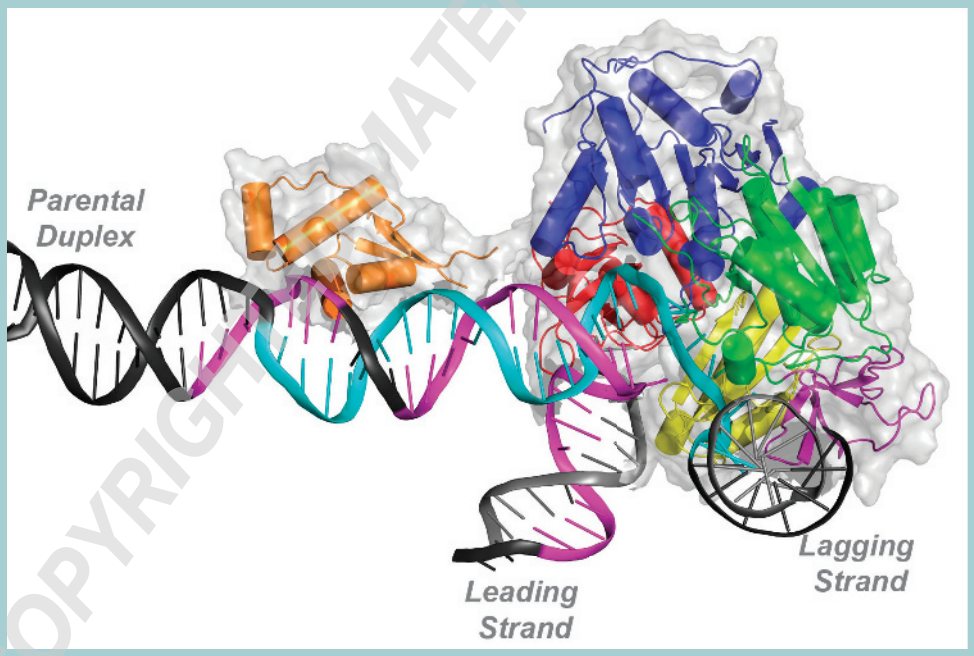


Model of the action of the PriA protein restarting a collapsed DNA replication fork. The DNA strands yet to be replicated (parental duplex) and the leading and lagging strands that have been replicated are labeled. The DNA strands shown in cyan and purple indicate regions that are believed to be bound by the PriA proteins based on biochemical experiments. The various subdomains of the PriA proteins are indicated in other colors. From Windgassen et al. (see Suggested Reading).

The Bacterial Chromosome: DNA Structure, Replication, and Segregation

DNA Structure

THE SCIENCE OF MOLECULAR GENETICS began with the determination of the structure of DNA. Experiments with bacteria and phages (i.e., viruses that infect bacteria) in the late 1940s and early 1950s, as well as the presence of DNA in chromosomes of higher organisms, had implicated this macromolecule as the hereditary material (see the introduction). In the 1930s, biochemical studies of the base composition of DNA by Erwin Chargaff established that the amount of guanine always equals the amount of cytosine and that the amount of adenine always equals the amount of thymine, independent of the total base composition of the DNA. In the early 1950s, X-ray diffraction studies by Rosalind Franklin and Maurice Wilkins showed that DNA is a double helix. Finally, in 1953, Francis Crick and James Watson put together the chemical and X-ray diffraction information in their famous model of the structure of DNA. This story is one of the most dramatic in the history of science and has been the subject of many historical treatments, some of which are listed at the end of this chapter.

Figure 1.1 illustrates the Watson-Crick structure of DNA, in which two strands wrap around each other to form a double helix. These strands can be extremely long, even in a simple bacterium, extending up to 1 mm—a thousand times longer than the bacterium itself. In a human cell, the strands that make up a single chromosome (which is one DNA molecule) are hundreds of millimeters, or many inches, long.

The Deoxyribonucleotides

If we think of DNA strands as chains, deoxyribonucleotides form the links. Figure 1.2 shows the basic structure of deoxyribonucleotides, called **deoxynucleotides** for short. Each is composed of a **base**, a **sugar**, and a **phosphate** group. The DNA bases are **adenine** (A), **cytosine** (C), **guanine** (G), and **thymine** (T), which have either one or two rings, as shown in Figure 1.2. The bases with two rings (A and G) are the **purines**, and those with only one ring (T and C) are **pyrimidines**. A third pyrimidine, uracil (U), replaces thymine in RNA. The carbons and nitrogens making up the rings of the bases are numbered sequentially, as shown in the figure. All four DNA bases are attached to the five-carbon sugar deoxyribose. This sugar is identical to ribose, which is found in RNA, except that it does not have an oxygen attached to the second

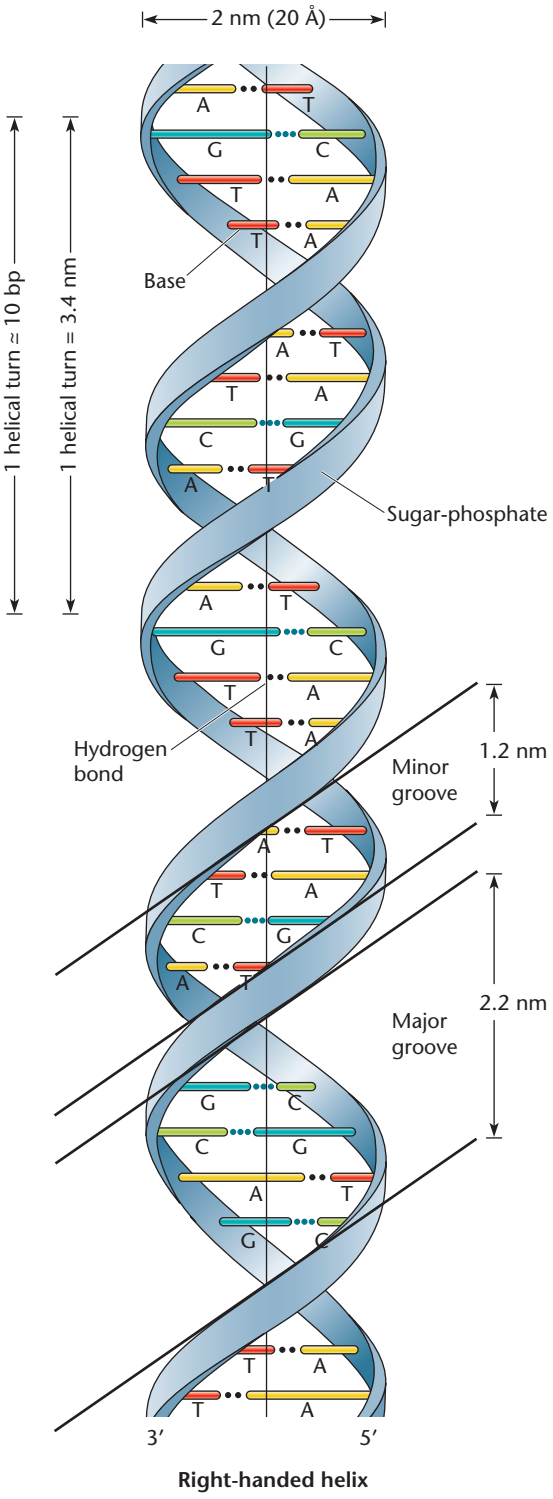


Figure 1.1 Schematic drawing of the Watson-Crick structure of DNA, showing the helical sugar-phosphate backbones of the two strands held together by hydrogen bonding between the bases. Also shown are the major and minor grooves and the dimensions of the helix.

carbon—hence the name deoxyribose. The carbons in the sugar of a nucleotide are also numbered 1, 2, 3, and so on, but they are labeled with "primes" to distinguish them from the carbons in the bases (Figure 1.2). The nucleotides also have one or more phosphate groups attached to a carbon of the deoxyribose sugar, as shown. The carbon to which the phosphate group is attached is indicated, although if the group is attached to the 5′ carbon (the usual situation), the carbon to which it is attached is often not stipulated.

The components of the deoxynucleotides have special names. A **deoxynucleoside** (rather than -tide) is a base attached to a sugar but lacking a phosphate. Without phosphates, the four deoxynucleosides are called **deoxyadenosine**, **deoxycytidine**, **deoxyguanosine**, and **deoxythymidine**. As shown in Figure 1.2, the deoxynucleotides have one, two, or three phosphates attached to the sugar and are known as deoxynucleoside monophosphates, diphosphates, or triphosphates, respectively. The individual deoxynucleoside monophosphates, called deoxyguanosine monophosphate etc., are often abbreviated dGMP, dAMP, dCMP, and dTMP, where the d stands for deoxy; the G, A, C, or T stands for the base; and the MP stands for monophosphate. In turn, the diphosphates are abbreviated dGDP, dADP, dCDP, and dTDP, and the triphosphates are abbreviated dGTP, dATP, dCTP, and dTTP. The phosphate attached to the sugar is called the α phosphate, while the next two are called the β and γ phosphates, respectively, as shown in the figure. Collectively, the four deoxynucleoside triphosphates are often referred to as dNTPs.

The DNA Chain

Phosphodiester bonds join each deoxynucleotide link in the DNA chain. As shown in Figure 1.3, the phosphate attached to the last (5′) carbon of the deoxyribose sugar of one nucleotide is attached to the third (3′) carbon of the sugar of the next nucleotide, thus forming one strand of nucleotides connected 5′ to 3′, 5′ to 3′, etc.

The 5′ and 3′ Ends

The nucleotides found at the ends of a linear piece of DNA have properties that are biochemically important and useful for orienting the DNA strand. At one end of the DNA chain, a nucleotide will have a phosphate attached to its 5′ carbon that does not connect it to another nucleotide. This end of the strand is called the 5′ end or the **5′ phosphate end** (Figure 1.3B). On the other end, the last nucleotide lacks a phosphate at its 3′ carbon. Because it has only a hydroxyl group (the OH in Figure 1.3B), this end is called the 3′ end or the **3′ hydroxyl end**.

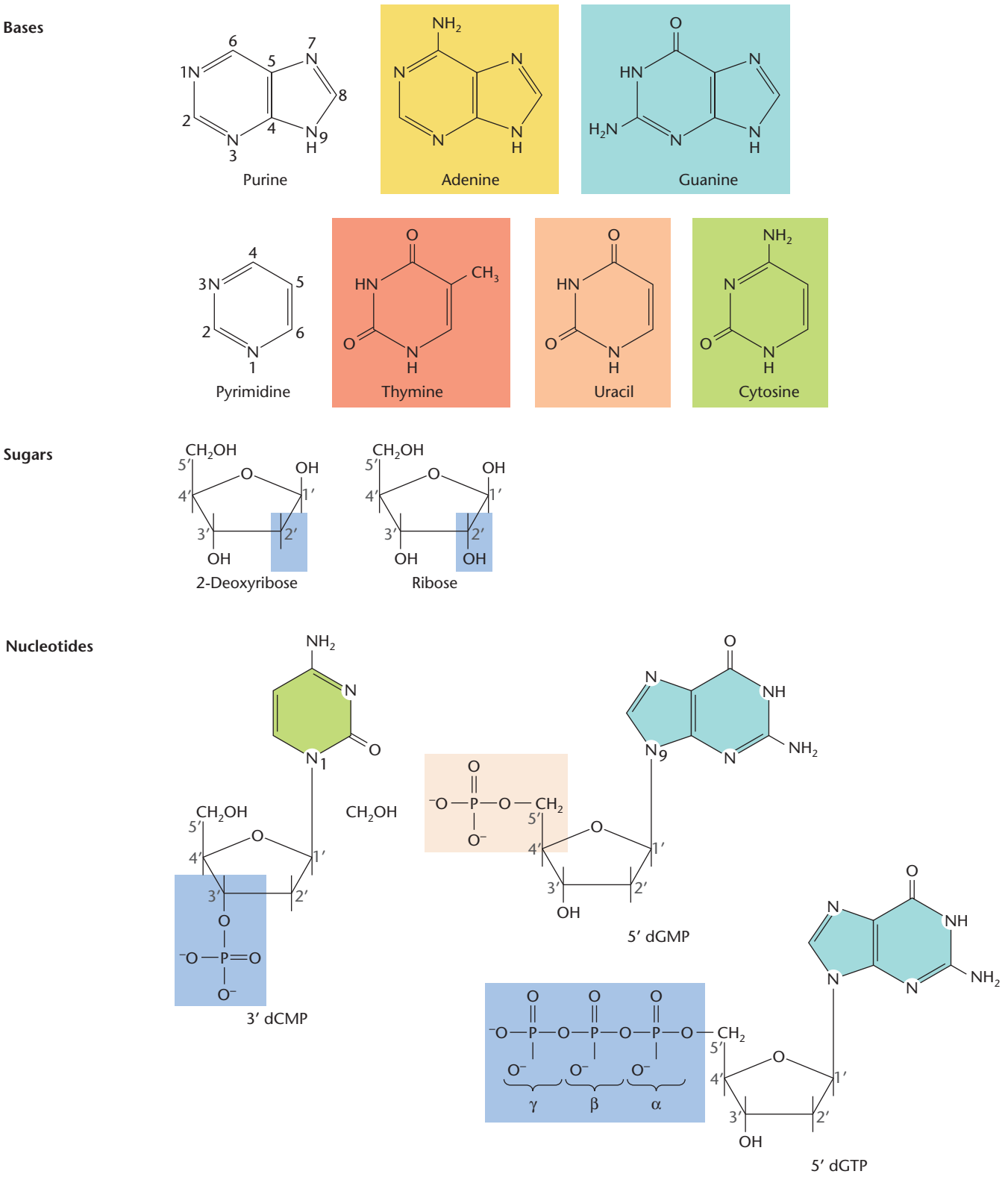


Figure 1.2 Chemical structures of deoxyribonucleotides, showing the bases and sugars and how they are assembled into a deoxyribonucleotide.

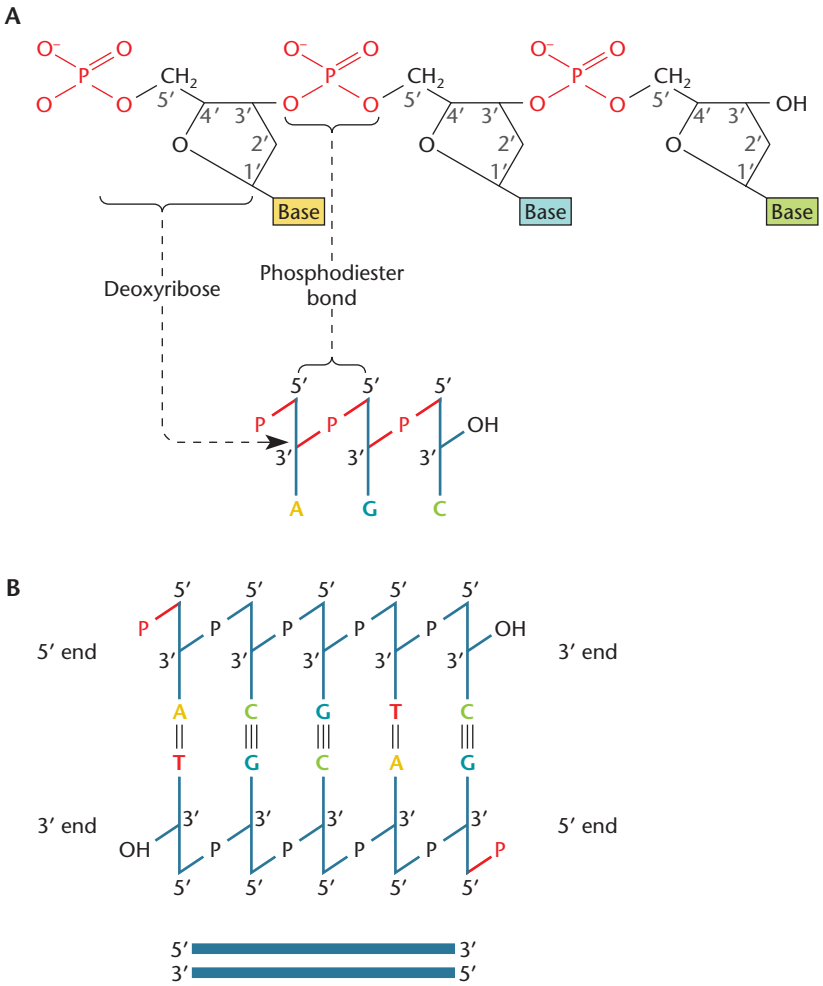


Figure 1.3 **(A)** Schematic drawing of a DNA chain, showing the 3′-to-5′ attachment of the phosphates to the sugars, forming phosphodiester bonds. **(B)** Two strands of DNA bind at the bases in an antiparallel arrangement of the phosphate-sugar backbones.

Base Pairing

The sugar and phosphate groups of DNA form what is often called a **backbone** to support the bases, which jut out from the chain. This structure allows the bases from one single strand of DNA to form hydrogen bonds with another strand of DNA, thereby holding together two separate nucleotide chains (Figure 1.3B). The first clue that pairing between specific bases could form the basis for the structure of DNA came from Erwin Chargaff's observation about the ratios of the bases; no matter the source of the DNA, the concentration of guanine (G) always equals the concentration of cytosine (C) and the concentration of adenine (A) always equals the concentration of thymine (T). These ratios, named Chargaff's rules, gave Watson and Crick one of the essential clues to the structure of DNA. They proposed that the two strands of the DNA are held together by specific hydrogen bonding between the bases in opposite strands, as shown in Figure 1.4. Thus, the amounts of A and T and of C and G are always the same because A's pair only with T's and G's pair only with C's to hold the DNA strands together. Each A-and-T pair and each G-and-C pair in DNA is called a **complementary base pair**, and the sequences of two strands of DNA are said to be complementary if one strand always has a T where there is an A in the other strand and a G where there is a C in the other strand.

It did not escape the attention of Watson and Crick that the complementary base-pairing rules essentially explain heredity. If A pairs only with T and G pairs only with C, then each strand of DNA can replicate to make a complementary copy, so that the two replicated DNAs will be exact copies of each other. Offspring containing the new DNAs would have the same sequence of nucleotides in their DNAs as their parents and thus would be exact copies of their parents.

Antiparallel Construction

As mentioned at the beginning of this section, the complete DNA molecule consists of two long chains wrapped around each other in a double helix (Figure 1.1). The

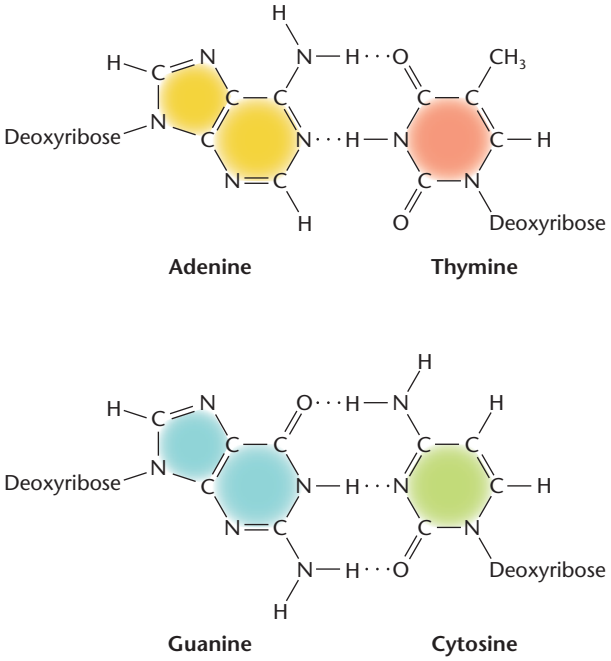


Figure 1.4 The two complementary base pairs found in DNA. Two hydrogen bonds form in adenine-thymine base pairs. Three hydrogen bonds form in guanine-cytosine base pairs.

double-stranded molecule can be thought of as being like a circular staircase, with the alternating phosphates and deoxyribose sugars forming the railings and the bases connected to each other forming the steps. However, the two chains run in opposite orientations, with the phosphates on one strand attached 5′ to 3′, 5′ to 3′, etc., to the sugars and those on the other strand attached 3′ to 5′, 3′ to 5′, etc. This arrangement is called **antiparallel**. In addition to phosphodiester bonds running in opposite directions, the antiparallel construction causes the 5′ phosphate end of one strand and the 3′ hydroxyl end of the other to be on the same end of the double-stranded DNA molecule (Figure 1.3B).

The Major and Minor Grooves

Because the two strands of DNA are wrapped around each other to form a double helix, the helix has two grooves between the two strands (Figure 1.1). One of these grooves is wider than the other, so it is called the **major groove**. The other, narrower groove is called the **minor groove**. Most of the modifications to DNA that are discussed in this and later chapters occur in the major groove of the helix.

The Mechanism of DNA Replication

The molecular details of DNA replication are probably similar in all organisms on Earth. The basic process of replication involves **polymerizing**, or linking, the nucleotides of DNA into long chains, or strands, using the sequence on the other strand as a guide. Because the nucleotides must be made before they can be put together into DNA, the nucleotides are an essential **precursor** of DNA synthesis.

Deoxyribonucleotide Precursor Synthesis

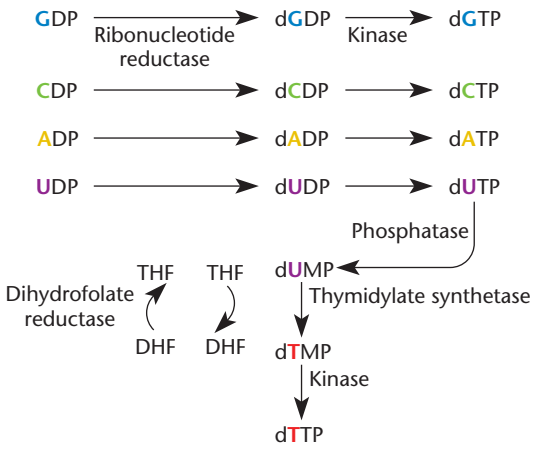
The precursors of DNA synthesis are the four deoxyribonucleoside triphosphates, dATP, dGTP, dCTP, and dTTP. The triphosphates are synthesized from the corresponding ribose nucleoside diphosphates by the pathway shown in Figure 1.5. In the first step, the enzyme **ribonucleotide reductase** reduces (i.e., removes an oxygen from) the ribose sugar to produce the deoxyribose sugar by changing the hydroxyl group at the 2′ position (the second carbon) of the sugar to a hydrogen. Then, an enzyme known as a **kinase** adds a phosphate to the deoxynucleoside diphosphate to make the deoxynucleoside triphosphate precursor.

The deoxynucleoside triphosphate dTTP is synthesized by a somewhat different pathway from the other three. The first step is the same. Ribonucleotide reductase synthesizes the nucleotide dUDP (deoxyuridine diphosphate) from the ribose UDP. However, from then on, the pathway differs. A phosphate is added to make dUTP, and the dUTP is converted to dUMP by a phosphatase that removes two of the phosphates. This molecule is then converted to dTMP by the enzyme thymidylate synthetase, using tetrahydrofolate to donate a methyl group. Kinases then add two phosphates to the dTMP to make the precursor dTTP.

Replication of the Bacterial Chromosome

Once the precursors of DNA replication are synthesized, they must be polymerized into long double-stranded DNA molecules. A very large complex of many enzymes

Figure 1.5 The pathways for synthesis of deoxynucleotides from ribonucleotides. Some of the enzymes referred to in the text are identified. THF, tetrahydrofolate; DHF, dihydrofolate.



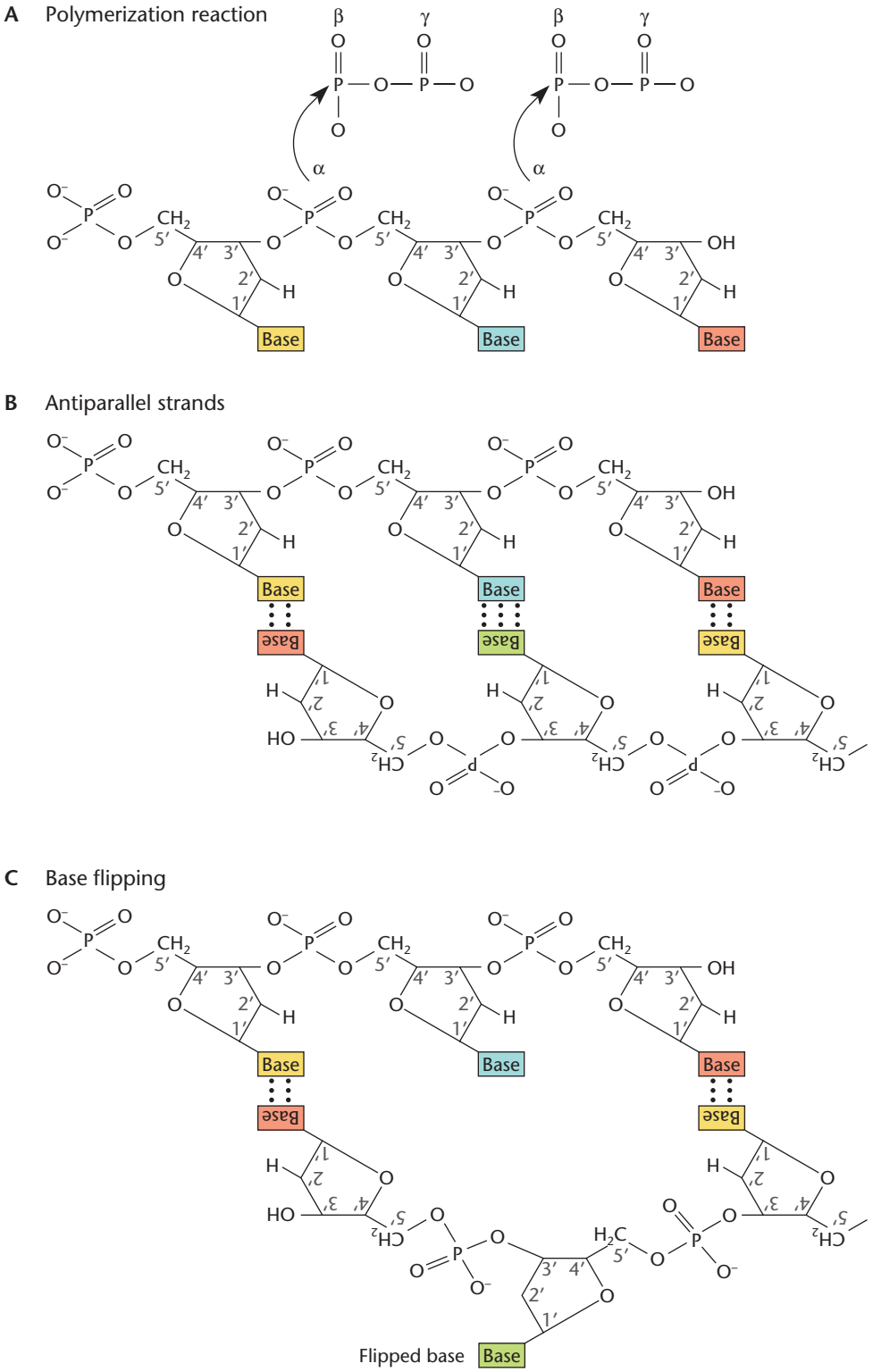


Figure 1.6 Features of DNA. **(A)** Polymerization of the deoxynucleotides during DNA synthesis. The β and γ phosphates of each deoxynucleoside triphosphate are cleaved off to provide energy for the polymerization reaction. **(B)** The strands of DNA are antiparallel. **(C)** A single base can be flipped out from the double helix, which could be important in recombination and repair.

assembles on the DNA and moves along the DNA, separating the strands and making a complementary copy of each of the strands. Thus, two strands of DNA enter the complex and four strands emerge on the other side, forming a branched structure. Each of the emerging branches contains one old "conserved" strand and one new strand, hence the name **semiconservation replication**. This branched structure where replication is occurring is called the **replication fork**. In this section, we discuss what is happening in the replication fork, including the mechanisms used to overcome obstacles to the progression of the replication fork on the chromosome.

DNA POLYMERASES

The properties of the DNA polymerases, the enzymes that join the deoxynucleotides together to make the long chains, are the best guides to an understanding of the replication of DNA. These enzymes make DNA by linking one deoxynucleotide to another to generate a long chain of DNA. This process is called DNA **polymerization,** hence the name DNA polymerases.

Figure 1.6 shows the basic process of DNA polymerization by DNA polymerase. The DNA polymerase attaches the first phosphate (the α phosphate) of one deoxynucleoside triphosphate to the 3′ carbon of the sugar of the next deoxynucleoside triphosphate, in the process releasing the last two phosphates (the β and γ phosphates) of the first deoxynucleoside triphosphate to produce energy for the reaction. Then the α phosphate of another deoxynucleoside triphosphate is attached to the 3′ carbon of this deoxynucleotide, and the process continues until a long chain is synthesized.

DNA polymerases also need a **template strand** to direct the synthesis of the new strand (Figure 1.7). As mentioned in "Base Pairing" above, complementary base pairing dictates that wherever there is a T in the template strand, an A is inserted in the strand being synthesized and so forth according to the base-pairing rules. The DNA polymerase can move only in the 3′-to-5′ direction on the template strand, linking deoxynucleotides in the new strand in the 5′-to-3′ direction. When replication is completed, the product is a new double-stranded DNA with antiparallel strands, one of which is the old template strand and one of which is the newly synthesized strand.

There are two DNA polymerases that participate in normal DNA replication in *Escherichia coli*; they are called DNA polymerase III and DNA polymerase I (Table 1.1). DNA polymerase III is a large protein complex in which the enzyme that polymerizes nucleotides works with numerous accessory proteins. In *E. coli* and other bacteria, DNA polymerase III is responsible for the bulk of DNA replication on both DNA strands. As discussed below, DNA polymerase I has a number of features that are important because replication is continually reinitiated on one of the DNA strands. It also plays a role in DNA repair, as discussed in chapter 10. Table 1.1 lists many of the DNA replication proteins, the genes encoding them, and their functions.

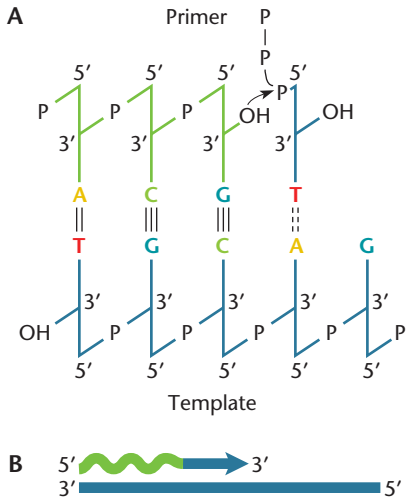


Figure 1.7 Functions of the primer and template in DNA replication. **(A)** The DNA polymerase adds deoxynucleotides to the 3′ end of the primer by using the template strand to direct the selection of each base. **(B)** Simple illustration of 5′-to-3′ DNA synthesis. The wavy green line represents the primer.

PRIMASES

One type of enzyme, called a **primase**, is required during DNA replication because DNA polymerases cannot start the synthesis of a new strand of DNA; they can only attach deoxynucleotides to a preexisting 3′ OH group. The 3′ OH group to which DNA polymerase adds a deoxynucleotide is called the **primer** (Figure 1.7). The requirement for a primer for DNA polymerase creates an apparent dilemma in DNA replication. When a new strand of DNA is synthesized, there is no DNA upstream (i.e., on the 5′ side) to act as a primer. Primase makes small stretches of RNA that are complementary to the template strand, which are in turn used to initiate, or prime, polymerization of a new strand of DNA (Figure 1.8). In some special situations, an RNA primer for DNA replication can be made by the RNA polymerase used for information processing (i.e., the RNA polymerase that makes all the other RNAs, including mRNA, tRNA, and rRNA; see chapter 2). Unlike DNA polymerase, primase and RNA polymerases do not require a primer to initiate the synthesis of new strands. During DNA replication, a special enzyme activity recognizes and removes the RNA primer (see below).

Table 1.1 Proteins involved in *Escherichia coli* DNA replication

Protein	Gene	Function
DnaA	*dnaA*	Initiator protein; primosome (priming complex) formation
DnaB	*dnaB*	DNA helicase
DnaC	*dnaC*	Delivers DnaB to replication complex
SSB	*ssb*	Binding to single-stranded DNA
Primase	*dnaG*	RNA primer synthesis
DNA ligase	*lig*	Sealing DNA nicks
DNA gyrase		Supercoiling
α	*gyrA*	Nick closing
β	*gyrB*	ATPase
DNA Pol I	*polA*	Primer removal; gap filling
DNA Pol III (holoenzyme)		
α	*dnaE*	Polymerization
ε	*dnaQ*	3′-to-5′ editing
RNase H	*rnhA*	Can aid in RNA primer removal
θ	*holE*	Present in core (αεθ)
β	*dnaN*	Sliding clamp
τ^a	*dnaX*	Organizes complex; joins leading and lagging DNA Pol III
γ^b	*dnaX*	Binds clamp loaders and single-strand-binding protein
δ	*holA*	Clamp loading
δ′	*holB*	Clamp loading
χ	*holC*	Binds single-strand-binding protein
φ	*holD*	Holds χ to the clamp loader

[a]Full-length product of the *dnaX* gene.
[b]Shorter product of the *dnaX* gene produced by translational frameshifting (see chapter 2).

NUCLEASES

Enzymes that degrade DNA strands by breaking the phosphodiester bonds are just as important in replication as the enzymes that polymerize DNA by forming phosphodiester bonds between the nucleotides. These bond-breaking enzymes, called **nucleases**, can be grouped into two major categories. One type can initiate breaks in the middle of a DNA strand and so are called **endonucleases**, from a Greek word meaning "within," and the other type can remove nucleotides only from the ends of DNA strands and so are called exonucleases, from a Greek word meaning "outside." A special type of endonuclease activity, called a **flap endonuclease** activity, is involved in primer removal by DNA polymerase I. The flap endonuclease activity appears to be common to all organisms for removing RNA primers. In *E. coli*, DNA polymerase I displaces the RNA primer, making a flap-like structure, and then the flap endonuclease activity of Pol I cleaves away the oligonucleotide as indicated (Figure 1.9). The **exonucleases** can be subdivided into two groups. Some exonucleases can degrade only from the 3′ end of a DNA strand, degrading DNA in the 3′-to-5′ direction. These are called 3′ exonucleases; one example of their activity is their role in the editing function associated with DNA polymerases I and III, which is discussed below. Other exonucleases, called 5′ exonucleases, degrade DNA strands only from the 5′ end.

DNA LIGASES

DNA **ligases** are enzymes that form phosphodiester bonds between the ends of separate presynthesized chains of DNA. This important function cannot be performed by any of the known DNA polymerases. During replication, ligase joins the 5′ phosphate at the end of one DNA chain to the 3′ hydroxyl at the end of another chain to make a longer, continuous chain (Figure 1.8).

ACCESSORY PROTEINS

Replication of large DNAs requires many functions that reside in proteins separate from the subunit used for polymerizing the chain of nucleotides. These functions include the coordination of multiple DNA polymerases and tethering of these components to the template DNA strands as a moving production platform. DNA polymerase III is the major DNA replication protein in *E. coli* responsible for polymerizing the new complementary DNA strands, and it functions with multiple **DNA polymerase accessory proteins** that travel along the template strand with the molecule of DNA polymerase III. The term **DNA polymerase III holoenzyme** can be used to describe

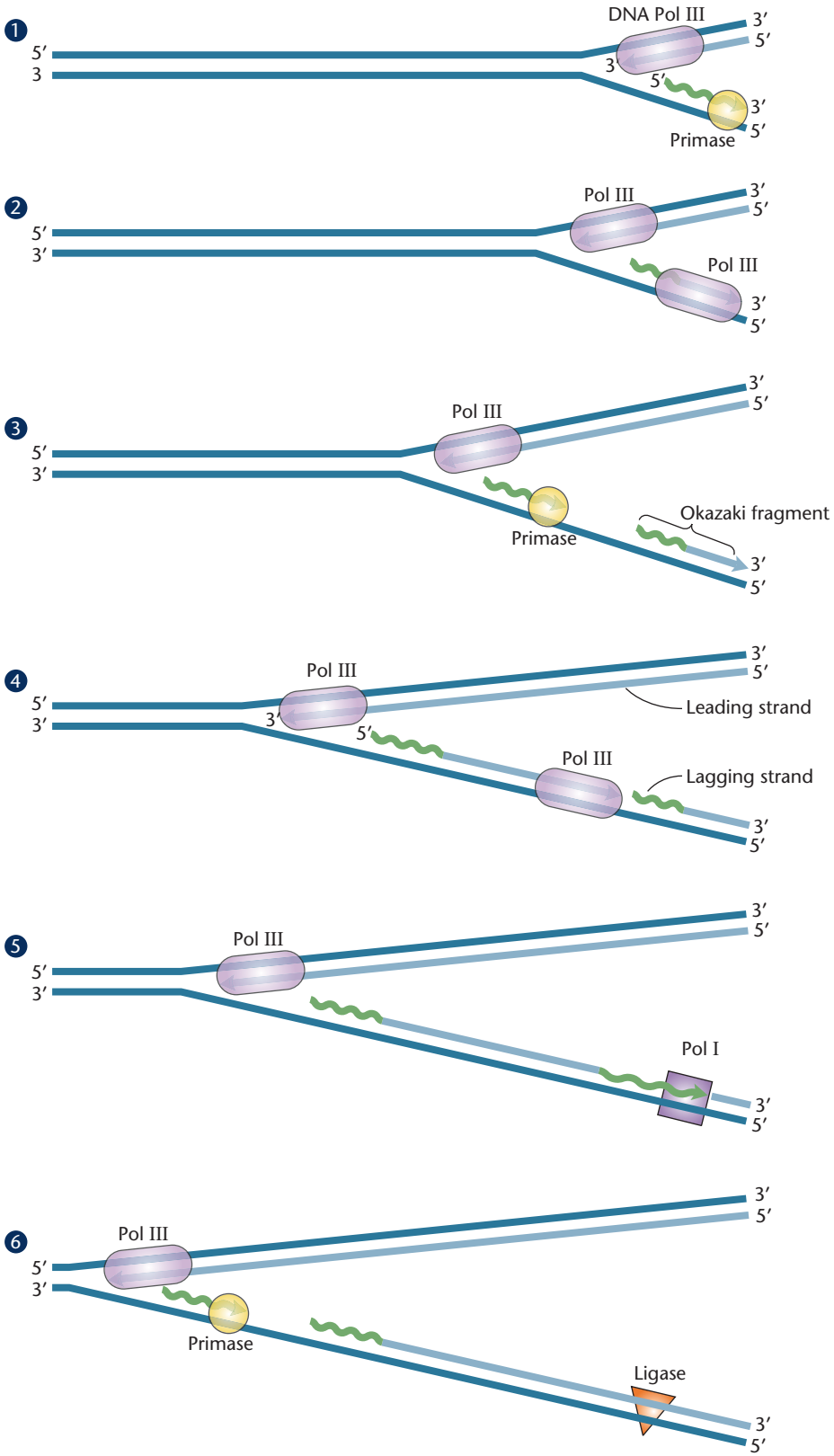


Figure 1.8 Discontinuous synthesis of one of the two strands of DNA during chromosome replication. **(1)** DNA polymerase (Pol) III replicates one strand, and the primase synthesizes RNA on the other strand in the opposite direction. **(2)** Pol III extends the RNA primer to synthesize an Okazaki fragment. **(3)** The primase synthesizes another RNA primer. **(4)** Pol III extends this primer until it reaches the previous primer. **(5)** Pol I removes the first RNA primer and replaces it with DNA. **(6)** DNA ligase seals the nick to make a continuous DNA strand, and the process continues. The strand that is synthesized continuously is the leading strand; the strand that is synthesized discontinuously is the lagging strand.

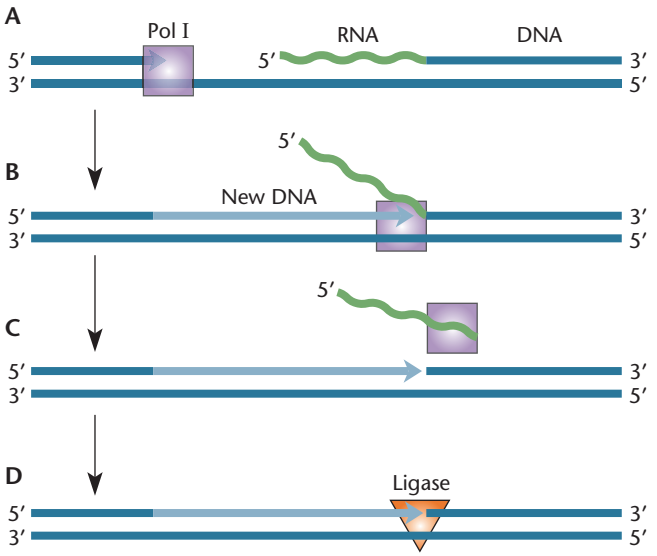


Figure 1.9 DNA polymerase I can remove an RNA primer by using strand displacement and endonuclease activity. **(A)** The DNA strand produced by DNA polymerase III holoenzyme is extended by DNA polymerase I until it encounters a previously synthesized RNA primer. **(B)** During the process of DNA replication, DNA polymerase I displaces the RNA primer. **(C)** An endonuclease activity in DNA polymerase I is used to cleave off the RNA primer. **(D)** Ligase joins the new Okazaki fragment to the previous Okazaki fragment to allow a contiguous DNA.

the entire complex of proteins. The various subunits and subassemblies of the DNA polymerase III holoenzyme were originally identified from fractionation procedures and were designated by Greek letters (Table 1.1).

One of the DNA polymerase accessory proteins forms a ring around the template DNA strand and is responsible for keeping DNA polymerase from falling off. Because this ring slides freely over double-stranded DNA and will not easily come off the DNA, it is also referred to as a sliding clamp, or β clamp. The β clamp provides the foundation of the mobile platform for DNA replication, allowing it to continue for long distances without being released. In bacteria, the β clamp is a product of the *dnaN* gene, where two head-to-toe molecules form the ring around the DNA. While it was first isolated as part of the DNA polymerase III holoenzyme, the β clamp protein is important for multiple DNA transactions.

A special subcomplex within DNA polymerase III is the **clamp loader**, which is responsible for loading β clamp proteins onto the DNA. The clamp-loading complex is also responsible for tethering proteins across the DNA replication fork; the clamp loader binds the DNA polymerases on both DNA template strands and the enzyme responsible for separating the DNA strands (see below). The clamp loader is a complicated structure that consists of one γ and two τ proteins and one each of δ and δ′, which form a five-sided structure, and two additional proteins, χ and ψ. The clamp loader complex is also responsible for removing β clamps. The rate of clamp removal allows many β clamps to reside on the DNA for a period of time after the replication fork passes (see Moolman et al., Suggested Reading). β clamps temporarily left behind on the newly replicated DNA play a role in helping to recruit other proteins responsible for various replication and repair functions described in this and other chapters.

Replication of Double-Stranded DNA

Additional complications of DNA replication come from the fact that the DNA is double stranded and the strands are antiparallel. The replication of all bacterial chromosomes begins at one point, called the **origin of replication**, with the replication enzymes moving in opposite directions from this point along the chromosome. In this process, both strands of DNA are replicated at the same time with a coordinated set of proteins. Replicating the antiparallel strands is further complicated by the above-mentioned fact that DNA polymerases can replicate only in the 5′-to-3′ direction. Therefore, one DNA strand is replicated in the same direction that the replication fork is moving, and in theory, replication of this strand could continue without the need for reinitiating in a process called **leading-strand** DNA synthesis. However, replication of the other DNA strand occurs in the opposite direction from the progression of the replication machinery. Replication of this strand must continually be reinitiated in a process known as **lagging-strand** DNA synthesis. Replication of double-stranded DNA requires coordination between multiple holoenzyme subunits and DNA polymerases, as well as a host of other replication proteins.

SEPARATING THE TWO TEMPLATE DNA STRANDS

To serve as templates for DNA replication, the two DNA strands must be separated, a task that DNA polymerase cannot perform on its own. The strands must be separated because the bases of the DNA are inside the double helix, where they are not available to pair with the incoming deoxynucleotides to direct which nucleotide will be inserted at each step. Proteins called DNA **helicases** separate the strands of DNA (see Singleton et al., Suggested Reading). Many of these proteins form a ring around one strand of DNA and propel the strand through the ring, acting as a mechanical wedge that strips the strands apart as it moves. It takes a lot of energy to separate the strands of DNA, and helicases cleave a lot of ATP for energy, forming ADP in the process. There are about 20 different helicases in *E. coli*, and each helicase works in only one direction, either the 3′-to-5′ or the 5′-to-3′ direction. The DnaB helicase that normally separates the strands of DNA ahead of the replication fork in *E. coli* is a large doughnut-shaped complex composed of six polypeptide products of the *dnaB* gene. It propels one

strand, the template for lagging-strand DNA replication, through the center of the complex in the 5′-to-3′ direction, opening strands of DNA ahead of the replication fork (Figure 1.10). The DnaB ring cannot load onto single-stranded DNA on its own to start a DNA replication fork; it requires the loading protein DnaC. Other helicases are discussed in later chapters in connection with recombination and repair.

Once the strands of DNA have been separated, they also must be prevented from coming back together (or from annealing to themselves if they happen to be complementary over short regions). Separation of the strands is maintained by proteins called single-strand-binding (SSB) proteins or, less frequently, helix-destabilizing proteins. They are proteins that bind preferentially to single-stranded DNA and prevent double-stranded helical DNA from reforming prematurely. Interestingly, SSB activity goes beyond this passive role. SSB is also responsible for recruiting a number of replication and repair proteins through a specific set of amino acids encoded in the very C-terminal end of SSB, allowing it to serve as an organizational hub for other processes.

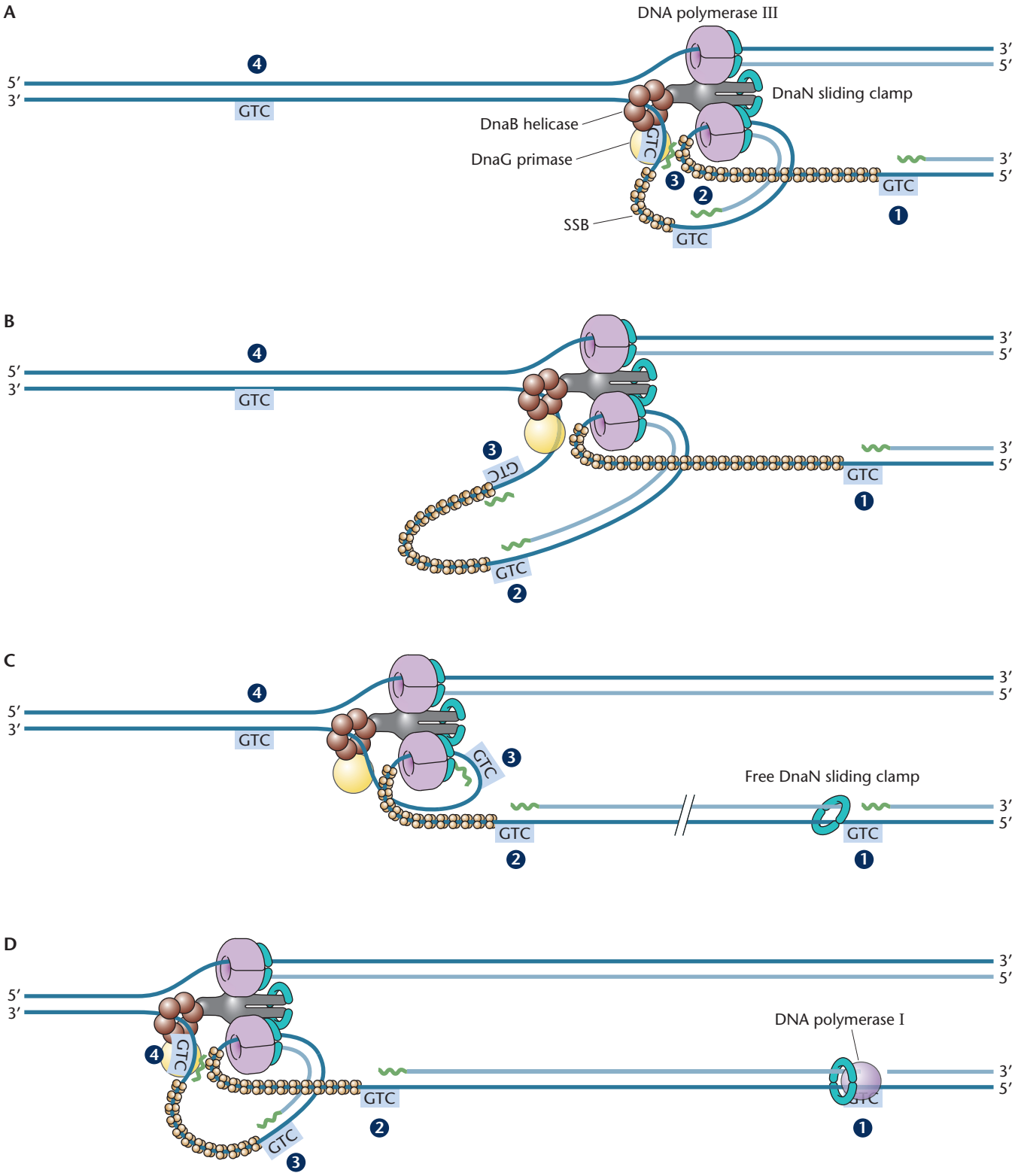
PROCESSING THE TWO TEMPLATE DNA STRANDS

As discussed above, the antiparallel configuration of DNA requires that the two DNA polymerases travel in two different directions while still allowing the larger replication machine to travel in one direction down the chromosome (Figure 1.8). This leads to fundamental differences in the natures of leading- and lagging-strand DNA replication. While replication of the leading-strand template can occur as soon as the strands are separated by the DnaB helicase, replication of the lagging-strand template is consistently reinitiated approximately every 1 to 2 kilobases (kb); this slows the process, hence the name lagging-strand synthesis. The short pieces of DNA produced from the lagging-strand template are called **Okazaki fragments**. Synthesis of each Okazaki fragment requires a new RNA primer about 10 to 12 nucleotides in length. In *E. coli*, these primers are synthesized by DnaG primase at the template sequence 3′-GTC-5′, beginning synthesis opposite the T. These RNA primers are then used to prime DNA synthesis by DNA polymerase III, which continues until it reaches the last RNA primer produced by DnaG (Figure 1.8). Before these short pieces of DNA that are annealed to the template can be joined to make a long, continuous strand of DNA, the short RNA primers must be removed. This process is carried out by DNA polymerase I using its flap exonuclease activity to displace and cleave the RNA strand (Figure 1.9). As DNA polymerase I displaces the RNA primer, it extends the upstream (i.e., 5′) DNA that was previously polymerized by DNA polymerase III (Figure 1.8). Ribonuclease (RNase) H may contribute to this process under some circumstances by using its ability to degrade the RNA strand of a DNA-RNA double helix (Table 1.1). The Okazaki fragments are then joined together by DNA ligase as the replication fork moves on, as shown in Figure 1.8. By using RNA rather than DNA to prime the synthesis of new Okazaki fragments, the cell likely lowers the mistake rate of DNA replication (see below).

What actually happens at the replication fork is more complicated than is suggested by the simple picture given so far. For one thing, this picture ignores the overall topological restraints on the replicating DNA. The **topology** of a molecule refers to its position in space. Because the circular DNA is very long and its strands are wrapped around each other, pulling the two strands apart introduces stress into other regions of the DNA in the form of **supercoiling**. If no mechanism existed to allow the two strands of DNA to rotate around each other, supercoiling would cause the chromosome to look like a telephone cord wound up on itself, an event that has been experimentally shown to eventually halt progression of the DNA replication fork. To relieve this stress, enzymes called **topoisomerases** work to help undo the supercoiling ahead of the replication fork. DNA supercoiling and topoisomerases are discussed below. The fork itself can also twist when the supercoiling that builds up ahead of the replication fork diffuses behind the replication fork, a process that twists the two new strands around one another and that is also sorted out by topoisomerases (see below).

COORDINATING REPLICATION OF THE TWO TEMPLATE STRANDS

The picture of the two strands of DNA replicating independently, as shown in Figure 1.8, does not take into consideration all of the coordination that must occur during DNA replication. The anatomy of the larger complex of replication factors remains unresolved; however, interactions among many of these components provide a hint as to how the larger complex functions (Figure 1.10). Rather than replicating independently, the DNA polymerases that produce the leading-strand and lagging-strand DNAs are joined to each other through the τ subunits of the holoenzyme (Table 1.1). In the holoenzyme there are two τ subunits and a derivative product called γ, which is incapable of interacting with DNA polymerase. The γ and τ subunits are encoded by the same gene, *dnaX*. Expression of the full gene results in production of the longer τ subunit, whereas a stutter in how the protein is produced from this gene, called a "frameshift," produces the shorter γ product. The configuration of having two τ and γ subunits may be important to ensure that only two DNA polymerase III molecules are at the replication fork, possibly facilitating the use of alternate polymerases for repair when needed (see below and Dohrmann et al., Suggested Reading).


A
DNA polymerase III
5′
3′
4
GTC
DnaN sliding clamp
DnaB helicase
DnaG primase
3′
5′
3
2
1
SSB
GTC
B
5′
3′
4
GTC
3′
5′
3
GTC
GTC
1
GTC
2
C
5′
3′
4
GTC
3′
5′
3
GTC
Free DnaN sliding clamp
GTC
GTC
2
1
D
5′
3′
3′
5′
DNA polymerase I
4
GTC
GTC
2
GTC
3
1

To accommodate the fact that the two DNA polymerases must move in opposite directions and still remain tethered, the lagging-strand template probably loops out as an Okazaki fragment is synthesized. The loop is then relaxed as the polymerase on the lagging-strand template is released from the β clamp, allowing the DNA polymerase to rapidly and efficiently "hop" ahead to the next RNA primer to begin synthesizing the next Okazaki fragment (Figure 1.10). The polymerase associates with a new β clamp assembled by the clamp loader at the site of the new RNA primer, while the old β clamp is left behind. The β clamp left on the last Okazaki fragment plays important roles in finishing synthesis and joining the fragments of lagging-strand DNA via interactions with DNA polymerase I, ligase, and repair proteins. β clamps are eventually recycled, possibly through the removal function of the δ subunit of the clamp loader. This model involving the looping out of the lagging-strand template has been referred to as the "trombone" model of replication because the loops forming and contracting at the replication fork resemble the extension and return of the slide of the musical instrument. The situation is probably similar in all bacteria and even the other domains of life, although in some other bacteria, including *Bacillus subtilis*, and in eukaryotes, different combinations of DNA polymerases are used to polymerize the leading and lagging strands (see Sanders et al., Suggested Reading). In the case of *B. subtilis*, an additional DNA polymerase interacts with DnaG and the replicative helicase, extending the RNA primer with DNA before handing the template off to the DNA polymerase used for the majority of DNA replication on both strands.

In addition to its role in loading β clamps onto template DNA, the clamp loader also plays an important role in coordinating the various replication components. Not only does the τ subunit of the clamp loader interact with DNA polymerase on the leading-strand and lagging-strand templates, it also interacts with the DnaB helicase (Figure 1.10). Further coordination on the lagging-strand template is facilitated by interactions between the DnaG primase and DnaB helicase (Figure 1.10). Coordination through the clamp loader helps to focus the energy from DNA polymerization with the energy that powers the helicase, allowing a high rate of DNA replication. The interaction between DNA polymerase III and DnaB governs the speed of unwinding so that it matches the rate of DNA polymerization to prevent undue exposure of single-stranded DNA.

THE GENES FOR REPLICATION PROTEINS

Most of the genes for replication proteins have been found by isolating mutants defective in DNA replication, but not RNA or protein synthesis. Since a mutant cell that cannot replicate its DNA will die, any mutation (for definitions of mutants and mutations, see "Replication Errors" below and chapter 3) that inactivates a gene whose product is required for DNA replication will kill the cell. Therefore, for experimental purposes, only a type of mutant called a **temperature-sensitive mutant** can be usefully isolated with mutations in DNA replication genes. These are mutants in which the product of the gene is active at one temperature but inactive at another. The mutant cells can be propagated at the temperature at which the protein is active (the permissive temperature). However, shifting to the other (nonpermissive) temperature can test the effects of inactivating the protein. The molecular basis of temperature-sensitive mutants is discussed in more detail in chapter 3.

The immediate effect of a temperature shift on a mutant with a mutation in a DNA replication gene depends on whether the product of the gene is continuously required for replication at the replication forks or is involved only in the initiation of new rounds of replication. For example, if the mutation is in a gene for DNA polymerase III or in the gene for the DnaG primase, replication ceases immediately. However, if the temperature-sensitive mutation is in a gene whose product is required only for initiation of DNA replication, for example, the gene for DnaA or DnaC (see "Initiation of Chromosome Replication" below), the replication rate for the population will

Figure 1.10 "Trombone" model for how both the leading strand and lagging strand might be simultaneously replicated at the replication fork. RNA primers are shown in green, and their initiation sites are shown as the sequence 3′-GTC-5′ boxed in blue. **(A)** The Pol III holoenzyme synthesizes lagging-strand DNA initiated from priming site 2 and runs into the primer at site 1. **(B)** The DNA strand undergoing lagging-strand replication loops out of the replication complex as the leading-strand polymerase progresses and the lagging-strand polymerase replicates toward the last Okazaki fragment. **(C)** Pol III has been released from the lagging-strand template at priming site 1 and has hopped ahead, leaving the old β clamps behind, and has reassembled with a new β clamp on the DNA at primer site 3 to synthesize an Okazaki fragment. Both the leading-strand and lagging-strand Pol III enzymes remain bound to each other and the helicase through interactions with τ during the release and reassembly process. **(D)** Pol III continues synthesis of the lagging strand from priming site 3 while Pol I is removing the primer at site 1 and replacing it with DNA. The Pol III holoenzyme hops to the primer at site 4 after reaching the primer at site 2. The primers and Okazaki fragments are not drawn to scale.

slowly decline. Unless the cells have been somehow synchronized in their cell cycles, each cell is at a different stage of replication, with some cells having just finished a round of replication and other cells having just begun a new round. Cells in which rounds of chromosome replication were under way at the time of the temperature shift will complete their replication cycle but will not start a new round. Therefore, the rate of replication decreases until the rounds of replication in all the cells are completed.

Replication Errors

To maintain the stability of a species, replication of the DNA must be almost free of error. Changes in the DNA sequence that are passed on to subsequent generations are called **mutations** (see chapter 3). Depending on where these changes occur, they can severely alter the protein products of genes or other cellular functions. To avoid such instability, the cell has mechanisms that reduce the error rate.

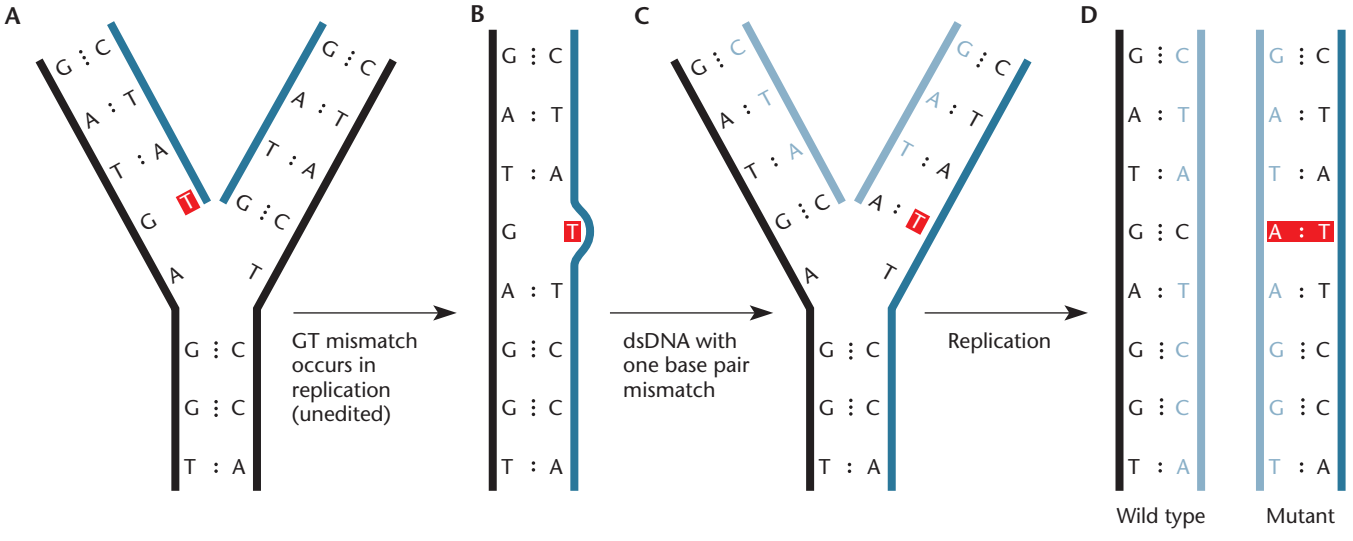
As DNA replicates, the wrong base is sometimes inserted into the growing DNA chain. For example, Figure 1.11 shows the incorrect incorporation of a T opposite a G. Such a base pair in which the bases are paired wrongly is called a **mismatch**. Mismatches can occur when the bases take on forms called **tautomers**, which pair differently from the normal form of the base (see chapter 3). After the first replication shown in Figure 1.11, the mispaired T pairs with an A, causing a GC-to-AT change in the sequence of one of the two progeny DNAs and thus changing the base pair at that position on all subsequent copies of the mutated DNA molecule.

Editing

One way the cell reduces mistakes during replication is through editing functions. With some DNA polymerases the editing function resides in the same protein, while in other cases a separate protein performs the editing function. Editing proteins are aptly named because they go back over the newly replicated DNA looking for mistakes and then recognize and remove incorrectly inserted bases (Figure 1.12). If the last nucleotide inserted in the growing DNA chain creates a mismatch, the editing function stops replication until the offending nucleotide is removed. DNA replication then continues, inserting the correct nucleotide. Because the DNA chain grows in the 5′-to-3′ direction, the last nucleotide added is at the 3′ end. The enzyme activity found in a DNA polymerase or one of its accessory proteins that removes this nucleotide is therefore called a 3′ exonuclease. The editing proteins probably recognize a mismatch because the mispairing (between T and G in the example in Figure 1.11) causes a minor distortion in the structure of the double-stranded helix of the DNA.

DNA polymerase I is an example of a DNA polymerase in which the 3′ exonuclease editing activity is part of the DNA polymerase itself. However, in DNA polymerase III, which replicates the bacterial chromosome, the editing function resides in an accessory protein encoded by a separate gene whose product travels along the DNA with the DNA polymerase during replication. In *E. coli*, the 3′ exonuclease editing function is encoded by the *dnaQ* gene (Table 1.1), and *dnaQ* mutants, also called *mutD* mutants (i.e., cells with a mutation in this gene that inactivates the 3′ exonuclease function), show much higher rates of spontaneous mutagenesis than do cells containing the wild-type, or normally functioning, *dnaQ*

Figure 1.11 Mistakes in base pairing can lead to changes in the DNA sequence called mutations. If a T is mistakenly placed opposite a G during replication **(A)**, it can lead to an AT base pair replacing a GC base pair in the progeny DNA **(B to D)**.



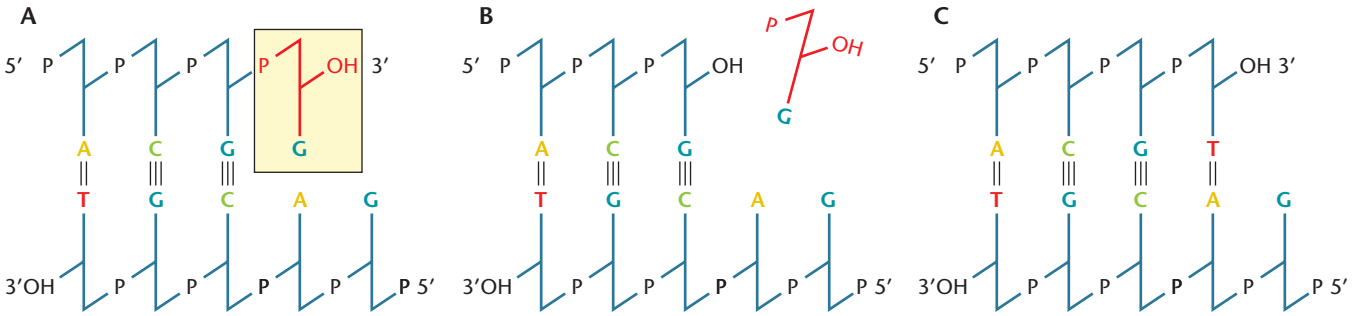


Figure 1.12 Editing function of DNA polymerase. **(A)** A G is mistakenly placed opposite an A while the DNA is replicating. **(B and C)** The DNA polymerase stops while the G is removed and replaced by a T before replication continues.

gene product. Because of their high spontaneous mutation rates, *mutD* mutants of *E. coli* can be used as a tool for mutagenesis, often combined with mutations in other genes whose products normally contribute to the correction of mismatches (see chapter 10).

RNA Primers and Editing

The importance of the editing functions in lowering the number of mistakes during replication may explain why DNA replication is primed by RNA rather than by DNA. When the replication of a DNA chain has just initiated, the helix may be too short for distortions in its structure to be easily recognized by the editing proteins. The mistakes may then go uncorrected. However, if the first nucleotides inserted in a growing chain are ribonucleotides rather than deoxynucleotides, an RNA primer is synthesized rather than a DNA primer. The RNA primer can be removed and resynthesized as DNA by using preexisting upstream DNA as a primer. Under these conditions, a distortion in the helix can be detected by the editing functions, and mistakes are avoided.

Another important system that safeguards the fidelity of the replication process is responsible for fixing mismatches after the growing DNA strand leaves the polymerase. In *E. coli* and its closest relatives, this process is guided by methylation and is termed **methyl-directed mismatch repair**. Related mismatch repair systems are used across all three domains of life, but the use of methylation signals is not widespread. The methyl-directed mismatch repair system is discussed in chapter 10.

Impediments to DNA Replication

While the process described above and diagrammed in Figure 1.10 would suffice for pristine DNA on a template that lacked any type of physical block to the progression of the DNA replication complex, in reality, the situation in the cell is rarely this tidy. DNA polymerases frequently encounter a number of different problems. Challenges to DNA polymerases include interruptions in the DNA template, bulky adducts that cannot be replicated by DNA polymerase III, and physical blocks mediated by supercoiling and proteins bound to, or acting on, the chromosome. One extreme form of impediment to a replication fork comes from **nicks** in either the leading-strand or lagging-strand template DNAs in which the phosphate-deoxynucleotide chain is broken in one strand of the DNA. These nicks cause the destruction of the nick-containing arm of the replication fork, resulting in a broken chromosome and collapse of the DNA replication fork. Bacteria possess a highly efficient mechanism for priming repair of this broken end by using the broken DNA itself as a primer to reinitiate DNA replication. This process was likely the original driving force for the evolution of recombination and is described in chapter 9.

Damaged DNA and DNA Polymerase III

DNA polymerase III replicates DNA with incredibly high fidelity. Much of the fidelity of the enzyme comes from the structure of the catalytic pocket, where there is a **pre-synthetic** check for base pairing between the template strand and the incoming nucleotide. A side effect of this small binding pocket is the inability of the polymerase to tolerate **lesions** in which chemical changes have occurred in the base, the deoxyribose sugar, or even the phosphate on the DNA. There are many mechanisms for DNA replication to continue even when a cell is grown under conditions that result in highly damaged DNA. While early work suggested that the polymerization of the leading and lagging strand was so tightly coupled that a lesion on one strand would stop the entire DNA replication fork, more recent work indicates flexibility. It is now clear that although the polymerases producing the leading strand and lagging strand are physically coupled, the two complexes can be momentarily uncoupled by leaving a single-strand DNA gap at the point of the lesion that blocked one of the DNA polymerases. Other processes can repair these gaps, and in extreme cases, where there is extensive damage in the chromosome, these gaps initiate a DNA damage response called the SOS response (see chapter 10).

Mechanisms To Deal with Impediments on Template DNA Strands

The mechanisms used for momentarily functionally uncoupling synthesis of the two strands differ depending on whether the lesion occurs on the leading-strand or lagging-strand template. The discontinuous nature of replication on the lagging-strand template affords the opportunity to circumvent lesions that halt DNA polymerase III. Typically, DNA polymerase III is recycled onto a new DNA primer when a new RNA primer is deposited (Figure 1.10). However, a stalled lagging-strand DNA polymerase III can also be recycled by premature release when it stalls at DNA damage (Figure 1.13A). The single-strand DNA gap left behind is repaired by another mechanism.

Under the historical model of the function of DnaG primase, primers are placed only on the lagging-strand template. However, biochemical studies indicate that in cases where the leading-strand polymerase stalls, primase can also produce an RNA primer on the leading-strand template, allowing replication to continue but leaving a gap on this strand (Figure 1.13B). This process of lesion skipping on the leading-strand template and the ability to utilize alternative polymerases (described below) to copy over DNA damage provide complementary mechanisms to deal with damaged DNA template strands (see Gabbai et al., Suggested Reading).

While DNA polymerase III and DNA polymerase I are important for high-fidelity DNA replication, other DNA polymerases are found in *E. coli* that allow replication through damaged DNA in a process known as **translesion synthesis**. Most translesion polymerases appear to come with a trade-off in which the ability to copy damaged DNA results in a lower fidelity of DNA replication. As expected, the expression of these polymerases is induced as a response to DNA damage in the cell. In addition to controlling the amount of translesion polymerase present in the cell, access to the DNA replication fork by polymerases other than DNA polymerase III is regulated by a process called **polymerase switching**, a process by which one DNA polymerase replaces a polymerase already found at the 3′ OH end of a primed DNA template (Figure 1.13C). In *E. coli*, DNA polymerases II, IV, and V can be recruited to temporarily step in for DNA polymerase III at damaged DNA (more details of this system are described in chapter 10). Each of these polymerases has different attributes, ranging from fairly accurate and highly processive (DNA polymerase II) to very inaccurate and not very processive (DNA polymerases IV and V). Processivity refers to how far a DNA polymerase moves on the template before falling off.

Having multiple DNA polymerases with different properties appears to be common in all living organisms. How accurate or processive a given DNA polymerase is may also depend on the nature of the damage found in the template DNA and/or the availability of various accessory proteins. The regulation of the use of these DNA polymerases is still incompletely understood, but there appear to be highly evolved processes in which the system has been fine-tuned by the process of natural selection over a long time so that the most appropriate DNA-copying mechanism is used for each environmental challenge.

Physical Blocks to Replication Forks

Proteins bound to, or otherwise acting on, the chromosome can also stop the progression of DNA polymerase III. A programmed block to DNA replication occurs in some bacteria to terminate DNA replication within one region of the chromosome (see below). However, unintended blocks can occur in other situations, such as when DNA polymerase encounters RNA polymerase carrying out transcription (see chapter 2) or when RNA polymerase is stalled at sites of damage in the chromosome. Transcription appears to be the most significant impediment to replication, but a variety of protein complexes need to be displaced ahead of the DNA replication fork.

In addition to DnaB, other helicases help DNA replication to proceed through obstacles on template DNAs. Experimentally, it was shown that two other *E. coli* helicases, Rep and UvrD, allow replication to proceed through a protein block on template DNA (see Guy et al., Suggested Reading). These helicases travel in the 3′-to-5′ direction, and therefore, they are likely to travel on the leading-strand template while DnaB progresses forward on the lagging-strand template (i.e., the 5′-to-3′ direction). Rep interacts directly with the DnaB helicase, probably as a normal part of the **replisome**. The replisome is the collection of proteins that interact with one another (through DNA or other proteins) and that are involved in carrying out DNA replication. In *E. coli*, UvrD may be a helicase of general use for helping out when DNA replication is blocked by proteins or to remove recombination structures from the chromosome to allow DNA replication to proceed. As explained in more detail in chapter 10, UvrD allows strand displacement during mismatch repair and plays additional roles in nucleotide excision repair in *E. coli* in a repair process that is coupled with transcription.

Replication of the Bacterial Chromosome and Cell Division

So far, we have discussed the details of DNA replication, but we have not discussed how bacterial DNA as a whole is replicated, nor have we discussed how the replication process is coordinated with division of the bacterial cell. To simplify the discussion, we first consider only bacteria that grow as individual cells and divide by binary fission to form two cells of equal size, even though this is far

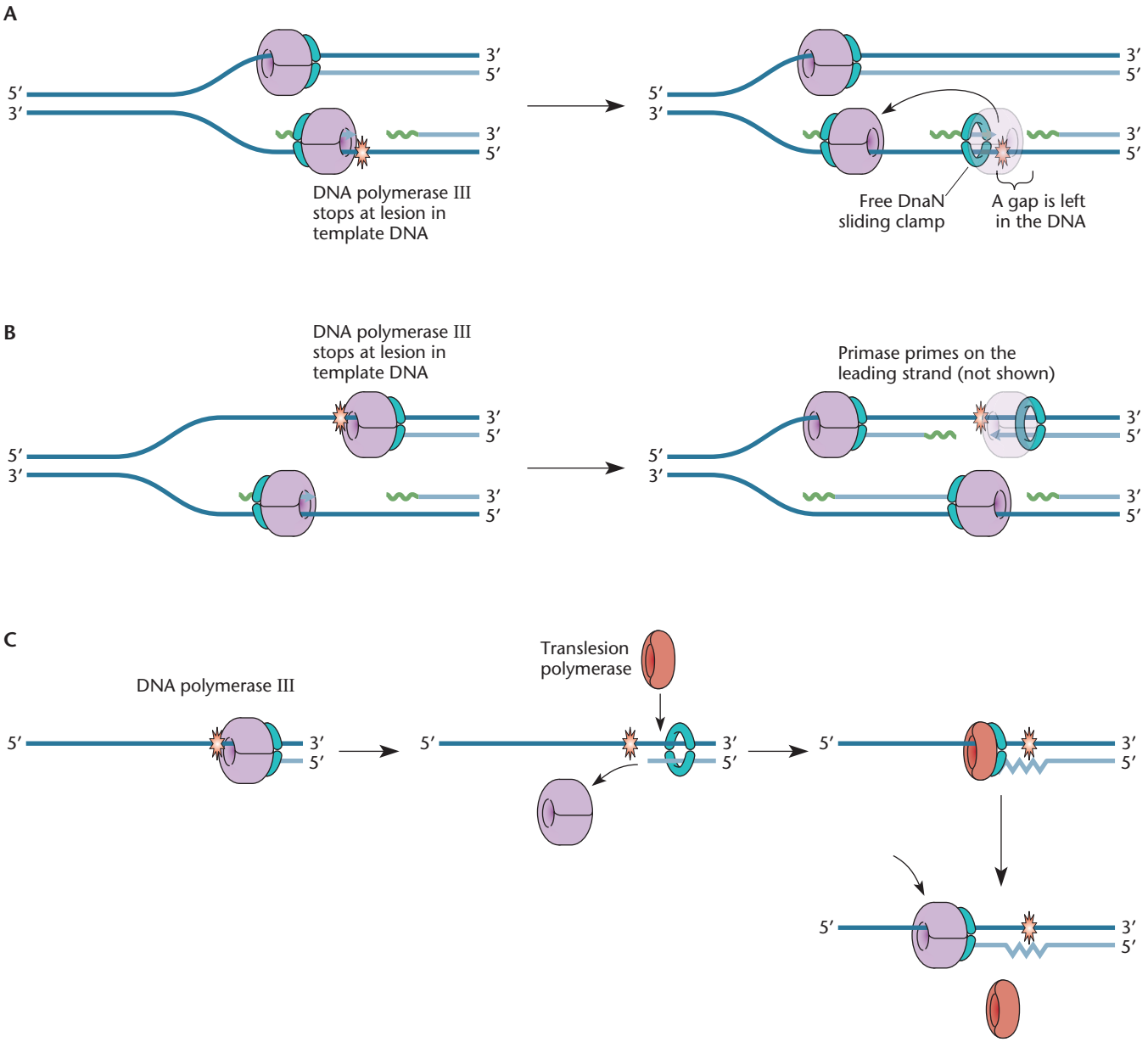


Figure 1.13 Physical blocks on template DNAs. **(A)** When DNA polymerase III stalls on the lagging-strand template strand, a new Okazaki fragment can be initiated and the stalled polymerase can be moved to the new RNA primer. The process leaves a gap that must be repaired by other means, probably involving RecFOR (see chapter 9). **(B)** When DNA polymerase III stalls on the leading strand, primase can restart DNA replication, leaving a gap that must be repaired by other means. **(C)** In some cases where DNA polymerase III stalls at damaged DNA, translesion polymerase can utilize the same sliding clamp to replicate through the lesion, often by error-prone DNA replication. The translesion polymerases have low processivity and fall off the template after a short distance, allowing the accurate DNA polymerase III replication to resume high-fidelity DNA replication.

from the only type of multiplication observed among bacteria.

The replication of the bacterial DNA occurs during the **cell division cycle**, which is the time during which a cell is born, grows larger, and divides into two progeny cells. Cell division is the process by which the larger cell splits into the two new cells. The **division time**, or **generation time**, is the time that elapses from the point when a cell is born until it divides. This time is usually approximately the same for all the individuals in the population

under a given set of growth conditions. The original cell before cell division is called the mother cell, and the two progeny cells after division are called the daughter cells.

Structure of Bacterial Chromosomes

The DNA molecule of a bacterium that carries most of its normal genes is commonly referred to as its **chromosome**, by analogy to the chromosomes of higher organisms. This name distinguishes the molecule from plasmid DNA, which in some cases can be almost as large as chromosomal DNA but usually carries genes that are not always required for growth of the bacterium (see chapter 4). Most bacteria have only one chromosome; in other words, there is only one unique DNA molecule per cell that carries most of the normal genes. There are exceptions, and it is estimated that 10% of bacteria have more than one chromosome, including *Vibrio cholerae*, the bacterium responsible for the disease cholera. Even in bacteria that contain multiple chromosomes, the second chromosome shows more characteristics of a plasmid than of a chromosome, particularly in how it initiates replication. There appear to be special molecular systems for managing multiple chromosomes (see Fournes et al., Suggested Reading).

As discussed below, when bacteria, such as *E. coli*, are reproducing very rapidly, new rounds of replication initiate before others are completed, temporally increasing the DNA content of the cells until cellular division returns the number of copies to one unit chromosome per cell. It is important to note, however, that these individual chromosomal DNAs are not unique since they are directly derived from each other by replication.

The structure of bacterial DNA differs significantly from that of the chromosomes of higher organisms. One difference is that the DNA in the chromosomes of most bacteria is circular in the sense that the ends are joined to each other (for exceptions, see Box 4.1). In contrast, eukaryotic chromosomes are usually linear with free ends. As discussed in chapter 4, the circularity of bacterial chromosomal DNA allows it to replicate in its entirety without using telomeres, as eukaryotic chromosomes do, or terminally redundant ends, as some bacteriophages do (see chapter 7). Even in cases where bacterial chromosomes are linear, they do not use the same mechanism, involving telomerases to replicate their ends, that is used by eukaryotic chromosomes. Another difference between the DNA of bacteria and eukaryotes is that the DNA in eukaryotes is wrapped around proteins called histones to form nucleosomes. Bacteria have the proteins HU, HN-S, Fis, and IHF, around which DNA is often wrapped, and archaea do have rudimentary histones related to those of eukaryotes. However, in general, DNA is much less structured in bacteria than in eukaryotes.

Replication of the Bacterial Chromosome

Replication of a circular bacterial chromosome initiates at a unique origin of chromosomal replication, or ***oriC***, and proceeds in both directions around the circle. On the *E. coli* chromosome, *oriC* is located at 84.3 min. As mentioned above, the place in DNA at which replication occurs is known as the replication fork. Two replication forks start at *oriC* and proceed around the circle until they meet and **terminate** chromosomal replication. The DNA polymerases responsible for replicating the leading and lagging strands associate as a single holoenzyme. However, there is no association between the DNA polymerases at the two DNA replication forks to help drive the separation of chromosomes, and therefore, other force-generating mechanisms must be at play. As discussed in "Termination of Chromosome Replication" below, some bacteria actively terminate replication at a unique site in the DNA; however, these systems are not widespread, and most bacteria terminate replication using an unknown mechanism or simply terminate DNA replication where the two replication forks meet. Each time the two replication forks proceed around the circle and meet, a **round of replication** has been completed, and two new DNAs, called the **daughter DNAs**, are generated.

Initiation of Chromosome Replication

Much has been learned about the molecular events occurring during the initiation of replication. Some of this information has a bearing on how the initiation of chromosome replication is regulated and serves as a model for the interaction of proteins and DNA.

Two types of functions are involved in the initiation of chromosome replication. One consists of the sites or sequences on DNA at which proteins act to initiate replication. These are called ***cis*-acting sites**. The prefix *cis* means "on this side of," and these sites act only on the same DNA. The proteins involved in initiation of replication are examples of ***trans*-acting functions**. The prefix *trans* means "on the other side of," and these functions can act on any DNA in the same cell, not just the DNA from which they were made. The concepts of *cis*- and *trans*-acting properties are common in molecular genetics, and these references are used throughout this book.

ORIGIN OF CHROMOSOMAL REPLICATION

One *cis*-acting site involved in DNA replication is the *oriC* site, at which replication initiates. The sequence of *oriC* is well defined in *E. coli*, and the basic components that make up *oriC* are broadly similar in most bacteria. Figure 1.14 shows the structure of the origin of replication of *E. coli*. Less than 260 base pairs (bp) of DNA is required for initiation at this site. Within *oriC* are a series of binding sites for various proteins; the most important

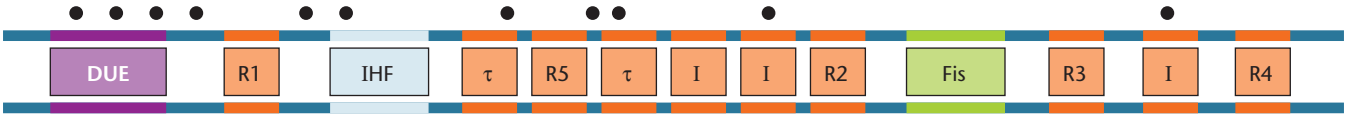


Figure 1.14 Structure of the origin of chromosomal replication (*oriC*) region of *Escherichia coli*. Shown are the positions of multiple types of DnaA-binding sequences, five DnaA boxes (R1 to R5) and other DnaA-binding sites (I and τ), and an AT-rich region that is unwound to allow loading of the replication apparatus, the DNA-unwinding element (DUE). Also shown are binding sites for the IHF and Fis proteins and a large number of GATC sites (black dots) that are important in regulating initiation by acting as sites of Dam methylation.

of these binding proteins is the master initiator protein in bacteria, called DnaA (see below). The **canonical** DnaA-binding sequences are 9 bp in length, and these sites are termed **DnaA boxes**. While five DnaA boxes exist within *oriC*, three of these sites bind DnaA particularly strongly, i.e., are of particularly high affinity, and are always bound by DnaA (Figure 1.14). The ability of DnaA to bind these high-affinity sites at all times can be considered analogous to the origin recognition complex associated with eukaryotes. Additional sites called "I" and "τ" sites, which differ from DnaA boxes, exist in the origin region but are occupied only by DnaA that is bound to ATP and not ADP (see below) (Figure 1.14). Finally, within an AT-rich region of DNA that is opened for initiation, called the DNA-unwinding element, are three additional sites for DnaA binding that are occupied only when DnaA is bound to ATP and not ADP. Binding sites for other DNA-binding proteins (IHF and Fis) are also found in this region.

INITIATION PROTEINS

Besides the *cis*-acting *oriC* site and DnaA, many *trans*-acting proteins are also required for the initiation of DNA replication, including the DnaB and DnaC proteins. DnaA is required only for initiation, allowing DnaC to load the DnaB helicase for establishing the DNA replication forks. Many proteins used in other cellular functions are also involved, such as the primase (DnaG), the normal RNA polymerase that makes most of the RNA in the cell, and the DNA-binding proteins IHF and Fis (Figure 1.14).

Figure 1.15 outlines how DnaA, DnaB, DnaC, and other proteins participate in the initiation of chromosome replication. As we will see throughout the remainder of this chapter, there are many points at which the initiation of DNA replication is controlled. One important regulatory consideration concerns the nucleotide-binding state of DnaA. While DnaA always binds some of the DnaA boxes, for initiation of DNA replication all of the boxes are bound, forming a special architecture that can open the DNA strands. This type of binding requires that DnaA be bound to ATP (DnaA-ATP). In biology, there are many examples where the nucleotide-bound state of a protein determines its activity. Proteins of this type have the capacity to hydrolyze nucleotides from the NTP to the NDP form, but the energy released is not directly used to actively carry out any particular task and instead allows the configuration of the proteins to change. In the case of DnaA, the ATP-bound form of the protein allows it to form a large multimer structure composed of many molecules of DnaA protein, where the DNA strands are opened through bending of the DNA by DnaA with the help of the IHF and Fis proteins. Within the special complex that productively opens the DNA strands in the origin, DnaA binding appears to take on a different form when it interacts with DUE (DNA unwinding element), preferentially engaging single-strand DNA in this region to facilitate strand opening (Figure 1.15). The binding and opening are also aided by supercoiling at the origin (see "Supercoiling" below) and by the SSB protein, which helps to keep the helix from reforming. DnaA binds directly to the helicase DnaB, and in a process involving DnaC, DnaB helicase is loaded onto *oriC*. Action of the DnaB helicase opens the strands further for priming and replication, and DnaC leaves the complex.

RNA Priming of Initiation

As described above, RNA primers are continuously needed during the DNA replication process. The complex that travels along the chromosome laying down RNA primers is called the **primosome** and contains DnaB and DnaG primase. The RNA polymerase that synthesizes most of the RNA molecules, including mRNA, in the cell (see chapter 2) is needed to initiate rounds of replication; however, transcription from RNA polymerase in this context from adjacent genes is involved in controlling the separation of the strands of DNA in the *oriC* region. DnaG primase is responsible for laying down RNA primers for DNA synthesis after replication is initiated at *oriC*.

Termination of Chromosome Replication

After replication of the chromosome initiates in the *oriC* region and proceeds around the circular chromosome in both directions, the two replication forks must meet somewhere on the chromosome and the two daughter chromosomes must separate. In some bacteria, including

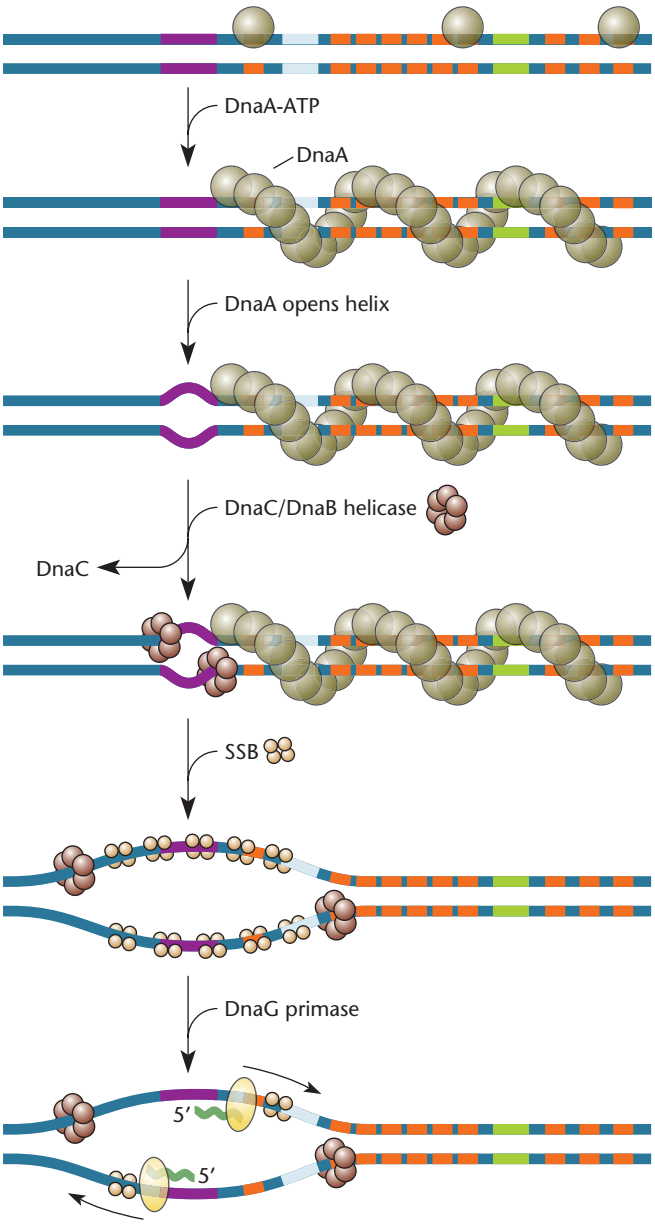
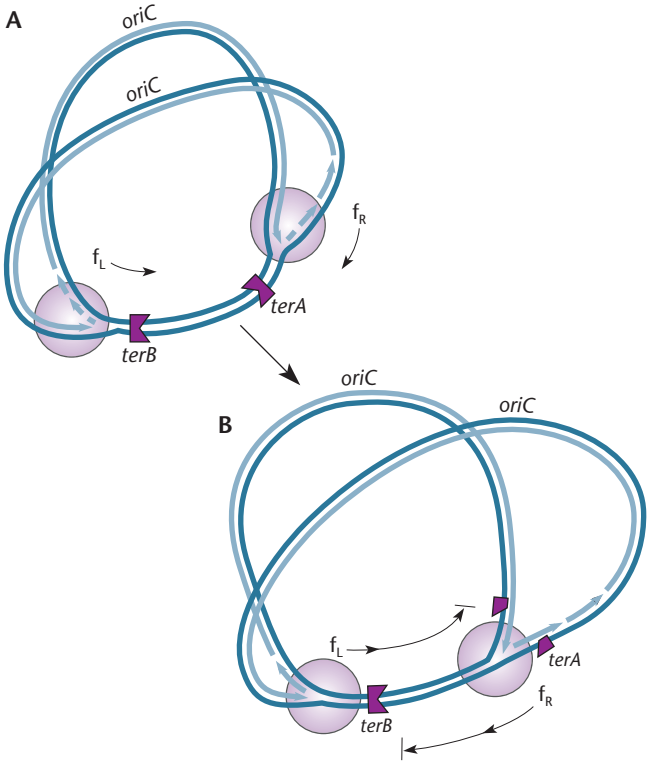


Figure 1.15 Initiation of replication at the *Escherichia coli* origin (*oriC*) region. DnaA is always bound to three DnaA boxes within *oriC*, even when DnaA is in its non-ATP-bound state acting as an origin recognition complex. About a dozen DnaA-ATP proteins bind to the origin, possibly by forming a type of helical filament and opening the helix and the DNA-unwinding element. DnaC helps the DnaB helicase to bind. The DnaG primase synthesizes RNA primers, initiating replication.

E. coli and *B. subtilis*, a specific system exists to control the region where replication forks meet. What happens in other organisms is less clear. As with most cellular processes, the process of termination of chromosome replication is especially well understood in *E. coli*. In this bacterium, termination is facilitated by DNA sequences, called *ter* sites, that are only ~22 bp long. These sites act somewhat like the one-way gates in an automobile parking lot, allowing the replication forks to pass through in one direction but not in the other.

Figure 1.16 shows how the one-way nature of *ter* sequences can cause replication to terminate in a specific region of the chromosome. In the illustration, two *ter* sites called *terA* and *terB* bracket the termination region. Replication forks are unaffected by the *terA* site in the clockwise direction but are terminated in the counterclockwise direction. The opposite is true for *terB*. Thus, the clockwise-moving replication fork progresses through *terA*, but if it gets to *terB* before it meets the counterclockwise-moving fork, it stalls, because it cannot move clockwise through *terB*. Similarly, the replication fork moving in the counterclockwise direction stalls at the *terA* site and waits for the clockwise-moving replication fork. When the counterclockwise and clockwise replication forks meet, at *terA*, *terB*, or somewhere between them, the two forks terminate replication, releasing the

Figure 1.16 Termination of chromosome replication in *Escherichia coli*. **(A)** The replication forks that start at the origin of chromosomal replication (*oriC*) can traverse *terA* and *terB* in only one direction, opposite that indicated by the black arrows. **(B)** When they meet, between or at one of the two clusters, chromosome replication terminates. f_L is the fork that initiated to the left and moved in a counterclockwise direction. f_R is the fork that initiated to the right and moved in a clockwise direction. Adapted from Camara JE, Crooke E, *in* Higgins MP, ed, *The Bacterial Chromosome* (ASM Press, Washington, DC, 2005), with permission.



two daughter DNAs. In *E. coli*, it is known that most DNA replication termination occurs at one *ter* site, possibly because it is oriented to terminate replication forks traveling in the clockwise direction, which is shorter in most laboratory *E. coli* strains.

Encountering a *ter* DNA sequence, by itself, is not sufficient to stop the replication fork. A protein is also required to terminate replication at *ter* sites. The protein that works with *ter* sites, called the terminus utilization substance (Tus) in *E. coli* and the replication terminator protein (RTP) in *B. subtilis*, binds to the *ter* sites and stops the replicating helicase (DnaB in *E. coli*) that is separating the strands of DNA ahead of the replication fork. In both *E. coli* and *B. subtilis*, multiple *ter* sites bracket the terminus region, helping to ensure that replication proceeds from *oriC* to the terminus region (see Box 1.1). While the *ter* systems are not absolutely essential for *E. coli* and *B. subtilis* growing in the laboratory setting, they are important for other aspects of genome stability that are important for maintaining genome integrity over time in the natural environment (see Rudolph et al., Suggested Reading). It has been argued that the active termination systems involving *ter* sites and a *trans* acting protein originated in plasmids (chapter 5) and were domesticated in some branches of bacteria for use in the bacterial chromosome (see Galli et al., Suggested Reading).

Chromosome Segregation

While bacteria do not contain a special membrane compartment for chromosomal DNA like the nucleus of eukaryotes, even in bacteria the chromosome does not freely diffuse within the cytoplasm. In fact, as we learn more about bacterial chromosomes, we are realizing that they are maintained with an incredible amount of organization. Even with the aid of only a standard laboratory microscope and DNA stain, a mass of chromosomal DNA is easily observed in the center of the cell in a structure called the **nucleoid**, which is very compact, considering that the DNA is about a thousand times as long as the cell.

Because of the large size of the chromosome, the process of moving the replicating chromosomes to daughter cells, called **segregation**, is not trivial. Chromosome segregation encounters a number of obstacles. Obvious initial obstacles are viscous forces and torsional stress associated with unwinding the template strands of DNA. Advances in microscopy and techniques that allow the localization of certain regions of the chromosome are revealing the choreography involved in coordinating DNA

BOX 1.1

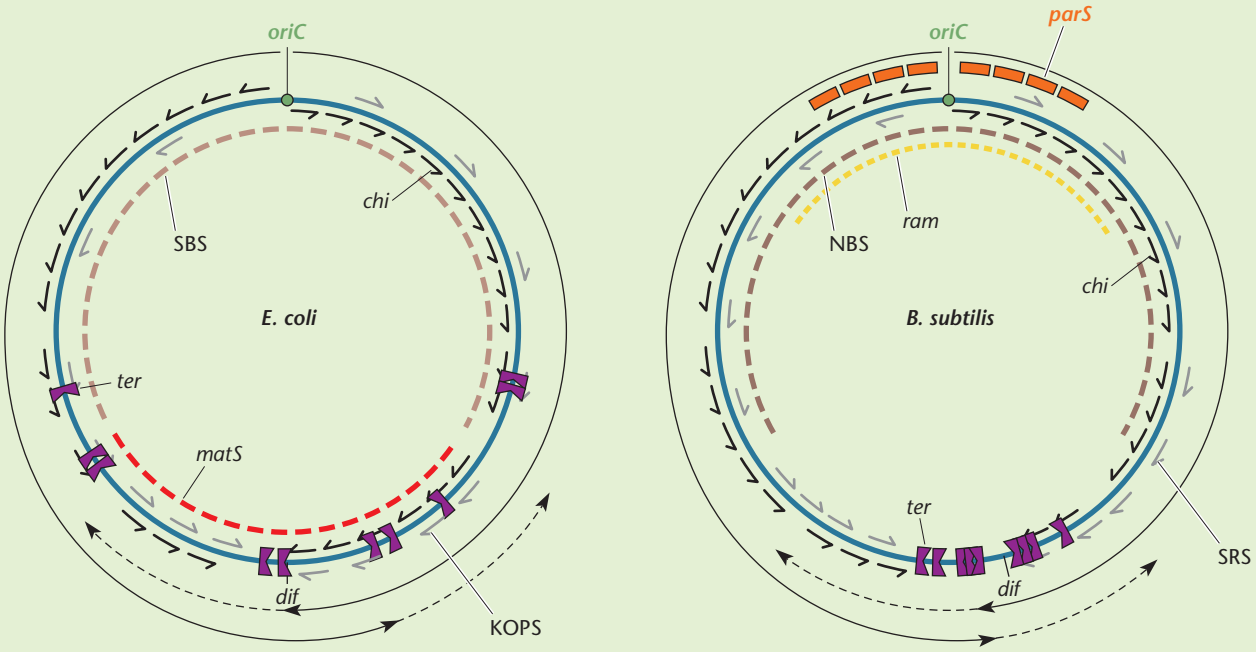
Structural Features of Bacterial Genomes

It is widely appreciated that the chromosomes are the information storehouse for an organism. What is less appreciated is that the chromosome as a structure has evolved sequence motifs that allow it to be efficiently replicated, repaired, and segregated into daughter cells. The distribution and orientation of these motifs are discussed here; the molecular biology of the systems that recognize these sequences is explained in greater detail in the text. The placement of these sequence motifs in the context of the chromosome is important for their function, as is the orientation of many of these sequences. Many of the motifs are oriented in one direction, which follows the direction of the DNA replication fork. DNA replication in *E. coli* and *B. subtilis* (and all bacteria studied to date) is initiated within a single *oriC* region and continues bidirectionally to a position on the chromosome equidistant from the origin (indicated by the long arrow-headed line in the figure). The *dif* site where the resolution of dimer chromosomes occurs is found near to where DNA replication normally terminates.

Certain DNA sequence motifs that guide DNA replication and DNA repair are polar in that they are not symmetrical and need to be in a specific 5′-to-3′ direction to carry out their functions. In other words, these sequences must be found in a certain orientation in the chromosome and will not work if they are flipped around. In *E. coli* and *B. subtilis*, DNA replication forks are actively terminated at specific sites called *ter* sites. The *ter* sites act as a trap for DNA replication forks, and these sites encompass a large portion of the chromosome, allowing DNA replication forks to pass when approaching from one direction but not the other. Multiple *ter* sites (10 in *E. coli* and 9 in *B. subtilis*) are found in the genome, and the redundancy of these sites may be important for catching replication events that get through the initial *ter* sites (indicated with a dashed line) or to stop replication forks that are initiated for DNA repair by recombination (see chapter 9). For unknown reasons, the central *ter* sites in *B. subtilis* are very close together, while the central sites in *E. coli* are separated by hundreds of thousands of base pairs. Along the path of DNA replication are sequence motifs involved in guiding the DNA translocase proteins involved in chromosome segregation: FtsK, found in *E. coli* (which recognizes motifs called KOPS), and SpoIIIE, found in *B. subtilis* (called SRS motifs). The StpA DNA translocase from *B. subtilis* may also recognize the SRS sites. Chromosomes also

(continued)

Structural Features of Bacterial Genomes



have polar DNA sequences that guide the recombination machinery (see chapter 9). DNA recombination is extremely important in bacteria as a way to repair DNA double-strand breaks that occur during DNA replication. Repair of these breaks occurs when recombination reestablishes a DNA replication fork using one broken end and the sister chromosome. Reestablishment of DNA replication involves an efficient processing event that utilizes polar sites called *chi* sites. The RecBCD complex in *E. coli* or the AddAB complex in *B. subtilis* carries out this processing activity using information found in the *chi* sites. *chi* sites are species specific and are common in genomes in one orientation from the origin to the terminus region on the leading strand (found about 1 every 5 kb in the *E. coli* chromosome).

Other polar sequence biases in the chromosome include an overrepresentation of the 5′-CTG-3′ sequence that primes lagging-strand DNA synthesis (not shown). Interestingly, the most common triplet codon is the CUG (5′-CTG-3′ in DNA) that codes for leucine, comprising almost 5% of all codons in *E. coli*. The 5′-CTG-3′ sequence is found in the *chi* sequence and all of the most frequent 8-bp sequences in the chromosome. It would be difficult to argue which came first, the use of this sequence by the primase or its frequency of use as a codon. Another type of sequence bias, but one that is not polar in nature, is a general sequence bias called the G/C skew, where G and C are overrepresented in the leading strand. The trend toward A and T in the lagging strand is believed not to have an adaptive value but to be a result of the way in which repair differs on the two strands.

There are other DNA sequences that are not polar but that show biases for regions of the chromosome. Around the origin of *B. subtilis*, an area recognized by the Spo0J protein for segregation of the origin region to daughter cells, the *parS* sequence, is found 8 times (the orange rectangular boxes in the figure). Also around the origin of *B. subtilis*, there is an enrichment of *ram* sites (short yellow dashes), a sequence recognized by the RacA protein for maintaining segregation during sporulation. Binding sites for the nucleoid occlusion proteins which prevent septum formation until division is nearly complete reside across the chromosome but are absent from the terminus region. SlmA-binding sites (SBS) in *E. coli* and Noc-binding sites (NBS) in *B. subtilis* (long brown dashes) are recognized by SlmA and Noc, respectively. The GATC sites (not shown), which are important for regulating the initiation of DNA replication at *oriC* in *E. coli*, also show enrichment in the *oriC* region, with a spacing that is important for SeqA binding. The organization of the large domain comprising the terminus region of the chromosome appears to be important in *E. coli*, where the MatP protein recognizes *matS* sites (red dashes in the *E. coli* diagram) found across this region.

References

Blattner FR, Plunkett G III, Bloch CA, Perna NT, Burland V, Riley M, Collado-Vides J, Glasner JD, Rode CK, Mayhew GF, Gregor J, Davis NW, Kirkpatrick HA, Goeden MA, Rose DJ, Mau B, Shao Y. 1997. The complete genome sequence of *Escherichia coli* K-12. *Science* **277:**1453–1462.

Touzain F, Petit M-A, Schbath S, El Karoui M. 2011. DNA motifs that sculpt the bacterial chromosome. *Nat Rev Microbiol* **9:**15–26.

replication and chromosome segregation. Microscopy experiments using green fluorescent protein (GFP) fused to replication proteins allow the localization of DNA replication forks, and GFP fused to proteins that bind to specific sites on DNA allows localization of the origin and terminus in the cell. These experiments show that soon after the initiation of DNA replication within the nucleoid, the origins start to move to the daughter cells. Once replication is complete, segregation would still be prevented if daughter chromosomes were joined by recombination, interlinked, or otherwise tangled during replication. Even if they were not physically joined, their separation would be very difficult if the two daughter chromosomes were randomly spread out throughout the cell. It is therefore not surprising that bacteria have a number of systems to ensure that their chromosomes segregate properly into the daughter cells during cell division. Molecular systems responsible for chromosome segregation are discussed separately below.

RESOLUTION OF DIMER CHROMOSOMES

During the process of DNA replication, two circular daughter chromosomes will periodically become joined, forming a chromosome **dimer**, in which they are joined end to end to form a double-length circle. Dimer chromosomes can result from recombination between the two replicating chromosomes and are fairly common. Recombination involved in restarting stalled DNA replication forks from a sister chromosome probably accounts for many of these events (see chapter 9). Such dimers prevent chromosome segregation because the two daughter chromosomes are part of the same large molecule.

If dimer chromosomes can be created by recombination, they can also be resolved into the individual chromosomes by a second recombination event. The general recombination system could in theory resolve the dimers between sister chromosomes by recombination; however, the general recombination system can both create and resolve dimers, depending on how many crossovers occur between the daughter DNAs. An odd number of crossovers occurring between any two sequences on the two daughter DNAs in the dimer will resolve the dimer, but an even number of crossovers will recreate a dimer.

All bacteria with a circular chromosome appear to have a system dedicated to dimer resolution. In *E. coli*, and in most bacteria, the so-called Xer recombination system is used to resolve chromosome dimers. Rather than using the general recombination system, the Xer systems involve a site-specific recombinase (see chapter 8) to resolve chromosome dimers. This system has evolved so that it resolves dimers into the individual chromosomes but does not create new dimers. Its action is also coordinated with division of the cell and other important chromosome-partitioning functions. The Xer recombination system consists of two proteins called XerC and XerD and a specific site in the chromosome called *dif*. If two copies of the *dif* site occur on the same DNA, as happens when the chromosome has formed a dimer, the Xer proteins promote recombination between the two *dif* sites, resolving the dimer into the individual chromosomes (Figure 1.17). The *dif* site is always found centrally located in the *ter* region in bacterial chromosomes. This is likely, in part, to help ensure that there is only one *dif* site in the cell until just before cell division so that it is not replicated until just before the chromosome has completed replication and just before the cell divides. As added insurance, the activity of the Xer site-specific recombination system is also dependent on the formation of the division septum through an interaction with the FtsK protein. As diagrammed in Figure 1.17, FtsK protein is localized to the region where the division septum pinches in during cell division, where it plays multiple roles, including facilitating dimer resolution. While the Xer proteins are needed for dimer resolution, they are actually active for full dimer resolution only when they interact with FtsK. As shown in Figure 1.17, the localization of FtsK at the septum limits the dimer resolution process temporally and spatially to when the daughter chromosomes are in the process of moving through the septum into daughter cells, a process that is facilitated and coordinated by FtsK itself (see below) (see Aussel et al., Suggested Reading).

The FtsK protein is a DNA translocase that can pump DNA through itself to help align the two *dif* sites in a dimer chromosomal DNA at the septum in the middle of the cell before they recombine, thereby facilitating segregation into daughter cells (Figure 1.17). See the reconstructed image of the structure of the motor domain of FtsK at the start of this chapter (see also Jean et al., Suggested Reading). An obvious question is how the FtsK protein "knows" which direction pumps to engage for transporting DNA in the correct direction to move a *dif* site to the septum. It does this by using sites on the DNA as an orientation cue. These sites are called the KOPS sites, for FtsK-orienting polar *s*equence. Polar means that they read differently in one direction on the DNA than in the other. These sequences are oriented in the DNA so that they will only be read progressing from the origin to the *dif* site in the chromosome; the FtsK protein, by translocating the DNA in the direction pointed to by the KOPS sites, will pump the DNA toward the *dif* sites close to the terminus of replication. Note in Figure 1.17 how the polar KOPS sites in the chromosome (shown as half-arrows) can be used as directional information by FtsK to mobilize the chromosome to the daughter cells while moving the *dif* sites toward the septum. This is just one example of how the chromosome, typically thought of as the informational storehouse of the cell, also contains structural information used by the many proteins

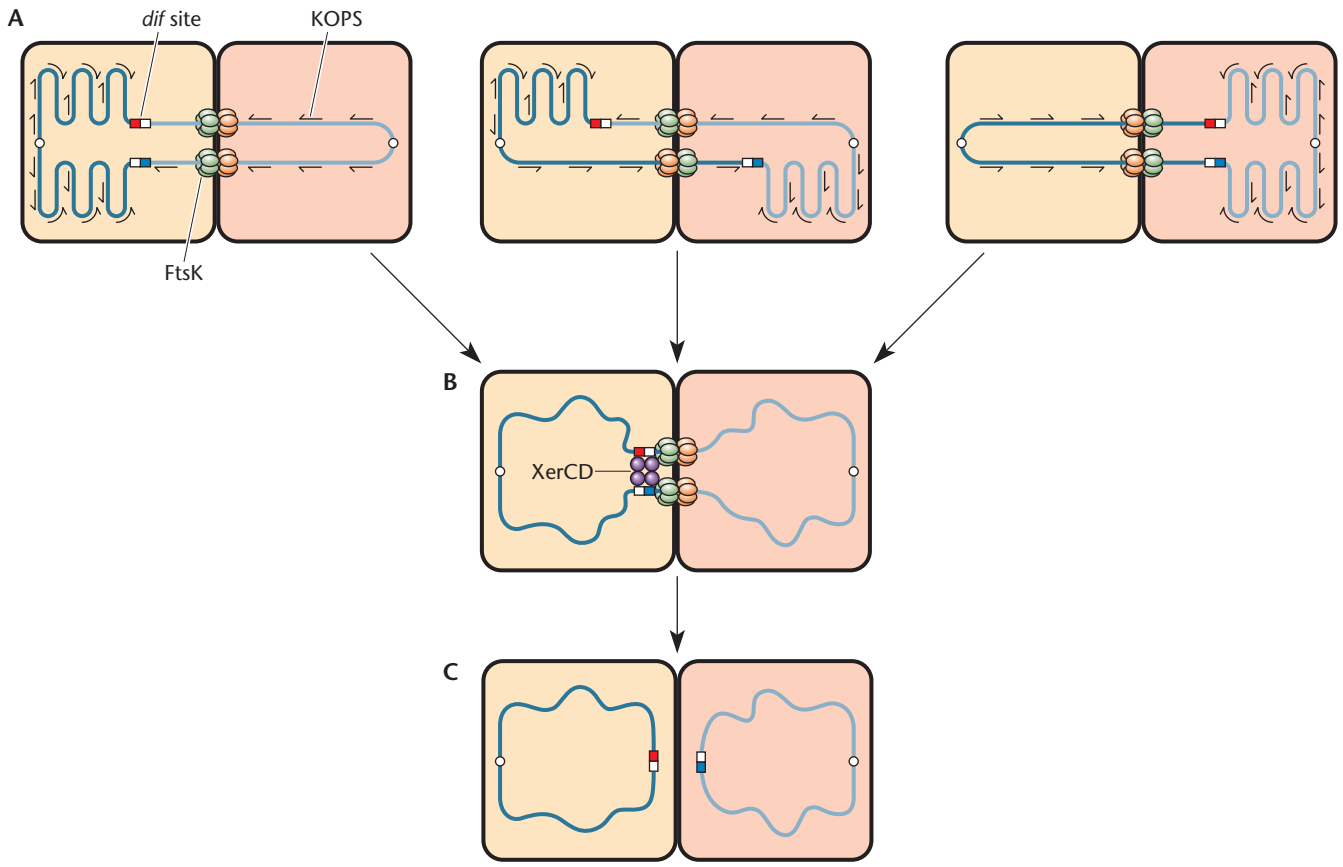


Figure 1.17 Model of the way in which chromosome translocation by FtsK coordinates chromosome segregation with dimer resolution. The two daughter chromosomes in a dimerized chromosome are shown in different shades of blue for emphasis. **(A)** Three possible distributions of the newly replicated dimer chromosome are shown following the start of cell septation. FtsK is a DNA translocase that can pump the chromosomes to the correct daughter cells using polar sequences in the chromosome called KOPS sites (shown as half arrows) while also moving the *dif* sites into alignment at the septum. Green indicates FtsK complexes that are actively pumping; orange indicates idle FtsK complexes where activity is set by the orientation of KOPS sites. **(B)** After aligning the *dif* sites, the FtsK protein also interacts with the XerCD enzyme, allowing it to resolve the dimer chromosomes at the *dif* sites. **(C)** The coordinated activities of the dimer resolution system and FtsK lead to monomer chromosomes that are capable of full segregation to daughter cells. From Camara JE, Crooke E, *in* Higgins NP (ed), *The Bacterial Chromosome* (ASM Press, Washington, DC, 2005).

that manage and repair the chromosome (Box 1.1). FtsK also has a domain that interacts with topoisomerase IV, a protein capable of untangling catenanes in the chromosomes (see below). Therefore, the FtsK protein coordinates a veritable clearinghouse of activities that help with chromosome management during chromosome replication and chromosome segregation.

Homologs of FtsK are widespread, and a homologous protein called SftA appears to carry out similar functions in *B. subtilis* (see Biller and Burkholder, Suggested Reading). Some bacteria have more than one FtsK-like protein, presumably for other specialized tasks. In the case of *B. subtilis*, another FtsK homolog, SpoIIIE, is responsible for translocating the final third of the chromosome into spores during spore development so the spore will get a complete copy of the chromosome (see chapter 12). How the roles of FtsK, SpoIIIE, and SftA differ and how the multiple members are involved in various DNA processing events in a single cell remain active areas of research with many questions still to be answered.

Interestingly, instead of XerC and XerD, *Streptococcus* and *Lactococcus* species have a system more closely related to bacteriophage integration systems that uses a single protein called XerS to carry out the same function (see Le Bourgeois et al., Suggested Reading). Archaea also seem to use a single protein (in this case called XerA) to resolve dimer chromosome at a *cis*-acting *dif* site (see Cortez et al., Suggested Reading). The regulation of dimer chromosome resolution in these systems and any involvement of FtsK-like proteins are unclear.

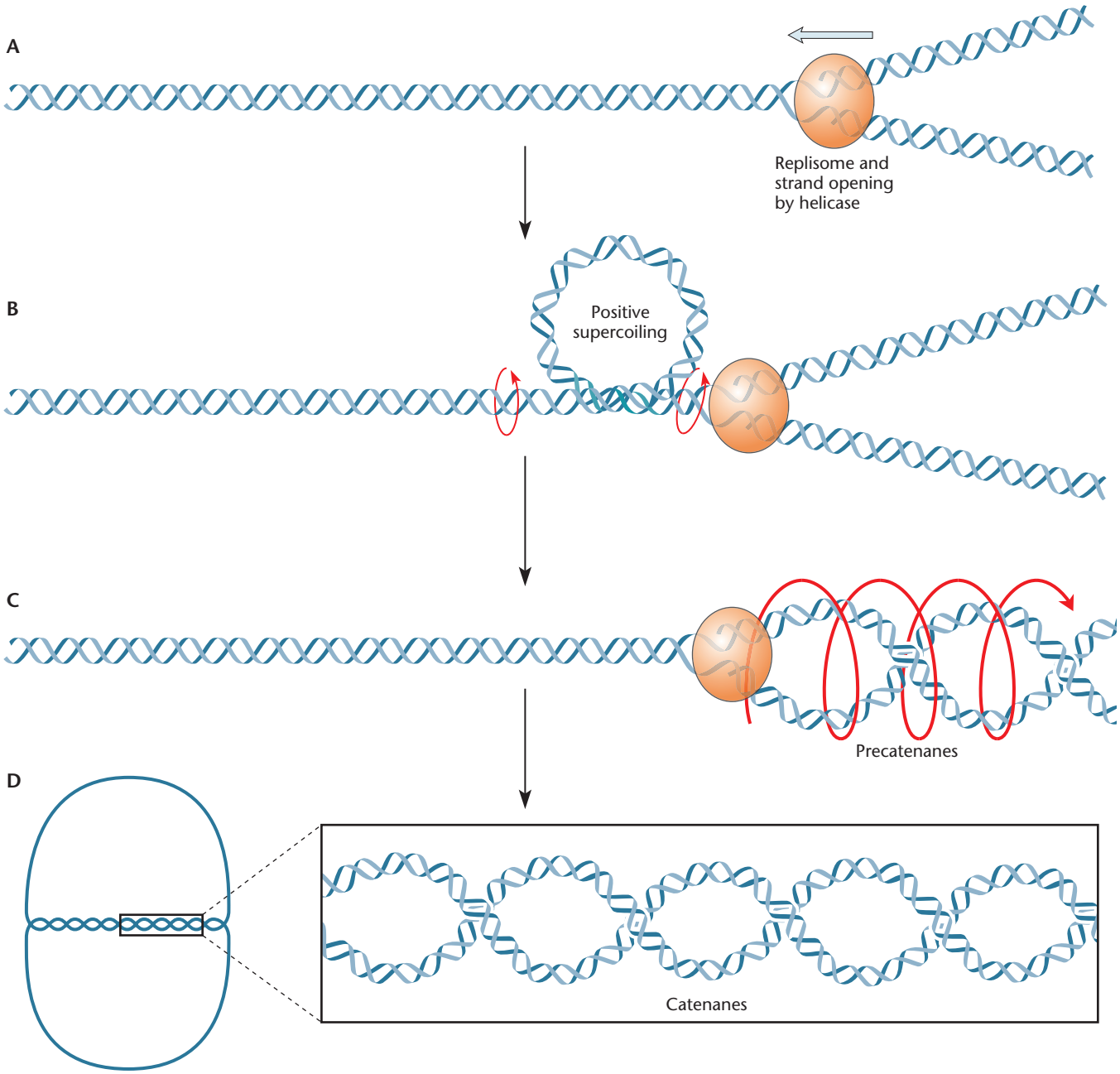


Figure 1.18 Model of the way in which unwinding of the template DNA strands can cause twists that can diffuse across the replication complex and twist the new DNA strands. **(A)** The replication machinery must open the double-stranded DNA to copy the template strands. **(B)** Unwinding the template DNA strands introduces twists (shown by thin red arrows) called positive supercoils ahead of the replication fork. **(C)** Some of the torsion that is generated ahead of the replication fork can be relieved by rotation of the replication complex itself. The torsional stress can spread behind the fork and intertwine the new copies of the chromosome (shown by a thick red arrow). The intertwined chromosomes are called precatenanes. **(D)** Precatenanes result in links in the daughter chromosomes called catenanes that must be unlinked for the chromosomes to separate into daughter cells.

DECATENATION

DNAs also become joined to each other through the formation of **catenanes**, where the daughter DNAs become interlinked like the links on a chain. These interlinks can form as a side effect of separating DNA strands during the process of DNA replication (Figure 1.18). As we discuss in "Supercoiling" below, DNA replication introduces a great deal of torsional stress ahead of the DNA replication fork (Figure 1.18A and B). This stress can be transferred across the DNA replication fork into the newly

formed DNA strands in twists between the two new daughter strands called precatenanes (Figure 1.18C), because the twists will eventually form catenanes when replication is complete (Figure 1.18D). Once such interlinks are formed, the only way to unlink them is to break both strands of one of the two DNAs and pass the two strands of the other DNA through the break. The break must then be resealed. This double-strand passage, called **decatenation**, is one of the reactions performed by type II topoisomerases (see below). A type II topoisomerase called **topoisomerase IV** (Topo IV) plays a major role in removing most of the interlinks between the daughter DNAs in *E. coli*. The act of removing these links appears to remove one of the major cohesive forces between the daughter chromosomes prior to segregation. One of the major points of regulation appears to be spatial, where the two subunits that make Topo IV are most likely to interact through association with other proteins. In one case, this occurs following replication, when the chromosomes are being translocated across the division septum by FtsK (Figure 1.19). FtsK helps to regulate the decatenation process as it pulls the chromosome to the septum, because one of the subunits of Topo IV interacts with FtsK. Interaction between Topo IV and the FtsK protein puts the enzyme in a very appropriate position for the removal of catenanes just before chromosome segregation. Topo IV is also regulated through an interaction with a protein involved in the condensation of chromosomes following DNA replication (see "Condensation" below).

CONDENSATION

Bacterial cells have an important mechanism to help manage chromosomes, which is to condense them after DNA replication. If the daughter chromosomes are condensed, they do not overlap as much in the cell and so are less apt to become interlinked. Condensation of chromosomes prior to mitosis has been known for a long time to occur in eukaryotic cells, where the chromosomes are only clearly visible just before mitosis. We now know that bacteria also condense their daughter DNAs to make them easier to manage, even though it is more difficult to visualize the condensation of bacterial chromosomes because of their smaller size.

Condensins

Proteins called **condensins,** which help to condense DNA in the cell, were first discovered in eukaryotes. They are also known as **SMC proteins** for structural maintenance of chromosome. Condensins are long, dumbbell-shaped proteins with globular domains at the ends and a long coiled-coil region holding them together. The long coiled-coil region has a hinge so that it can fold back on itself and the two globular domains can bind together. Condensins work with partner proteins to help condense

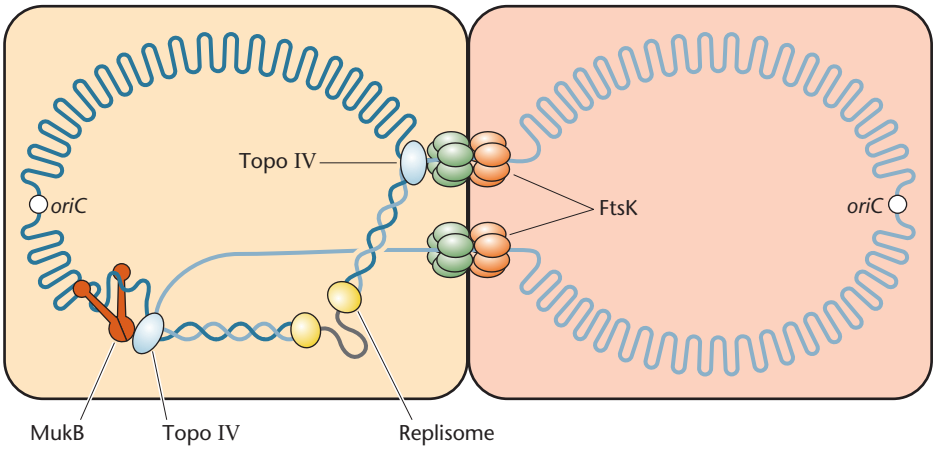


Figure 1.19 Model of the way in which chromosome decatenation by topoisomerase IV (Topo IV) is coordinated with chromosome condensation by MukB and chromosome translocation with FtsK. The two daughter chromosomes that have been replicated are indicated by light and dark blue lines, and the unreplicated portion of the chromosome is shown in gray. Unwinding of the template DNA strands that is associated with DNA replication twists the newly replicated strands of DNA, forming precatenanes that go on to become catenanes if not unlinked by the action of Topo IV. Topo IV can interact with the chromosome condensation protein MukB to remove catenanes before DNA is condensed. Topo IV also interacts with the FtsK translocase to coordinate decatenation with chromosome segregation. The replisome is shown as a yellow circle, and double-stranded DNA is shown as a single line for simplicity.

DNA into a higher-order structure. While still somewhat controversial, bacterial condensins may encircle DNA to gather regions on the DNA together that can subsequently be held together by a variety of other sequence-nonspecific DNA-binding proteins discussed elsewhere, such as HU, IHF, H-NS, and Fis.

The condensin of *E. coli*, called MukB, was found because mutations in its gene interfere with chromosome segregation. MukB was suspected of condensing the DNA because the protein and supercoiling of the DNA can compensate for each other in allowing proper segregation of the daughter chromosomes into daughter cells. As indicated above, supercoiling can lead to the formation of precatenanes and catenanes that are removed by Topo IV. The removal of precatenanes and catenanes also appears to be regulated with the condensation of chromosomes by an interaction between one of the Topo IV subunits and MukB (Figure 1.19) (see Hayama and Marians, Suggested Reading). MukB interacts with DNA through association with the partner proteins MukF and MukE. *B. subtilis* also has a condensin, which is more similar in amino acid sequence to the eukaryotic condensins and so was also named SMC protein (see Britton et al., Suggested Reading). It also has partner proteins named ScpA and ScpB. In *B. subtilis*, the link between chromosome condensation and partitioning is becoming clearer with the finding that proteins that directly recognize the region around the origin and are involved in partitioning are able to recruit the condensin SMC in this organism (see Thanbichler, Suggested Reading).

Supercoiling

Another way bacteria condense DNAs is through supercoiling. In bacteria, all DNAs are negatively supercoiled, which means that DNA is twisted in the opposite direction to the Watson-Crick helix, creating underwinds. As discussed in more detail below, the underwinds introduce stress into the DNA, causing it to wrap up on itself, much like a rope wraps up on itself if the two ends are rotated in opposite directions. This twisting occurs in loops in the DNA, causing the DNA to be condensed into a smaller space.

KEEPING NEW SISTER CHROMOSOMES SEPARATE INVOLVES COORDINATING MULTIPLE PROCESSES

Many processes ensure that the chromosomes are physically separate, yet are gathered together in a manageable way in the bacterial cell. With circular chromosomes, an uneven number of recombination events will regularly form dimer chromosomes, where the two circles join to form one large circle that cannot be subdivided into daughter cells. The highly controlled process of dimer resolution with XerC, XerD, and FtsK in *E. coli* ensures that dimer chromosomes are kept separate without accidently making dimers out of separate chromosomes before they can be passed to the daughter cells. Circular chromosomes can also become conjoined like the rings on a chain that cannot be passed on to daughter cells. Another highly regulated process by Topo IV removes these links between the chromosomes. Regulation of Topo IV activity with condensation and FtsK transport, as shown in Figure 1.19, plays an important role in making sure the enzyme separates interlinked chromosomes and does not link them together. A number of processes compact the chromosomes in bacteria. Condensins gather regions of the chromosome together that are, in turn, held together by a number of nonspecific DNA-binding proteins to allow compaction of the bacterial chromosome. This process is facilitated by supercoiling that allows the chromosome to twist around itself. In the next section, we learn how chromosomes are efficiently partitioned into each daughter cell during the process of cell division.

CHROMOSOME PARTITIONING

Not only must the two daughter chromosomes be segregated after replication, they also must be segregated in such a way that each daughter cell gets only one of the two copies of the chromosome. Otherwise, one daughter cell would get two chromosomes and the other would be left with no chromosome and eventually would die. The apportionment of one daughter chromosome to each of the two daughter cells is called **partitioning**. Daughter cells that lack a chromosome after division are very rare, indicating that partitioning is a very efficient process in bacteria. Because of the importance of chromosome segregation, redundant mechanisms may have evolved to ensure that it occurs accurately. Indeed, many of the mechanisms that allow condensation of chromosomes can contribute to partitioning once the origin regions are located in the nascent daughter cells. While broad themes that describe partitioning across all bacteria have eluded our understanding, some important model systems are fairly well understood. In this section, we discuss what is known in the model bacteria.

The Par Proteins

Early work concentrated on the functions of the so-called partitioning proteins, the products of the *par* genes. The Par functions were first discovered in plasmids, which are small DNA molecules that are found in bacterial cells and that replicate independently of the chromosome (see chapter 4). Because they exist independently of the chromosome, plasmids usually must also have a system for partitioning; otherwise, they would often be lost from cells during the process of division. The Par systems of plasmids are known to fall into two families, one represented

by the Par system of plasmid R1 and the other, much larger family represented by plasmids P1, F, and many others. It is the second of these families to which the known Par functions of chromosomes belong. The molecular details of Par systems are addressed in chapter 4, but some of the basics are described here as they apply to chromosome segregation systems.

The region of DNA that is to be partitioned, whether it is the origin region of the chromosome or a plasmid, contains a series of *cis*-acting sites called *parS* sites (Figure 1.20). One of two proteins in this system, ParB, binds to the *parS* sites. Partitioning of the ParB-bound *parS* DNA occurs because it is attracted to the other protein component of the system, active ParA*, which binds DNA. ParA* binds DNA nonspecifically and basically coats the entire nucleoid. The system works because ParB-*parS* is attracted to DNA-bound ParA*, but upon interacting with ParA* it helps to convert the ParA* into inactive ParA that does not bind DNA and is therefore released from the chromosome. As ParB-*parS* follows the gradient of ParA* across the nucleoid, it keeps the two *oriC* regions separate from one another. ParA eventually converts back to its active ParA* form and binds elsewhere on the nucleoid.

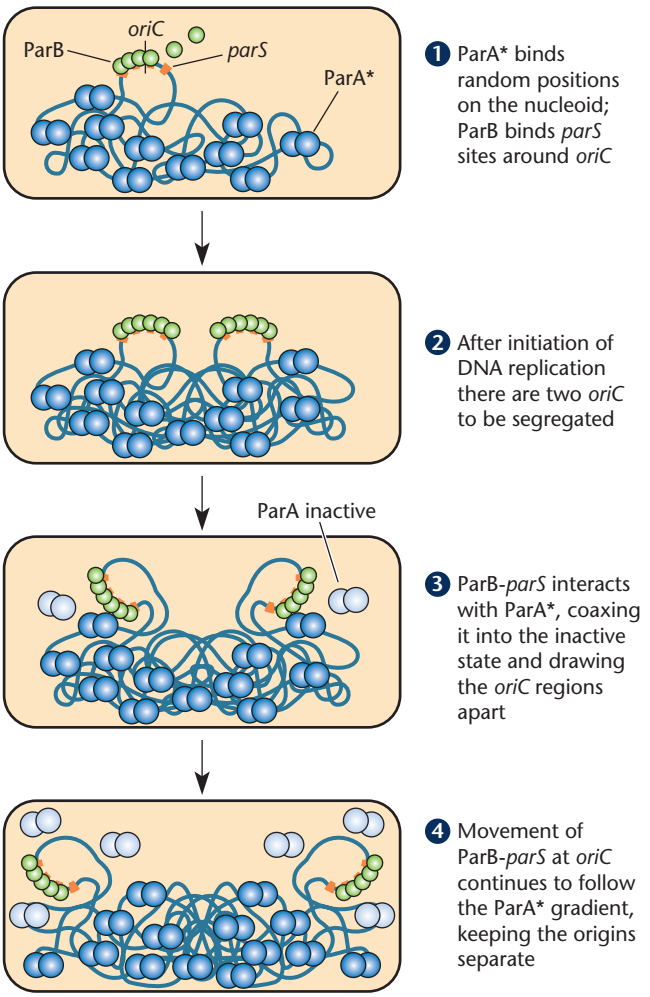


Figure 1.20 Model of how an origin region containing *parS* sites bound by the ParB protein is segregated by its attraction to DNA-bound ParA proteins (ParA*). Inactivation and displacement of ParA by the ParB-*parS* complex provide a mechanism to separate the origin region-containing ParB-*parS* complexes in the dividing cell.

Par functions in *B. subtilis*. The ParAB/*parS* system in *B. subtilis* provides some insight into how partitioning of *oriC* can work with condensation functions. In *B. subtilis*, the proteins analogous to the ParA and ParB proteins are called Soj and Spo0J, respectively. These names come from early genetic studies of *B. subtilis* sporulation, where *spo0J* was identified as a gene required for sporulation and *soj* was identified as a suppressor of *spo0J*. The *parS* sites close to the origin of chromosome replication (see Box 1.1) are bound by Spo0J(ParB), and these are pulled apart following the gradient to active Soj(ParA*), similar to what is shown in Figure 1.20. In *B. subtilis*, Spo0J(ParB) is capable of recruiting condensin protein SMC. Recruitment of SMC may play a critical role in organizing the chromosomes to help extrude the origin regions as they are gathered together with the SMC protein and other nonspecific DNA-binding proteins.

Macrodomains

Surprisingly, the chromosomes of *E. coli* and other *Enterobacteriaceae* do not have a recognizable Par-like system, despite the fact that plasmids with ParAB/*parS* systems are common in *Enterobacteriaceae*, and it seems likely that another system is responsible for the active partitioning of the chromosomes. Work in *E. coli* suggests that an incompletely understood system that hinges on large independently organizing regions called **macrodomains** functions in *Enterobacteriaceae*. Macrodomains (MD) were first established as large regions in the *E. coli* chromosome where recombination between two sites within a given macrodomain region occurs at a much greater frequency than outside the macrodomain region (see Valens et al. 2004, Suggested Reading). These regions localize to a specific subregion within the larger nucleoid and show individualized segregation properties. Four regions were identified as having these properties, one encompassing the origin region (Ori-MD), one encompassing the terminus region (Ter-MD), and two regions to the right (Right-MD) and left (Left-MD) of the terminus region (Figure 1.21). Two regions surrounding the origin macrodomain appear to not be as structured. Extensive molecular details exist for the Ter macrodomain, but little is known about the other macrodomain regions.

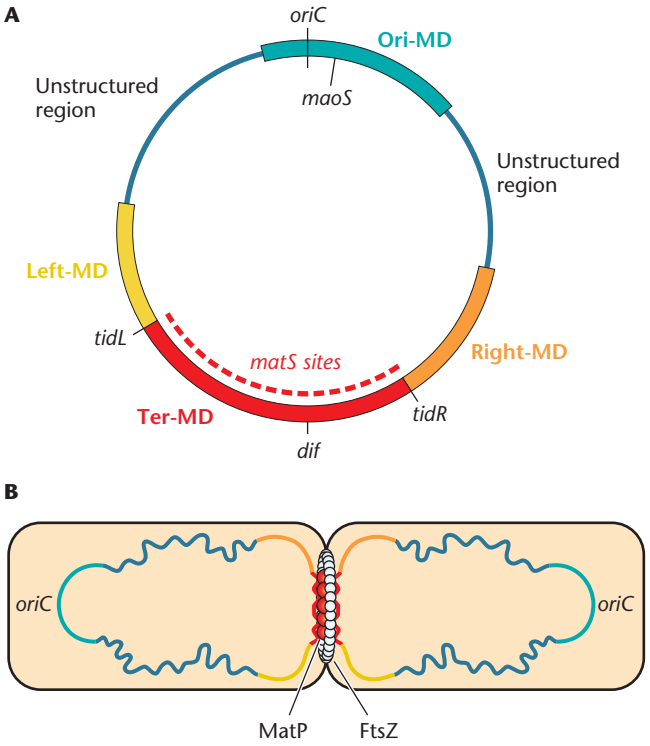


Figure 1.21 The *E. coli* chromosome has four structured regions called macrodomains (MDs) that aid in segregation of the chromosome. **(A)** Approximate genetic positions of the Ori-MD, the Right- and Left-MD, the Ter-MD, and the two unstructured regions of the *E. coli* chromosome. The Ter-MD is the best understood and is organized by the MatP protein via an association with its *matS* sites across the terminus region of the chromosome. The *tidL* and *tidR* sites, which are recognized by the YfbV protein (not shown), constrain the MatP protein from extending outside this region. **(B)** MatP compacts and localizes the Ter-MD through an association with ZapB (not shown), which is localized around the FtsZ ring region. The Ori-MD is organized via an association of the MaoP protein with a single site near the origin, *maoS*, via an unknown mechanism. The molecular mechanism responsible for organizing the Right- and Left-MDs and any associated proteins is unknown.

The Ter macrodomain: MatP and *matS*. The molecular basis for the Ter macrodomain involves a series of 23 *matS* sites found across an 800-kb region that are recognized by the protein MatP (see Mercier et al., Suggested Reading). MatP-*matS* complexes compact this region where they also associate with a specific set of proteins localized at the center, facilitating the process of orderly cell division (see below) (Figure 1.21B). The structure formed with MatP-*matS* appears to be constrained from spreading by two sites (*tidL* and *tidR*) which interact with a partner protein suggested to associate with the cell membrane (Figure 1.21A). Presumably, the MatP-*matS* system would coordinate the segregation of the terminus region as a late step in chromosome segregation. The molecular basis for the left and right macrodomains remains unknown.

The origin macrodomain: *maoS* and MaoP. Recombination studies were used to establish two important players in the origin macrodomain (see Valens et al. 2016, Suggested Reading). The explanation for the origin macrodomain is less clear, but appears to involve a single *cis*-acting site called *maoS*, found about 22 kb away from *oriC*, that is recognized by the MaoP protein (Figure 1.20). It remains unclear how a single site allows the formation of this large macrodomain or how it would function. The MaoP-*maoS* system is functionally distinct from the ParAB-*parS* partitioning systems found in plasmids and most other bacterial chromosomes, and placing the MaoP-*maoS* system onto a plasmid does not recapitulate the segregation found with the origin region, suggesting that important pieces are still missing from this story. A 25-bp site called *migS* was identified for its role in orienting the origin region within the larger nucleoid, but any role this site plays in segregation or any *trans*-acting factor that works with this site has yet to be identified (see Yamaichi and Niki, Suggested Reading).

Interestingly, a number of systems found in *E. coli* and other enteric bacteria appear to have coevolved: MatP/*matS*, MaoP/*maoS*, MukBEF, SeqA, and Dam methyltransferase along with a number of other proteins are only found in this subgroup of bacteria (see Brézellec et al., Suggested Reading). It is unclear how these systems functionally replace the ParAB-*parS* system found in most bacteria, and it will be interesting to discover how the systems relate and function in *Enterobacteriaceae*.

Coordinating Cell Division and Chromosome Partitioning in *E. coli* and *B. subtilis*

Much has also been learned about how the bacterial division septum forms. This process is called **cytokinesis**. A protein called FtsZ, which forms a ring around the midpoint of the cell, performs the primary step in this process (Figure 1.21b). This protein is related to tubulin of eukaryotes and forms filaments that grow and shorten by adding and removing shorter filaments, called protofilaments, to its ends in the presence of GTP. Before the cell is ready to divide, the FtsZ protein exists as helical filaments that traverse the cell. When the cell is about to divide, these filaments converge on the middle of the cell and form a ring at the site of the future septum. The FtsZ ring then attracts many other proteins, including the DNA translocase FtsK discussed above. FtsZ helps form the division septum, which eventually squeezes the mother cell at its center to allow the formation of the two daughter cells. The following major questions may be asked: why does the septum form only in the middle of the cell, and why does septum formation not occur over the nucleoid prior to chromosome segregation? The answers to these questions lie, at least in part, in two types of systems: the Min systems and the nucleoid occlusion systems.

The Min Proteins

In *E. coli*, three proteins called MinC, MinD, and MinE are known to be involved in localizing the division septum at the center of the cell. The *min* genes of *E. coli* were found because mutations in these genes can cause division septa to form in the wrong places, sometimes pinching off smaller cells called minicells. Apparently, in the absence of the Min proteins, division septa can form in places other than the middle of the cell. When this happens, smaller minicells that lack a chromosome are pinched off, hence the name Min proteins, for minicell-producing. It was predicted that the Min proteins would be localized in the ends of the *E. coli* cell, where they could prevent FtsZ from forming a division septum anywhere but the middle of the cell. However, when the localization of the Min proteins was studied using GFP fusions to the Min proteins, a very surprising result was revealed: the Min proteins oscillate from one pole of the cell to the other during the cell cycle. A model used to account for this finding held that oscillations of MinD and MinE drive the oscillation of MinC, which interacts with MinD (see MinCD in Figure 1.22) and is ultimately responsible for preventing FtsZ ring formation at the cell poles and enforcing the formation of a single FtsZ ring at mid-cell. The molecular mechanism that drives the redistribution of the proteins within the cell stems from the interaction of MinD (an ATPase) and MinE, which stimulates ATPase activity in MinD. MinD interacts with the membrane only in the ATP-bound state. More recent work with the system suggests that changes in the nature of MinD and MinE membrane-bound complexes and the states found in the cytoplasm are important for setting a distribution of these proteins (see Vecchiarelli et al., Suggested Reading). Ultimately, the concentration gradient set by the dynamic behavior of MinE and MinD sets a low concentration of MinCD at the center of the cell, allowing FtsZ to form a ring at the center of the cell.

Regulation of septum formation in *B. subtilis* differs from that found in *E. coli*. In *B. subtilis*, MinE is lacking and MinC and MinD do not oscillate. Instead, MinCD appears to tether directly to the cell poles by binding to another protein at the cell poles, called DivIVA. This binding creates a gradient of concentration of MinCD in the cell and similarly only allows formation of a single FtsZ ring at the center of the cell. Therefore, these two model bacteria use somewhat different mechanisms to establish a gradient of MinCD concentration and thereby restrict FtsZ ring formation to the center of the cell (Figure 1.22).

Nucleoid Occlusion

Not only should the FtsZ ring form only in the center of the cell, it also should not initiate the assembly of a division septum while the nucleoid is still occupying the center of the cell, or it might constrict the membrane around the chromosome. In fact, it was observed in *E. coli* that FtsZ rings never formed in the center of the cell when it was still occupied by the nucleoid, which had not yet segregated. Proteins that inhibit FtsZ ring formation in the presence of the nucleoid were discovered in both *E. coli* and *B. subtilis* at about the same time and were named **nucleoid occlusion proteins**. Both proteins were found because they are essential only if the Min system is inactivated. There is evidence that another process drives formation of the FtsZ at the midcell and that having both Min and nucleoid occlusion systems serves to increase the accuracy of the division process. The nucleoid exclusion protein in *B. subtilis*, named Noc, was found serendipitously, because its gene, *noc*, is adjacent to the genes for the Par functions, *soj* and *spo0J*, and it was observed that mutations in this gene could not be combined with mutations in the *minD* gene without making the cells very sick. The reason *noc* mutant strains were sick was that

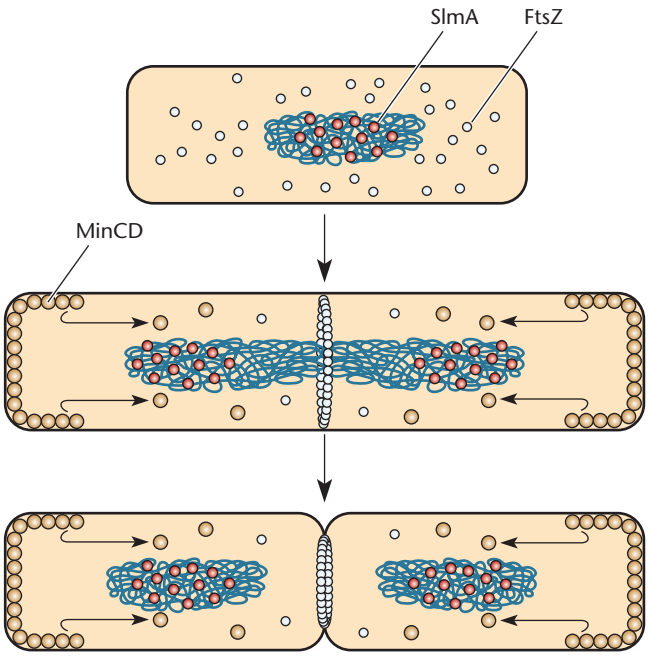


Figure 1.22 The MinCDE and nucleoid occlusion systems control placement of the FtsZ ring in *E. coli*. The FtsZ ring is an important marker of the central division site in a soon-to-divide cell which should not assemble until the chromosome is ready to segregate. MinC, which interacts with MinD, prevents a FtsZ ring from forming anywhere but at the center of the cell. Different mechanisms are used in bacteria to limit the concentration of MinCD at the center of the cell, as described in the text, to prevent FtsZ ring formation outside the center of the elongating cell. The nucleoid occlusion protein of *E. coli*, SlmA, is an important system that ensures that the FtsZ ring does not form over the nucleoid until the chromosomes are nearly completely replicated and segregated into daughter cells. Because the SlmA-binding sites are absent from the terminus region, the FtsZ ring can start to form once chromosomal replication is nearly complete and the origin region has progressed into the daughter cells away from the division site. See the text for details.

they were forming long filaments of cells because they were not dividing properly. The nucleoid occlusion protein in *E. coli*, named SlmA, was found directly by a synthetic lethal screen (see Bernhardt and de Boer, Suggested Reading). A synthetic lethal screen is a powerful genetic tool in which mutations in genes whose products are required only if another gene product is absent are isolated. The investigators looked for mutations in genes that were required only in the absence of the products of the *min* genes. They expressed the *min* genes from an inducible promoter and looked for mutants that were sick and failed to form colonies only in the absence of inducer. Some of these mutants had mutations in a gene that was named *slmA* by the investigators. While mutants deficient in Min proteins had more Z rings, they were never over the nucleoids (Figure 1.22). However, mutants that lacked both the Min proteins and SlmA often formed FtsZ rings over the nucleoids, as expected for a mutant deficient in nucleoid occlusion. The use of inducible promoters and other examples of synthetic phenotypes are discussed in more detail in later chapters.

The Noc and SlmA proteins seem to act by binding to DNA and then inhibiting FtsZ ring formation close to the DNA to which they are bound. There are known to be DNA sequences called Noc-binding sites (NBS) that are bound by Noc, and SlmA-binding sites (SBS) that are bound by SlmA; these sites are distributed across the chromosome and concentrated in the origin region but are absent from the terminus region (Box 1.1). This allows SlmA and Noc to help protect the nucleoid from the division septum until the final moments prior to the completion of DNA replication (see Wu et al., Suggested Reading) (Figure 1.22).

Note that this entire section has focused on simple division in rod-shaped cells. Fascinating adaptations to these simple ideas are known to occur in systems where the cells resulting from division are morphologically distinct. A particularly well-studied system in *Caulobacter crescentus* shows many adaptations when a mother cell gives rise to smaller motile daughter cells. Additionally, round or coccoid cells have distinct mechanisms that allow them to divide with clear cell poles as are found in rod-shaped cells. Given the extreme morphological variation known to occur across bacteria (see Kysela et al., Suggested Reading), interesting adaptations for division likely await discovery.

Coordination of Cell Division with Replication of the Chromosome

It is not sufficient to know how chromosomes replicate and then segregate into the daughter cells prior to division. Something must coordinate the replication of the chromosome with division of the cells. If the cells divided before the replication of the chromosome was completed, there would not be two complete chromosomes to segregate into the daughter cells, and one cell would end up without a complete chromosome. The mechanism by which cell division is coordinated with replication of the DNA is still not completely understood, but there is a lot of relevant information.

TIMING OF REPLICATION IN THE CELL CYCLE

It is important to know when replication occurs during the cell cycle. Experiments were designed to determine the relationship between the time of chromosome replication and the cell cycle in *E. coli* (see Helmstetter and Cooper, Suggested Reading). The conclusions are still generally accepted, so it is worth going over them in some detail. The investigators recognized that if the DNA content of cells at different stages in the cell cycle could be measured, it would be possible to determine how far chromosome replication had proceeded at that time in the cell cycle. Since bacterial cells are too small to allow measurements of DNA content in a single cell by the methods they had, it was necessary to measure the DNA content in a large number of pooled cells. However, cells growing in culture are all at different stages in their cell cycles. Therefore, to know how far replication had proceeded at a certain stage in the cell cycle, it was necessary to synchronize the cells in the population so that all were the same age or at the same point in their life cycles at the same time.

Helmstetter and Cooper accomplished this by using what they called a bacterial "baby machine." Their idea was to first label the DNA of a growing culture of bacterial cells by adding radioactively labeled nucleosides and then fix the bacterial cells on a membrane. When the cells on the filter divided, one of the two daughter cells would no longer be attached and would be released into the medium. All of the daughter cells released at a given time would be newborns and so would be the same age. This means that cells that divided to release the daughter cells at a given time would also be the same age and would have DNA in the same replication state. The amount of radioactivity in the released cells was then a measure of how much of the chromosome had replicated in cells of that age. This experiment was done under different growth conditions to show how the timing of replication and the timing of cell division are coordinated under different growth conditions.

Figure 1.23 shows the results of these experiments. For convenience, the following letters were assigned to each of the intervals during the cell cycle. The letter I denotes the time from when the last round of chromosome replication initiated until a new round begins. C is the time it takes to replicate the entire chromosome, and D is the time from when a round of chromosome replication is completed until cell division occurs. The top of the figure shows the relationship of I, C, and D when cells are growing slowly, with a generation time of 70 min. Under these conditions, I is 70 min, C is 40 min, and D is 20 min.

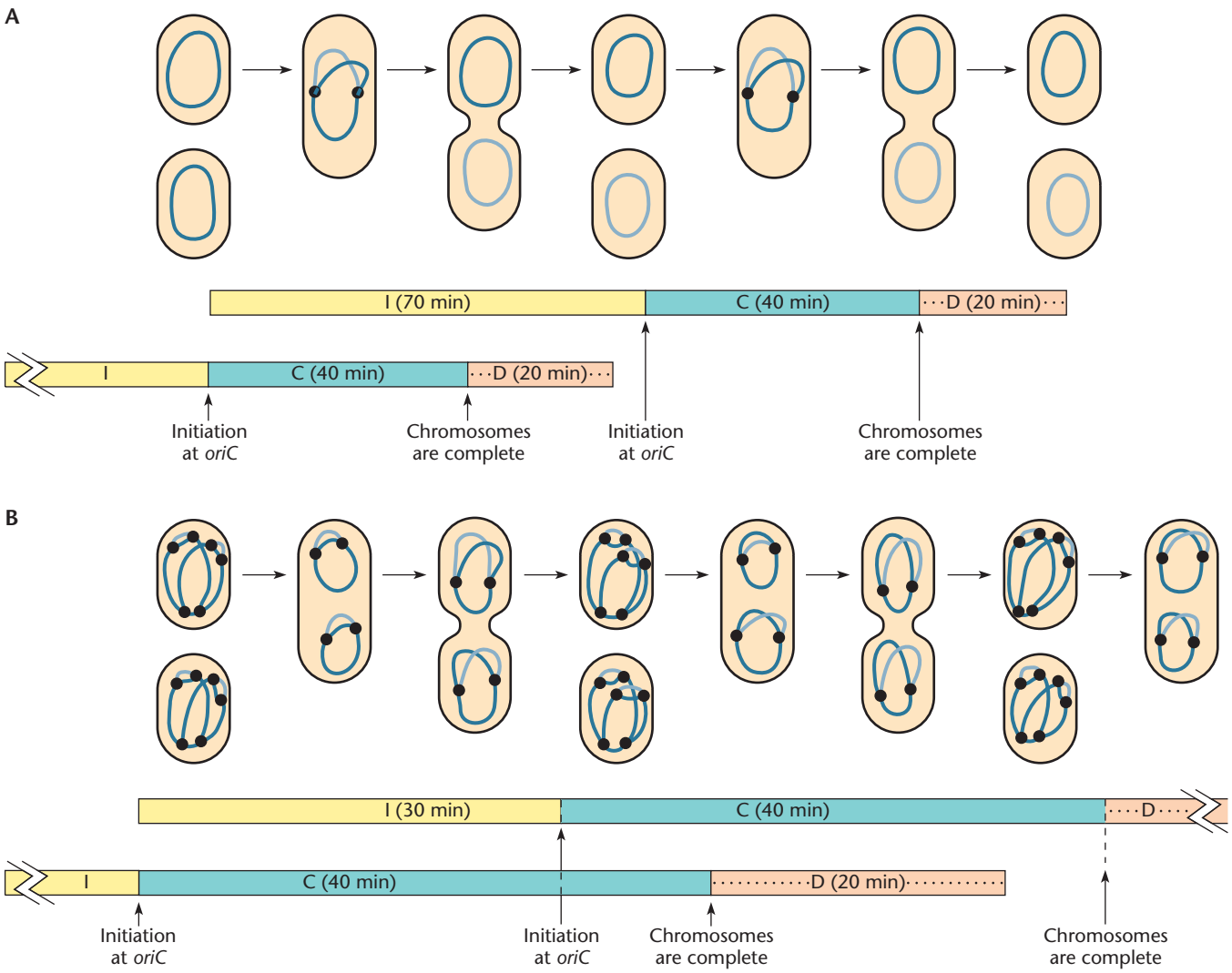


Figure 1.23 Timing of DNA replication during the cell cycle, with two different generation times **(A and B)**. The time between initiations (I) is the only time that changes. See the text for definitions of I, C, and D.

However, when the cells are growing in a richer medium and are dividing more rapidly, with a generation time of only 30 min, the pattern changes. The C and D intervals remain about the same, but the I interval is much shorter, only about 30 min.

Some conclusions may be drawn from these data. One conclusion is that the C and D intervals remain about the same independent of the growth rate. At 37°C, the time it takes the chromosome to replicate is always about 40 min, and it takes about 20 min from the time a round of replication terminates until the cell divides. However, the I interval gets shorter when the cells are growing faster and have shorter generation times. In fact, the I interval is approximately equal to the generation time—the time it takes a newborn cell to grow and divide. This makes sense because, as discussed below, initiation of chromosome replication appears to occur every time the cells reach a certain size. They reach this size once every generation time, independent of how fast they are growing.

Another point apparent from the data is that in cells growing rapidly with a short generation time, the I interval can be shorter than the C interval. If I is shorter than C, a new round of chromosomal DNA replication will begin before the old one is completed. This explains the higher DNA content of fast-growing cells as compared to slow-growing cells. It also explains the observation that genes closer to the origin of replication are present in more copies than are genes closer to the replication terminus.

Despite providing these important results, this elegant analysis does not allow us to tell whether division is coupled to initiation or termination of chromosomal DNA replication. The fact that the I interval always equals the generation time suggests that the events leading up to division are set in motion at the time a round of chromo-

some replication is initiated and are completed 60 min later independent of how fast the cells are growing. However, it is also possible that the act of termination of a round of chromosome replication sets in motion a cell division 20 min later. Multiple laboratories are working to resolve these issues.

Timing of Initiation of Replication

A new round of replication must be initiated each time the cell divides, or the amount of DNA in the cell would increase until the cells were stuffed full of it or decrease until no cell had a complete copy of the chromosome. Clearly, initiation of replication is exquisitely timed. In cells growing very rapidly, in which the next rounds of replication initiate before the last ones are completed, so that the cells contain a number of origins of replication, all of the origins in a cell "fire" simultaneously, indicating tight control.

A number of attempts have been made to correlate the timing of initiation of chromosome replication with other cellular parameters during the cell cycle. Most evidence from such attempts points to initiation of replication being tied to cell mass. After cells divide, their mass, or weight, continuously increases until they divide again. The initiation of chromosome replication occurs each time the cell achieves a certain mass, the initiation mass. If cells are growing faster in richer medium, they are larger and achieve the initiation mass sooner than do smaller, slower-growing cells, explaining why new rounds of chromosome replication occur before the termination of previous rounds in faster-growing cells but not in slower-growing cells. However, these experiments by themselves do not explain what it is about the cell mass that triggers initiation.

ROLE OF THE DnaA PROTEIN

DnaA is essential for initiating DNA replication at *oriC* and along with other components also helps to regulate the timing of DNA replication. DnaA primarily affects the frequency of initiation by changes in the ATP-binding versus ADP-binding state of DnaA and by access of DnaA to *oriC*. DnaA must be in its ATP-bound state to form the structure used to separate the DNA strands at *oriC* to initiate DNA replication. DnaA hydrolyzes ATP slowly; therefore, other factors are required to stimulate its ATPase activity to inactivate it until it is needed for the next round of DNA replication. The affinity of DnaA for the various binding sites within *oriC* (see "Origin of Chromosomal Replication" above) allows a mechanism for the origin to be identified but not immediately used as an origin until all conditions are in place for replication. A number of processes control the pool of DnaA found in the ATP-bound state. Reducing the pool of DnaA-ATP following initiation is an important mechanism for delaying reinitiation, thereby allowing replication to proceed in an orderly fashion. The production of new DnaA also plays a role in the ratio of ATP-bound versus ADP-bound DnaA, and because of the high ATP concentration in cells, newly made DnaA will be bound by ATP.

Inactivation of DnaA by hydrolysis to DnaA-ADP

Conversion to the DnaA-ADP state occurs primarily by two mechanisms following initiation of DNA replication. As mentioned above, only the ATP-bound form can bind all of the DnaA boxes in *oriC*. Only when all of the sites are bound can the structure form that will be needed to open the DUE to allow replication to initiate (Figure 1.15). This helps the DnaA protein to act as a switch that is largely independent of its concentration because other cellular inputs can control the ATP-bound versus ADP-bound state of DnaA. One cellular input that reduces the pool in the DnaA-ATP state involves the presence of the assembled replication fork at the origin region immediately after initiation of replication. The presence of the β sliding-clamp protein causes DnaA to hydrolyze ATP to ADP by interacting with another protein, a relative of DnaA called Hda (for homology to DnaA) (see Camara et al., Suggested Reading). This process is sometimes referred to as regulatory inactivation of DnaA, or RIDA. A second input that reduces the pool of DnaA-ATP involves a locus in the chromosome called *datA*. The *datA* locus, which is about 1 kb in length, has five DnaA-binding sites and a binding site for IHF. Binding of DnaA and IHF to this locus induces an endogenous ATPase activity of the DnaA protein, coaxing it into the inactive DnaA-ADP form in a process abbreviated as DDAH (*datA*-dependent DnaA-ATP hydrolysis). How specifically DDAH is controlled with the cell cycle remains unknown.

Reactivation of DnaA by nucleotide exchange to DnaA-ATP

Production of new DnaA in the ATP-bound form is one mechanism for increasing the DnaA-ATP/DnaA-ADP ratio in cells. There are also two mechanisms involved in "recycling" DnaA by encouraging it to exchange a nucleotide, putting it back into the DnaA-ATP form; one involves an exchange catalyzed by acidic phospholipids, and the other involves two sites found in the chromosome called DnaA-reactivating sequences (see Fujimitsu et al., Suggested Reading). The numerous inputs work together to limit the initiation of DNA replication to help keep the number of chromosomes consistent with the requirements for cell division.

SeqA-MEDIATED HEMIMETHYLATION AND SEQUESTRATION

In some types of bacteria, including *E. coli*, there is yet another means of delaying the initiation of new rounds of chromosome replication. As with the mechanisms described above, replication itself plays a role in this regulatory pathway where methylation of the DNA helps

delay initiation. In *E. coli* and other enteric bacteria, the A's in the symmetric sequence GATC/CTAG are methylated at the N6 position. These methyl groups are added to the bases by the enzyme deoxyadenosine methylase (Dam or Dam methylase), but this occurs only after the nucleotides have been incorporated into the DNA. Since DNA replicates by a semiconservative mechanism, the A in the GATC/CTAG sequence in the newly synthesized strand remains temporarily unmethylated after replication of a region containing this sequence (Figure 1.24). The DNA at this site is said to be hemimethylated if the bases on only one strand are methylated.

The hemimethylated state is important in the context of regulation of initiation because a *trans*-acting protein called SeqA only binds with high affinity to hemimethylated GATC/CTAG sequences. SeqA is an essential facilitator of hemimethylation control of replication initiation and is found only in bacteria that have Dam. The sequence GATC/CTAG is found 11 times within *oriC*, much more often than would be expected by chance alone (Figure 1.15). GATC/CTAG sequences are also associated with the low-affinity DnaA-binding sites across the origin. Furthermore, the promoter region of the *dnaA* gene, the region in which mRNA synthesis initiates for the DnaA protein, also has GATC/CTAG sequences. SeqA is able to bind all of these strategically located GATC/CTAG sites after they are replicated (and therefore rendered hemimethylated), which has the effect of delaying the conversion to full methylation at these sites for about one-third of a cell cycle. SeqA also blocks binding of DnaA to the low-affinity sites and inhibits expression of the *dnaA* gene. SeqA bound to hemimethylated DNA may associate with the cell membrane to sequester the *oriC* region after initiation of DNA replication (Figure 1.25). Sequestration of the hemimethylated *oriC* region is predicted to render it temporarily nonfunctional for the initiation of a new round of replication and delays its methylation by the Dam methylase (see Slater et al., Suggested Reading).

SeqA activities outside of *oriC*

In addition to the GATC/CTAG sequences associated with *oriC* and the *dnaA* promoter, SeqA also interacts with the GATC/CTAG sequences as the replication forks progress around the chromosome, effectively marking the location of the replisome (Figure 1.25). SeqA bound to transiently hemimethylated GATC/CTAG sites immediately behind the DNA replication forks may be capable of bringing together the nascent sister chromosomes, and it also appears to negatively regulate the decatenation activity of topoisomerase (see Joshi et al., Suggested Reading). Both of these processes may mediate a form of sister chromosome cohesion that in turn may help to regulate processing of the new DNAs by positioning them for DNA repair and recombination. Interestingly, SeqA is essential in mutants that lack the major pathways of DNA recombination involving RecA, supporting an additional role in protecting genome integrity. Additionally, chromosomes appear to be vulnerable to many types of mobile DNA elements during DNA replication, especially when

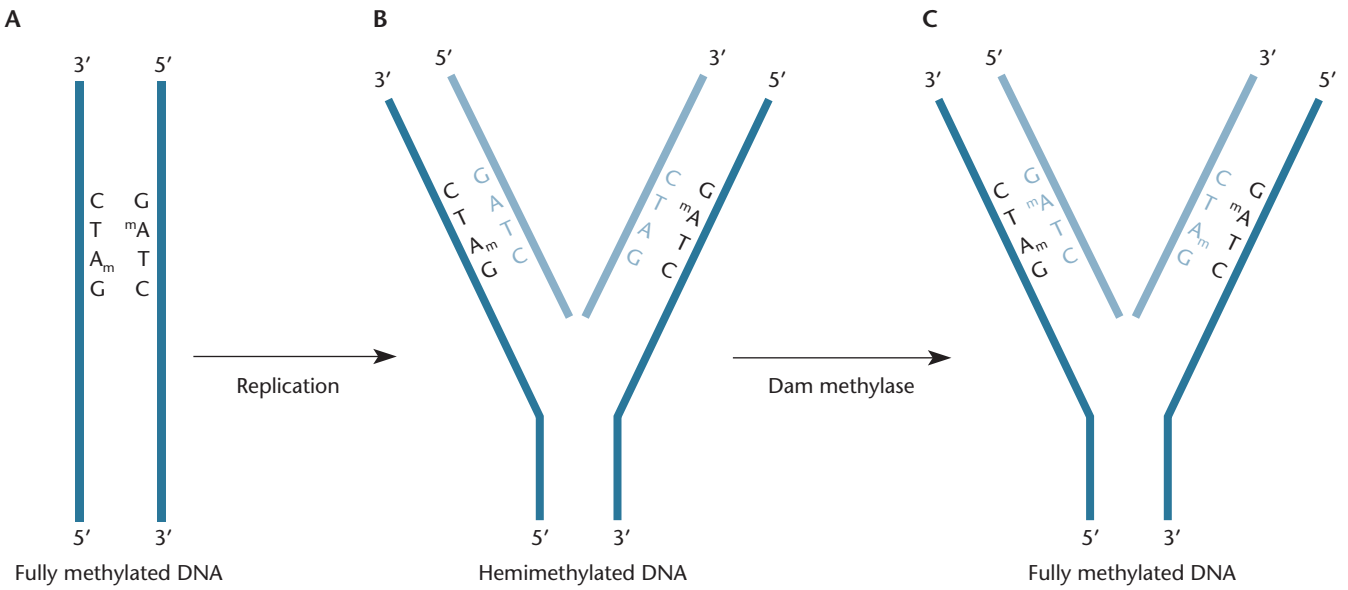


Figure 1.24 Replication creates hemimethylated DNA. **(A)** The A in the sequence GATC is methylated on both strands (A_m and ^mA). **(B)** After replication, the A in GATC in the new strand is not immediately methylated by the Dam methylase. **(C)** Eventually, GATC sites in the new strand are methylated, converting the DNA back to the fully methylated state.

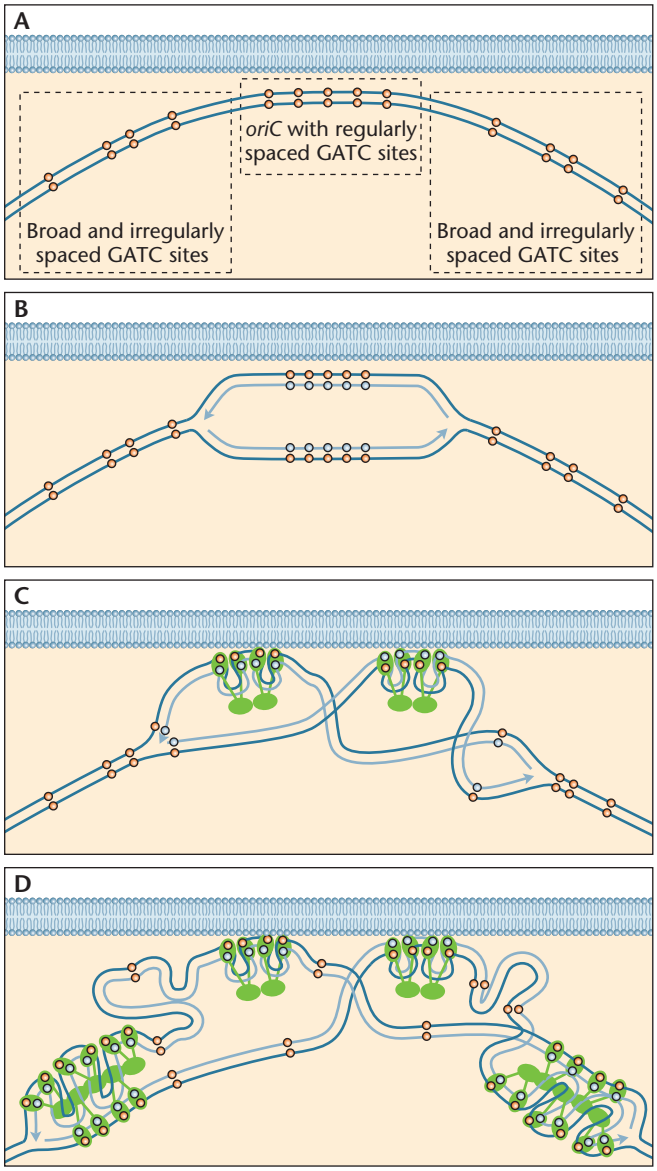


Figure 1.25 Model showing the possible functional consequences of SeqA binding with regularly and closely spaced GATC sites in the *oriC* region and broadly and irregularly spaced GATC sites outside the *oriC* region in the *E. coli* chromosome. Methylated GATC sites in a strand are represented by orange circles. Unmethylated GATC sites in a strand are represented by blue circles. **(A)** Before initiation, the *oriC* region is methylated in both strands. **(B)** After initiation, only one of the two strands is methylated; hence, the region is hemimethylated. The newly synthesized strand is shown in a lighter shade of blue. **(C)** SeqA binds the regularly spaced hemimethylated GATC sites in *oriC* and is able to interact with the membrane, thereby sequestering *oriC* and preventing further initiation and drastically slowing methylation. **(D)** Outside of the *oriC* region, where GATC sites are not regularly spaced and are situated farther apart, SeqA may be a bridge between the two hemimethylated strands to help coordinate processing events important for facilitating repair and recombination.

replication forks stall (see Fricker and Peters, Suggested Reading). SeqA-facilitated processes may help to protect DNA replication forks from mobile elements as one of multiple mechanisms of protecting genome integrity.

The Bacterial Nucleoid

The nucleoid was described with respect to chromosome segregation above. Indeed, experiments in many of the model systems indicate that bacteria carefully coordinate the position of the chromosome in the cell. Through techniques in which individual positions in the chromosome can be localized in whole cells, it has been shown that genes are located in the chromosome in roughly the same order as one would presume by looking at the DNA sequence. A variety of techniques are providing insight into how the structure of the chromosome is maintained in the cell. The molecular mechanisms that maintain the chromosome structure remain a mystery, but specific systems are likely to exist to ensure that the chromosome is available for transcription, recombination, and other functions.

Supercoiling in the Nucleoid

Supercoiling is one of the mechanisms that help compact and organize the chromosome. Supercoiling also affects the expression level of many genes. However, supercoiling will be lost if one of the strands of the DNA is cut, thereby allowing the strands to rotate around each other. The phosphodiester bond connecting the two deoxyribose sugars on the other strand serves as a swivel and rotates, resulting in relaxed (i.e., not supercoiled) DNA. A DNA with a phosphodiester bond broken in one of the two strands is said to be nicked. A variety of experiments suggest that the nucleoid is packaged in subregions that constrain supercoiling. Topological barriers would prevent the entire chromosome from losing supercoiling when there is a nick in the genome. Figure 1.26 illustrates supercoiling. In this example, the ends of a DNA molecule have been rotated in opposite directions, and the DNA has become twisted up on itself to relieve the stress. The DNA remains supercoiled as long as its ends are constrained and so cannot rotate, and a circular DNA has no free ends that can rotate. A linear DNA will not maintain supercoiling unless regions flanking the supercoiling are somehow otherwise constrained. A break or nick in a circular DNA should relax the whole DNA unless portions of the molecule are periodically attached to barriers that prevent rotation of the strands. Through the examination of the expression level of over 300 genes that are sensitive to supercoiling following the introduction of breaks in the chromosome, it has been estimated that the topologically isolated loops are about 10 kb in size. This is in good agreement with domain size as determined by directly measuring the lengths of loops under the microscope (see

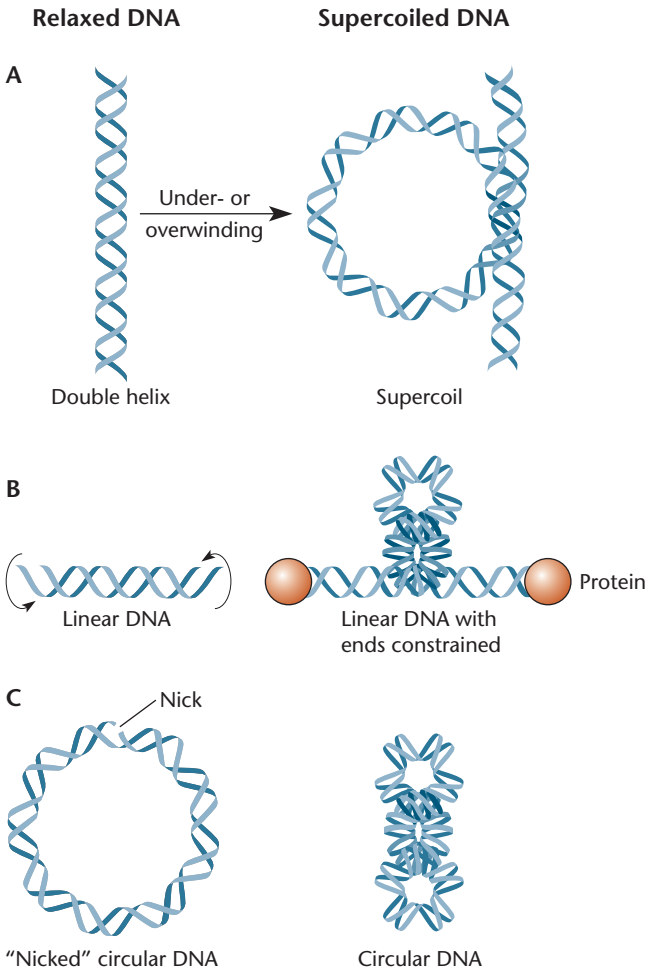


Figure 1.26 **(A)** Supercoiled DNA. **(B)** Twisting of the ends in opposite directions causes linear DNA to wrap up on itself. The supercoiling is lost if the ends of the DNA are not somehow constrained. **(C)** A break, or nick, in one of the two strands of a circular DNA relaxes the supercoils.

Postow et al., Suggested Reading). Although the exact mechanism or mechanisms that restrict topology are unresolved, this work indicates that the barriers to rotation are not fixed at certain places in the chromosome.

SUPERCOILING OF NATURAL DNAs

It is possible to estimate the extent of supercoiling of natural DNAs. According to the Watson-Crick structure, the two strands are wrapped around each other about once every 10.5 bp to form the double helix. Therefore, in a DNA of 2,100 bp, the two strands should be wrapped around each other about 2,100/10.5, or 200, times. In a supercoiled DNA of this size, however, the two strands are wrapped around each other either more or less than 200 times. If they are wrapped around each other more than once every 10.5 bp, the DNA is said to be **positively supercoiled**; if less than once every 10.5 bp, it is **negatively supercoiled**.

Most DNA in bacteria is negatively supercoiled, with an average of one negative supercoil for every 300 bp, although there are localized regions of higher or lower negative supercoiling. Also, in some regions, such as ahead of a transcribing RNA polymerase, the DNA may be positively supercoiled (see above).

SUPERCOILING STRESS

Some of the stress due to supercoiling of the DNA, which causes it to twist up on itself, can be relieved if the DNA is wrapped around something else, such as proteins. Sailors know about this effect: if you twist a rope in the right direction as you roll it up to store it, it does not try to unroll itself again when you are finished. Wrapping DNA around proteins in the cell is called constraining the supercoils. Unconstrained supercoils cause stress in the DNA, which can be relieved by twisting the DNA up on itself, as shown in Figure 1.26, and making the DNA more compact. The stress due to unconstrained supercoils can have other effects, as well, for example, helping to separate the strands of DNA during reactions such as replication, recombination, and initiation of RNA synthesis at promoters.

Topoisomerases

The supercoiling of DNA in the cell is modulated by topoisomerases (see Wang, Suggested Reading). Topoisomerases are discussed above, but not the molecular details of these enzymes. All organisms have these proteins, which manage to remove the supercoils from a circular DNA without permanently breaking either of the two strands. They perform this feat by binding to DNA, breaking one or both of the strands, and passing the DNA strands through the break before resealing it. As long as the enzyme holds the cut ends of the DNA so that they do not rotate, this process, known as **strand passage**, either introduces or removes supercoils in DNA.

The topoisomerases are classified into two groups, type I and type II (Figure 1.27). These two types differ in how many strands are cut and how many strands pass through the cut. The type I topoisomerases cut one strand and pass the other strand through the break before resealing the cut. The type II topoisomerases cut both strands and pass two other strands from somewhere else in the DNA, or even another DNA, through the break before resealing it. This basic difference changes how supercoiling is affected by these enzymes, as shown in Figure 1.27.

TYPE I TOPOISOMERASES

Bacteria have several type I topoisomerases. The major bacterial type I topoisomerase removes negative supercoils from DNA. In *E. coli* and *Salmonella enterica* serovar Typhimurium, the *topA* gene encodes this type I topoisomerase. As expected, DNA isolated from *E. coli* with a *topA* mutation is more negatively supercoiled than normal.

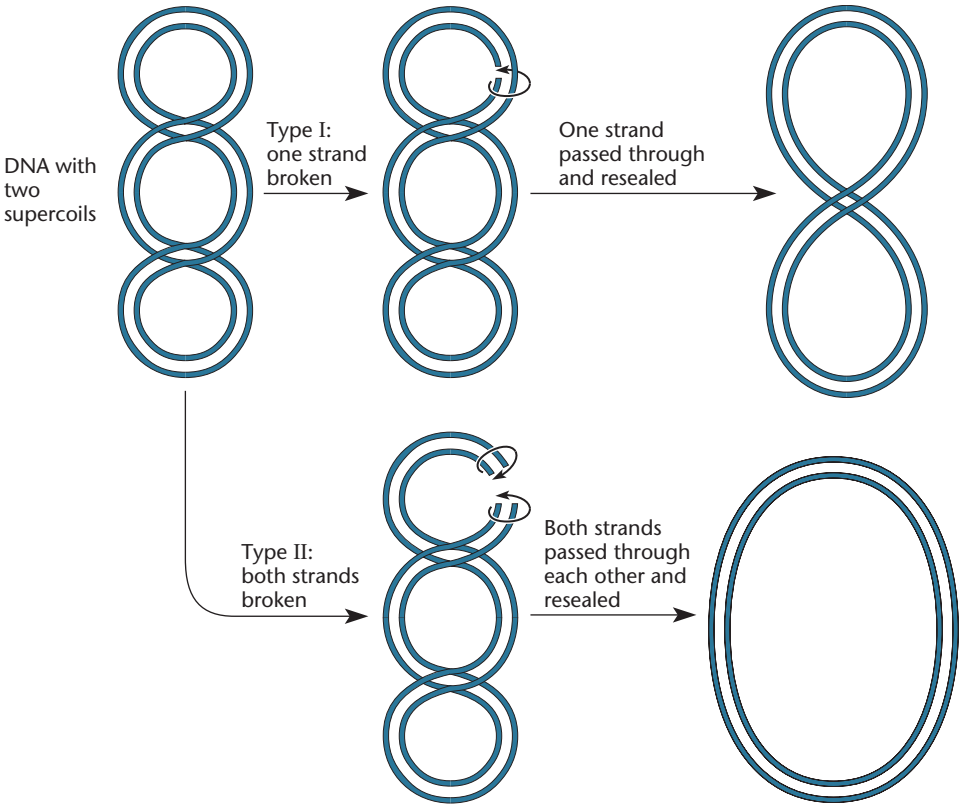


Figure 1.27 Action of the two types of topoisomerases. The type I topoisomerases break one strand of DNA and pass the other strand through the break, removing one supercoil at a time. The type II topoisomerases break both strands and pass another part of the same DNA through the breaks, introducing or removing two supercoils at a time.

TYPE II TOPOISOMERASES

Bacteria also have more than one type II topoisomerase. Because type II topoisomerases can break both strands and pass two other DNA strands through the break, they can either separate two linked circular DNA molecules or link them up. Linkage sometimes happens after replication or recombination. One major type II topoisomerase in *E. coli*, Topo IV (see above), decatenates daughter chromosomes after DNA replication, releasing the major source of cohesion between the chromosomes and allowing them to be segregated into the daughter cells. While most type II topoisomerases remove negative supercoils, a special type II topoisomerases in bacteria, called gyrase, plays an essential role in adding negative supercoils. Gyrase acts by first wrapping the DNA around itself and then cutting the two strands before passing another part of the DNA through the cuts, thereby introducing two negative supercoils. Adding negative supercoils increases the stress in the DNA and thus requires energy; hence, gyrase needs ATP for this reaction.

The gyrase of *E. coli* is made up of four polypeptides, two of which are encoded by the *gyrA* gene and two of which are encoded by *gyrB*. These genes were first identified by mutations that make the cell resistant to antibiotics that affect gyrase (Table 1.2 and Box 1.2). The GyrA subunits seem to be responsible for breaking the DNA and holding it as the strands pass through the cuts. The GyrB subunits have the ATP site that furnishes the energy for the supercoiling.

Table 1.2 Antibiotics that block replication

Antibiotic	Source	Target
Trimethoprim	Chemically synthesized	Dihydrofolate reductase
Hydroxyurea	Chemically synthesized	Ribonucleotide reductase
5-Fluorodeoxyuridine	Chemically synthesized	Thymidylate synthetase
Nalidixic acid	Chemically synthesized	*gyrA* subunit of gyrase
Novobiocin	*Streptomyces sphaeroides*	*gyrB* subunit of gyrase
Mitomycin C	*Streptomyces caespitosus*	Cross-links DNA

BOX 1.2

Antibiotics That Affect Replication and DNA Structure

Antibiotics are substances that block the growth of cells. Many antibiotics are naturally synthesized chemical compounds made by soil microorganisms, especially actinomycetes, that may help them compete with other soil microorganisms. There are also other ideas as to why bacteria make antibiotics, including for their use in intercellular communication, especially in highly organized structures like biofilms. Many antibiotics specifically block DNA replication or change the structure of DNA. Because some parts of the replication machinery have remained relatively unchanged throughout evolution, many of these antibiotics work against essentially all types of bacteria. Some even work against eukaryotic cells and so are used as antifungal agents and in antitumor chemotherapy.

Antibiotics That Block Precursor Synthesis

As discussed above, DNA is polymerized from the deoxynucleoside triphosphates. Any antibiotic that blocks the synthesis of these deoxynucleotide precursors will block DNA replication.

Inhibition of dihydrofolate reductase

Some of the most important precursor synthesis blockers are antibiotics that inhibit the enzyme dihydrofolate reductase. One such compound, trimethoprim, works very effectively in bacteria, and the antitumor drug methotrexate (amethopterin) inhibits the dihydrofolate reductase of eukaryotes. Methotrexate is used as an antitumor agent, among other uses such as in the treatment of inflammatory arthritis.

Antibiotics like trimethoprim that inhibit dihydrofolate reductase kill the cell by depleting it of tetrahydrofolate, which is needed for many biosynthetic reactions. This inhibition is overcome, however, if the cell lacks the enzyme thymidylate synthetase, which synthesizes dTMP; therefore, most mutants that are resistant to trimethoprim have mutations that inactivate the *thyA* thymidylate synthetase gene. The reason is apparent from the pathway for dTMP synthesis shown in Figure 1.5. Thymidylate synthetase is solely responsible for converting tetrahydrofolate to dihydrofolate when it transfers a methyl group from tetrahydrofolate to dUMP to make dTMP. The dihydrofolate reductase is the only enzyme in the cell that can restore the tetrahydrofolate needed for other biosynthetic reactions. However, if the cell lacks thymidylate synthetase, there is no need for a dihydrofolate reductase to restore tetrahydrofolate. Therefore, inhibition of dihydrofolate reductase by trimethoprim has no effect, thus making *thyA* mutant cells resistant to the antibiotic. Of course, if the cell lacks a thymidylate synthetase, it cannot make its own dTMP from dUMP and must be provided with thymidine in the medium so that it can replicate its DNA.

There is more than one mechanism by which cells can achieve trimethoprim resistance. They can have an altered dihydrofolate reductase to which trimethoprim cannot bind, or they can have more copies of the gene so that they make more enzyme than there is trimethoprim to inhibit it. Some plasmids and transposons carry genes for resistance to trimethoprim. These genes encode dihydrofolate reductases that are much less sensitive to trimethoprim and so can act even in the presence of high concentrations of the antibiotic.

Inhibition of ribonucleotide reductase

The antibiotic hydroxyurea inhibits the enzyme ribonucleotide reductase, which is required for the synthesis of all four precursors of DNA synthesis (Figure 1.5). The ribonucleotide reductase catalyzes the synthesis of the deoxynucleoside diphosphates dCDP, dGDP, dADP, and dUDP from the ribonucleoside diphosphates, an essential step in deoxynucleoside triphosphate synthesis. Mutants resistant to hydroxyurea have an altered ribonucleotide reductase.

Competition with deoxyuridine monophosphate

5-Fluorodeoxyuridine and the related 5-fluorouracil have monophosphate forms resembling dUMP, the substrate for thymidylate synthetase. By competing with the natural substrate for this enzyme, they inhibit the synthesis of deoxythymidine monophosphate. Mutants resistant to these compounds have an altered thymidylate synthetase. These are useful antibiotics for the treatment of fungal, as well as bacterial, infections.

Antibiotics That Block Polymerization of Deoxynucleotides

The polymerization of deoxynucleotide precursors into DNA would also seem to be a tempting target for antibiotics. However, there seem to be surprisingly few antibiotics that directly block this process. Most antibiotics that block polymerization do so indirectly, by binding to DNA or by mimicking the deoxynucleotides and causing chain termination, rather than by inhibiting the DNA polymerase itself.

Deoxynucleotide precursor mimics

Dideoxynucleotides are similar to the normal deoxynucleotide precursors, except that they lack a hydroxyl group on the 3′ carbon of the deoxynucleotide. Consequently, they can be incorporated into DNA, but then replication stops because they cannot link up with the next deoxynucleotide. These compounds are not useful antibacterial agents, probably because

they are not phosphorylated well in bacterial cells. However, this property of prematurely terminating replication has made them the basis for one of the first methods for DNA sequencing (see chapter 13).

Cross-linking

Mitomycin C blocks DNA synthesis by cross-linking the guanine bases in DNA to each other. Sometimes the cross-linked bases are in opposing strands. If the two strands are attached to each other, they cannot be separated during replication. Even one cross-link in DNA that is not repaired prevents replication of the chromosome. This antibiotic is also a useful antitumor drug, probably for the same reason. DNA cross-linking also affects RNA transcription.

Antibiotics That Affect DNA Structure

Acridine dyes

The acridine dyes include proflavine, ethidium, and chloroquine. These compounds insert between the bases of DNA and thereby cause frameshift mutations and inhibit DNA synthesis. Their ability to insert themselves between the bases in DNA has made acridine dyes very useful in genetics and molecular biology. Some of these applications are discussed in later chapters. In general, acridine dyes are not useful as antibiotics because of their toxicity due to their ability to block DNA synthesis in the mitochondria of eukaryotic cells. Some members of this large family of antibiotics have long been used as antimalarial drugs because of their ability to block DNA synthesis in the mitochondria (kinetoplasts) of trypanosomes. This is the basis for the antimalarial activity of the tonic water in a gin and tonic.

Thymidine mimic

5-Bromodeoxyuridine (BUdR) is similar to thymidine and is efficiently phosphorylated and incorporated in its place into DNA. However, BUdR incorporated into DNA often mispairs and increases replication errors. DNA containing BUdR is also more sensitive to some wavelengths of ultraviolet (UV) light (which makes BUdR useful in enrichment schemes for isolating mutants [see chapter 3]). Moreover, DNA containing BUdR has a different density from DNA exclusively containing thymidine (another feature of BUdR that is useful in experiments).

Antibiotics That Affect Gyrase

Many antibiotics and antitumor drugs affect topoisomerases. The type II topoisomerase, gyrase, in bacteria is a target for many different antibiotics. Because this enzyme is similar among all bacteria, these antibiotics have a broad spectrum of activity and kill many types of bacteria.

GyrA inhibition

Nalidixic acid specifically binds to the GyrA subunit, which is involved in cutting the DNA and in strand passage. This activity makes nalidixic acid and its many derivatives, including oxolinic acid and chloromycetin, very useful antibiotics. Another antibiotic that binds to the GyrA subunit, ciprofloxacin, is used for treating gonorrhea, anthrax, and bacterial dysentery. However, because these antibiotics can induce prophages, they may actually make some diseases worse (see chapter 7).

The mechanism of killing by these antibiotics is not completely understood. They are known to cause degradation of DNA and can cause the DNA to become covalently linked to gyrase, presumably by trapping it in an intermediate state in the process of strand passage. Bacteria resistant to nalidixic acid have an altered *gyrA* gene.

GyrB inhibition

Novobiocin and its more potent relative coumermycin bind to the GyrB subunit, which is involved in ATP binding. These antibiotics do not resemble ATP, but by binding to the gyrase they somehow prevent ATP cleavage, perhaps by changing the conformation of the enzyme. Mutants resistant to novobiocin have an altered *gyrB* gene.

The Bacterial Genome

The discussion of the replication and structure of the bacterial genome ignores the complexity of the sequences and the functions they encode. Advances in DNA-sequencing technologies are drastically reducing both the time it takes to sequence DNA and the cost of DNA sequencing. We are now at a time when determining the genomic sequence of a newly discovered bacterium is often one of the earliest steps in trying to fully characterize the functions of a new species. In addition, there are also a large number of studies where DNA from a particular environment or a consortium of organisms is chosen for DNA sequencing instead of that of a single organism derived from pure culture. Bacterial and archaeal genomes are in some ways more amenable to genome analysis than eukaryotic genomes because they are relatively small, ranging from ca. 0.5 megabases (Mb) for some obligate parasites to around 10 Mb for some free-living bacteria. They contain few introns and much less repetitive DNA than eukaryotic genomes. Chapter 13 deals with ideas for how genomes are structured and the tools we use to study them.

Summary

1. DNA consists of two strands wrapped around each other in a double helix. Each strand consists of a chain of nucleotides held together by phosphates joining their deoxyribose sugars. Because the phosphate joins the third carbon of one sugar to the fifth carbon of the next sugar, the DNA strands have directionality, or polarity, and have distinct 5′ phosphate and 3′ hydroxyl ends. The two strands of DNA are antiparallel, so that the 5′ end of one is on the same end as the 3′ end of the other.

2. DNA is synthesized from the precursor deoxynucleoside triphosphates by DNA polymerase. The first phosphate of each nucleotide is attached to the 3′ hydroxyl of the next deoxynucleotide, giving off the terminal two phosphates to provide energy for the reaction.

3. DNA polymerases require both a primer and a template strand. The pairing of the bases between the incoming deoxynucleotide and the base on the template strand dictates which deoxynucleotide will be added at each step, with A always pairing with T and G always pairing with C. The DNA polymerase synthesizes DNA in the 5′-to-3′ direction, moving in the 3′-to-5′ direction on the template.

4. DNA polymerases cannot put down the first deoxynucleotide, so RNA is usually used to prime the synthesis of a new strand. Afterward, the RNA primer is removed and replaced by DNA, using upstream DNA as a primer. The use of RNA primers helps reduce errors by allowing editing.

5. DNA polymerase does not synthesize DNA by itself but needs other proteins to help it replicate DNA. These other proteins are helicases that separate the strands of the DNA, ligases to join two DNA pieces together, primases to synthesize RNA primers, and other accessory proteins to keep the DNA polymerase on the DNA and reduce errors.

6. Both strands of double-stranded DNA are replicated from the same end, so that the overall direction of DNA replication is from 5′ to 3′ on one strand and from 3′ to 5′ on the other strand. Because DNA polymerase can polymerize only in the 5′-to-3′ direction, it must replicate one strand in short pieces and ligate them afterward to form a continuous strand. The short pieces are called Okazaki fragments. The two DNA polymerases replicating the leading and lagging template strands remain bound to each other in a process called the trombone model of replication.

7. Lesions on DNA, proteins bound to DNA, and transcription can stall DNA replication. Primase can reinitiate replication to prevent the entire holoenzyme from stalling but leaving a gap that is repaired by other mechanisms. DNA helicases and other enzymes can be used to help DNA polymerase past bound proteins and to allow the polymerase to continue after it stalls from collisions with RNA polymerase.

8. The DNA in a bacterium that carries most of the genes is called the bacterial chromosome. The chromosome of most bacteria is a long, circular molecule that replicates in both directions from a unique origin of replication, *oriC*. Replication of the chromosomes initiates each time the cells reach a certain size. For fast-growing cells, new rounds of replication initiate before old ones are completed. This accounts for the fact that fast-growing cells have a higher DNA content than slower-growing cells.

9. Chromosome replication terminates, and the two daughter DNAs separate, when the replication forks meet. In some bacteria, multiple *ter* sites that act as "one-way gates" delay movement of the replication forks on the chromosome. Proteins that bind to the *ter* sites that are inhibitors of the DnaB helicase stop replication at these sites.

10. To separate the daughter DNAs after replication, dimerized chromosomes, created by recombination between the daughter DNAs, are resolved by XerC and XerD, a site-specific recombination system that promotes recombination between the *dif* sites on the daughter chromosomes. The FtsK protein is a DNA translocase that promotes XerC-XerD recombination at *dif* sites to prevent dimerized chromosomal DNA from being guillotined by the forming septum. Topo IV decatenates the intertwined daughter DNAs by passing the double-stranded DNAs through each other, and its activity is also regulated temporally and spatially via an interaction with FtsK.

11. The daughter chromosomes are segregated by condensing the DNAs through supercoiling by DNA gyrase and by condensins that hold the DNA in large supercoiled loops.

12. The FtsZ protein forms a ring at the midpoint of the cell, attracting other proteins, which form the division septum.

13. The Min proteins prevent the formation of FtsZ rings anywhere in the cell other than in the middle. Nucleoid occlusion proteins prevent the formation of FtsZ rings over nucleoids.

14. Initiation of a round of chromosome replication occurs once every time the cell divides. Initiation occurs when the ratio of active DnaA protein to origins of replication reaches a critical number. After replication initiates, its ATPase is upregulated by interaction with the β clamp and Hda, converting it to the inactive ADP-bound state to prevent reinitiation. In some bacteria, including *E. coli* and related enteric bacteria, new initiations are prevented by hemimethylation of the newly

Summary (continued)

replicated DNA at the origin and by sequestration until the replication fork has left the origin.

15. The chromosomal DNA of bacteria is usually one long, continuous circular molecule about 1,000 times as long as the cell itself. This long DNA is condensed in a small part of the cell called the nucleoid. In this structure, the DNA loops out of a central condensed core region. Some of these loops of DNA are negatively supercoiled. In *E. coli*, most DNAs have one supercoil about every 300 bases.

16. The enzymes that modulate DNA supercoiling in the cell are called topoisomerases. There are two types of topoisomerases in cells. Type I topoisomerases can remove supercoils one at a time by breaking only one strand and passing the other strand through the break. Type II topoisomerases remove or add supercoils two at a time by breaking both strands and passing another region of the DNA through the break. The enzyme responsible for adding the negative supercoils to DNA in bacteria is a type II topoisomerase called gyrase. Topo IV decatenates daughter DNAs after replication.

QUESTIONS FOR THOUGHT

1. Some viruses, such as adenovirus, avoid the problem of lagging-strand synthesis by replicating the individual strands of the DNA in the leading-strand direction simultaneously from both ends so that eventually the entire molecule is replicated. Why do bacterial chromosomes not replicate in this way?

2. Why are DNA molecules so long? Would it not be easier to have many shorter pieces of DNA? What are the advantages and disadvantages of a single long DNA molecule?

3. Why do cells have DNA as their hereditary material instead of RNA, like some viruses?

4. What effect would shifting a temperature-sensitive mutant with a mutation in the *dnaA* gene for initiator protein DnaA have on the rate of DNA synthesis? Would the rate drop linearly or exponentially? Would the slope of the curve be affected by the growth rate of the cells at the time of the shift? Explain.

5. The gyrase inhibitor novobiocin inhibits the growth of almost all types of bacteria. What would you predict about the gyrase of the bacterium *Streptomyces sphaeroides*, which makes this antibiotic? How would you test your hypothesis?

6. How do you think chromosome replication and cell division are coordinated in bacteria like *E. coli*? How would you go about testing your hypothesis?

7. Why is termination of chromosome replication so sloppy that the *ter* region is nonessential for growth and there has to be more than one *ter* site in each direction to completely stop the replication fork? What are the advantages of not having a definite site on the chromosome at which replication always terminates?

SUGGESTED READING

Aussel L, Barre F-X, Aroyo M, Stasiak A, Stasiak AZ, Sherratt D. 2002. FtsK is a DNA motor protein that activates chromosome dimer resolution by switching the catalytic state of the XerC and XerD recombinases. *Cell* **108:**195–205.

Bernhardt TG, de Boer PAJ. 2005. SlmA, a nucleoid-associated, FtsZ binding protein required for blocking septal ring assembly over chromosomes in *E. coli*. *Mol Cell* **18:**555–564.

Brézellec P, Hoebeke M, Hiet MS, Pasek S, Ferat JL. 2006. DomainSieve: a protein domain-based screen that led to the identification of *dam*-associated genes with potential link to DNA maintenance. *Bioinformatics* **22:**1935–1941.

Biller SJ, Burkholder WF. 2009. The *Bacillus subtilis* SftA (YtpS) and SpoIIIE DNA translocases play distinct roles in growing cells to ensure faithful chromosome partitioning. *Mol Microbiol* **74:**790–809.

Blakely G, May G, McCulloch R, Arciszewska LK, Burke M, Lovett ST, Sherratt DJ. 1993. Two related recombinases are required for site-specific recombination at *dif* and *cer* in *E. coli* K12. *Cell* **75:**351–361.

Britton RA, Lin DC, Grossman AD. 1998. Characterization of a prokaryotic SMC protein involved in chromosome partitioning. *Genes Dev* **12:**1254–1259.

Camara JE, Breier AM, Brendler T, Austin S, Cozzarelli NR, Crooke E. 2005. Hda inactivation of DnaA is the predominant mechanism preventing hyperinitiation of *Escherichia coli* DNA replication. *EMBO Rep* **6:**736–741.

Cortez D, Quevillon-Cheruel S, Gribaldo S, Desnoues N, Sezonov G, Forterre P, Serre M-CM. 2010. Evidence for a Xer/dif system for chromosome resolution in archaea. *PLoS Genet* **6:**e1001166.

Dervyn E, Suski C, Daniel R, Bruand C, Chapuis J, Errington J, Jannière L, Ehrlich SD. 2001. Two essential DNA polymerases at the bacterial replication fork. *Science* **294:**1716–1719.

Dohrmann PR, Correa R, Frisch RL, Rosenberg SM, McHenry CS. 2016. The DNA polymerase III holoenzyme contains γ and is not a trimeric polymerase. *Nucleic Acids Res* **44:**1285–1297.

Fournes F, Val M-E, Skovgaard O, Mazel D. 2018. Replicate once per cell cycle: replication

control of secondary chromosomes. *Front Microbiol* **9:**1833.

Fricker AD, Peters JE. 2014. Vulnerabilities on the lagging-strand template: opportunities for mobile elements. *Annu Rev Genet* **48:**167–186.

Fujimitsu K, Senriuchi T, Katayama T. 2009. Specific genomic sequences of *E. coli* promote replicational initiation by directly reactivating ADP-DnaA. *Genes Dev* **23:**1221–1233.

Gabbai CB, Yeeles JTP, Marians KJ. 2014. Replisome-mediated translesion synthesis and leading strand template lesion skipping are competing bypass mechanisms. *J Biol Chem* **289:**32811–32823.

Galli E, Ferat J-L, Desfontaines J-M, Val M-E, Skovgaard O, Barre FX, Possoz C. 2019. Replication termination without a replication fork trap. *Sci Rep* **9:**8315. http://doi.org/10.1038/s41598-019-43795-2

Guy CP, Atkinson J, Gupta MK, Mahdi AA, Gwynn EJ, Rudolph CJ, Moon PB, van Knippenberg IC, Cadman CJ, Dillingham MS, Lloyd RG, McGlynn P. 2009. Rep provides a second motor at the replisome to promote duplication of protein-bound DNA. *Mol Cell* **36:**654–666.

Hayama R, Marians KJ. 2010. Physical and functional interaction between the condensin MukB and the decatenase topoisomerase IV in *Escherichia coli*. *Proc Natl Acad Sci USA* **107:**18826–18831.

Heller RC, Marians KJ. 2006. Replication fork reactivation downstream of a blocked nascent leading strand. *Nature* **439:**557–562.

Helmstetter CE, Cooper S. 1968. DNA synthesis during the division cycle of rapidly growing *Escherichia coli* B/r. *J Mol Biol* **31:**507–518.

Jean NK, Rutherford TJ, Löwe J. 2019. FtsK in motion reveals its mechanism for double-stranded DNA translocation. *bioRxiv* 1–24.

Joshi MC, Magnan D, Montminy TP, Lies M, Stepankiw N, Bates D. 2013. Regulation of sister chromosome cohesion by the replication fork tracking protein SeqA. *PLoS Genet* **9:**e1003673.

Kasho K, Katayama T. 2013. DnaA binding locus *datA* promotes DnaA-ATP hydrolysis to enable cell cycle-coordinated replication initiation. *Proc Natl Acad Sci USA* **110:**936–941.

Kohanski MA, Dwyer DJ, Hayete B, Lawrence CA, Collins JJ. 2007. A common mechanism of cellular death induced by bactericidal antibiotics. *Cell* **130:**797–810.

Kysela DT, Randich AM, Caccamo PD, Brun YV. 2016. Diversity takes shape: understanding the mechanistic and adaptive basis of bacterial morphology. *PLoS Biol* **14:**e1002565.

Le Bourgeois P, Bugarel M, Campo N, Daveran-Mingot M-L, Labonté J, Lanfranchi D, Lautier T, Pagès C, Ritzenthaler P. 2007. The unconventional Xer recombination machinery of streptococci/lactococci. *PLoS Genet* **3:**e117.

Liu N-J, Dutton RJ, Pogliano K. 2006. Evidence that the SpoIIIE DNA translocase participates in membrane fusion during cytokinesis and engulfment. *Mol Microbiol* **59:**1097–1113.

Mercier R, Petit MA, Schbath S, Robin S, El Karoui M, Boccard F, Espéli O. 2008. The MatP/*matS* site-specific system organizes the terminus region of the *E. coli* chromosome into a macrodomain. *Cell* **135:**475–485.

Moolman MC, Krishnan ST, Kerssemakers JWJ, van den Berg A, Tulinski P, Depken M, Reyes-Lamothe R, Sherratt DJ, Dekker NH. 2014. Slow unloading leads to DNA-bound β2-sliding clamp accumulation in live *Escherichia coli* cells. *Nat Commun* **5:**5820.

Neidhardt FC, Curtiss R III, Ingraham JL, Lin ECC, Low KB, Magasanik B, Reznikoff WS, Riley M, Schaechter M, Umbarger HE (ed). 1996. Escherichia coli *and* Salmonella*: Cellular and Molecular Biology*, 2nd ed. ASM Press, Washington, DC.

Olby R. 1974. *The Path to the Double Helix*. Macmillan Press, London, United Kingdom.

Pomerantz RT, O'Donnell M. 2008. The replisome uses mRNA as a primer after colliding with RNA polymerase. *Nature* **456:**762–766.

Postow L, Hardy CD, Arsuaga J, Cozzarelli NR. 2004. Topological domain structure of the *Escherichia coli* chromosome. *Genes Dev* **18:**1766–1779.

Reddy CA, Beveridge TJ, Breznak JA, Marzluf G, Schmidt TM, Snyder LR (ed). 2007. *Methods for General and Molecular Microbiology*, 3rd ed. ASM Press, Washington, DC.

Reyes-Lamothe R, Sherratt DJ, Leake MC. 2010. Stoichiometry and architecture of active DNA replication machinery in *Escherichia coli*. *Science* **328:**498–501.

Rudolph CJ, Upton AL, Stockum A, Nieduszynski CA, Lloyd RG. 2013. Avoiding chromosome pathology when replication forks collide. *Nature* **500:**608–611.

Sambrook J, Russell D. 2006. *The Condensed Protocols from Molecular Cloning: a Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Sanders GM, Dallmann HG, McHenry CS. 2010. Reconstitution of the *B. subtilis* replisome with 13 proteins including two distinct replicases. *Mol Cell* **37:**273–281.

Shih Y-L, Rothfield L. 2006. The bacterial cytoskeleton. *Microbiol Mol Biol Rev* **70:**729–754.

Singleton MR, Dillingham MS, Wigley DB. 2007. Structure and mechanism of helicases and nucleic acid translocases. *Annu Rev Biochem* **76:**23–50.

Slater S, Wold S, Lu M, Boye E, Skarstad K, Kleckner N. 1995. *E. coli* SeqA protein binds *oriC* in two different methyl-modulated reactions appropriate to its roles in DNA replication initiation and origin sequestration. *Cell* **82:**927–936.

Thanbichler M. 2009. Closing the ring: a new twist to bacterial chromosome condensation. *Cell* **137:**598–600.

Thanbichler M, Shapiro L. 2006. MipZ, a spatial regulator coordinating chromosome segregation with cell division in *Caulobacter*. *Cell* **126:**147–162.

Valens M, Penaud S, Rossignol M, Cornet F, Boccard F. 2004. Macrodomain organization of the *Escherichia coli* chromosome. *EMBO J* **23:**4330–4341.

Valens M, Thiel A, Boccard F. 2016. The MaoP/*maoS* site-specific system organizes the Ori region of the *E. coli* chromosome into a macrodomain. *PLoS Genet* **12:**e1006309.

Vecchiarelli AG, Li M, Mizuuchi M, Hwang LC, Seol Y, Neuman KC, Mizuuchi K. 2016. Membrane-bound MinDE complex acts as a toggle switch that drives Min oscillation coupled to cytoplasmic depletion of MinD. *Proc Natl Acad Sci USA* **113:**E1479–E1488.

Viollier PH, Thanbichler M, McGrath PT, West L, Meewan M, McAdams HH, Shapiro L. 2004. Rapid and sequential movement of individual chromosomal loci to specific subcellular locations during bacterial DNA replication. *Proc Natl Acad Sci USA* **101:**9257–9262.

Wang JC. 1996. DNA topoisomerases. *Annu Rev Biochem* **65:**635–692.

Watson JD. 1968. *The Double Helix*. Atheneum, New York, NY.

Watson JD, Crick FHC. 1953. Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid. *Nature* **171:**737–738.

Windgassen TA, Leroux M, Satyshur KA, Sandler SJ, Keck JL. 2018. Structure-specific DNA replication fork recogntion directs helicase and replication restart activities of the *Escherichia coli* PriA helicase. *Proc Natl Acad Sci USA* **115:**E9075–E9084.

Wu LJ, Ishikawa S, Kawai Y, Oshima T, Ogasawara N, Errington J. 2009. Noc protein binds to specific DNA sequences to coordinate cell division with chromosome segregation. *EMBO J* **28:**1940–1952.

Yamaichi Y, Niki H. 2004. *migS*, a *cis*-acting site that affects bipolar positioning of *oriC* on the *Escherichia coli* chromosome. *EMBO J* **23:**221–233.