after nucleotide binding, but before the nucleotide is covalently added to the growing chain, the enzyme must undergo a conformational change in which its "fingers" tighten around the active site (see Figure 5–4). Because this change occurs more readily with correct than incorrect base-pairing, it allows the polymerase to "double-check" the exact base-pair geometry before it catalyzes the addition of the nucleotide.

The next error-correcting reaction, known as *exonucleolytic proofreading*, takes place immediately after those rare instances in which an incorrect nucleotide is covalently added to the growing chain. DNA polymerase enzymes

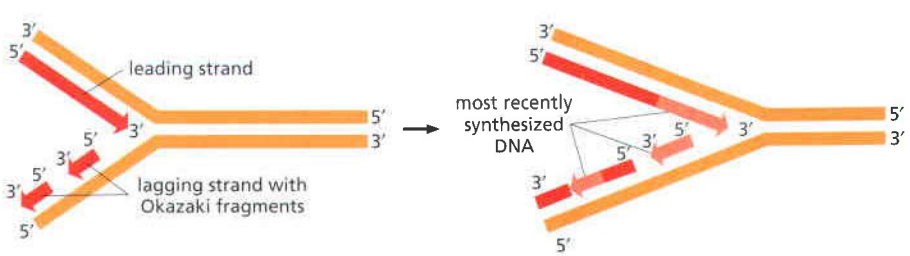

Figure 5–7 The structure of a DNA replication fork. Because both daughter DNA strands are polymerized in the 5′-to-3′ direction, the DNA synthesized on the lagging strand must be made initially as a series of short DNA molecules, called *Okazaki fragments*. On the lagging strand, the Okazaki fragments are synthesized sequentially, with those nearest the fork being the most recently made.

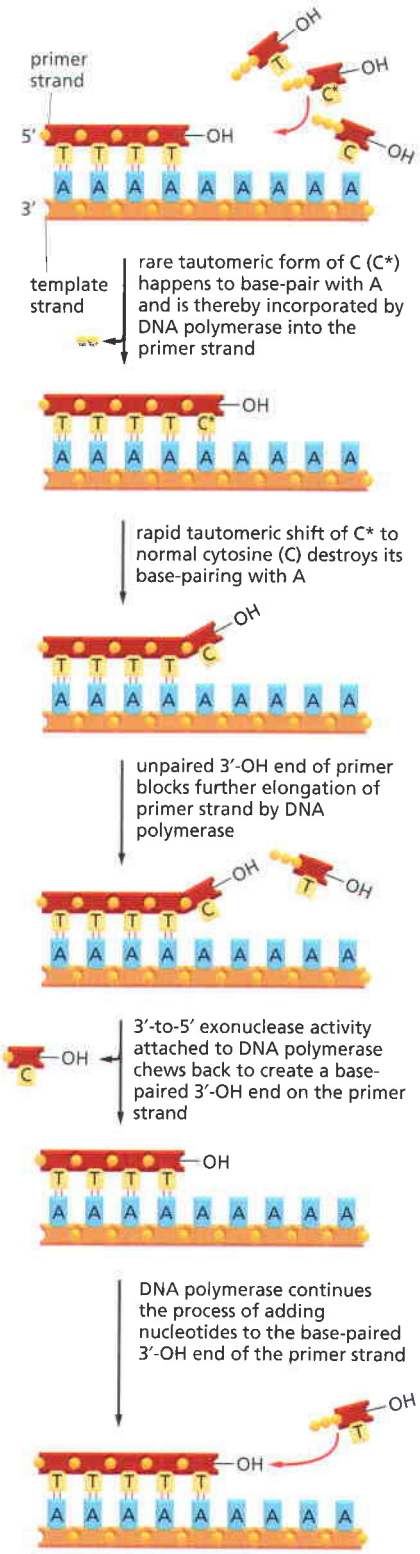

Figure 5–8 Exonucleolytic proofreading by DNA polymerase during DNA replication. In this example, the mismatch is due to the incorporation of a rare, transient tautomeric form of C, indicated by an asterisk. But the same proofreading mechanism applies to any misincorporation at the growing 3′-OH end. The part of DNA polymerase that removes the misincorporated nucleotide is a specialized member of a large class of enzymes, known as *exonucleases,* that cleave nucleotides one at a time from the ends of polynucleotides.

are highly discriminating in the types of DNA chains they will elongate: they absolutely require a previously formed base-paired 3′-OH end of a *primer strand* (see Figure 5–4). Those DNA molecules with a mismatched (improperly base-paired) nucleotide at the 3′-OH end of the primer strand are not effective as templates because the polymerase cannot extend such a strand. DNA polymerase molecules correct such a mismatched primer strand by means of a separate catalytic site (either in a separate subunit or in a separate domain of the polymerase molecule, depending on the polymerase). This *3′-to-5′ proofreading exonuclease* clips off any unpaired residues at the primer terminus, continuing until enough nucleotides have been removed to regenerate a correctly base-paired 3′-OH terminus that can prime DNA synthesis. In this way, DNA polymerase functions as a "self-correcting" enzyme that removes its own polymerization errors as it moves along the DNA (**Figure 5–8** and **Figure 5–9**).

The self-correcting properties of the DNA polymerase depend on its requirement for a perfectly base-paired primer terminus, and it is apparently not possible for such an enzyme to start synthesis *de novo.* By contrast, the RNA polymerase enzymes involved in gene transcription do not need such an efficient exonucleolytic proofreading mechanism: errors in making RNA are not passed on to the next generation, and the occasional defective RNA molecule that is produced has no long-term significance. RNA polymerases are thus able to start new polynucleotide chains without a primer.

There is an error frequency of about 1 mistake for every 10^4 polymerization events both in RNA synthesis and in the separate process of translating mRNA sequences into protein sequences. This error rate is 100,000 times greater than that in DNA replication, where a series of proofreading processes makes the process unusually accurate (**Table 5–1**).

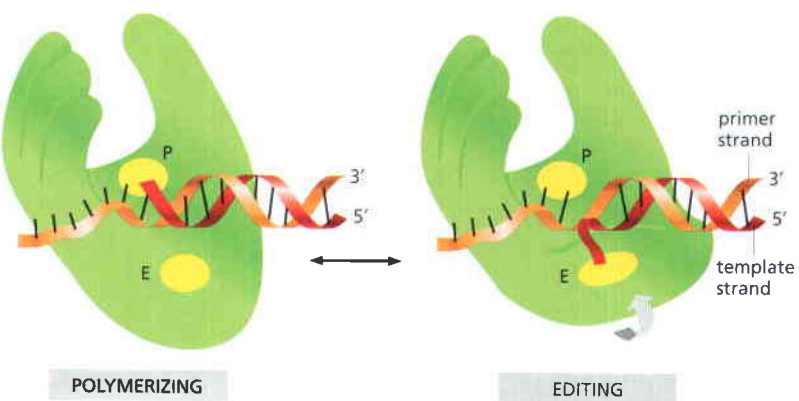

Figure 5–9 Editing by DNA polymerase. Outline of the structures of DNA polymerase complexed with the DNA template in the polymerizing mode *(left)* and the editing mode *(right).* The catalytic site for the exonucleolytic (E) and the polymerization (P) reactions are indicated. In the editing mode, the newly synthesized DNA transiently unpairs from the template and the polymerase undergoes a conformational change, moving the editing catalytic site into place to remove the most recently added nucleotide. <GATT>