REPLICATION OF LINEAR ds DNA

The Complete "Replication Apparatus" Is Complex

When Watson and Crick worked out the double-helix structure of DNA, they immediately recognized that the complementary nature of the two strands provided a simple basis for the faithful duplication of genetic material. Meselson and Stahl's demonstration of the semiconservative replication of the *E. coli* chromosome solidified the concept that the two strands of the double helix unwind and serve as templates for the

synthesis of complementary strands. Thus, a parental double helix directs the synthesis of two identical progeny double helices. Kornberg's isolation of an enzyme, DNA polymerase I, capable of synthesizing DNA *in vitro* appeared to provide the final link in what was thought to be an elegantly simple mechanism for the replication of the genetic material—but such was not the case. Twenty years later, scientists are still trying to work out the details of the mechanism of DNA replication.

DNA replication is complex. It is carried out by a ***multienzyme complex***, often called the ***replication apparatus*** or the ***replisome***. In eukaryotes, the components of the replication machinery are just begin-

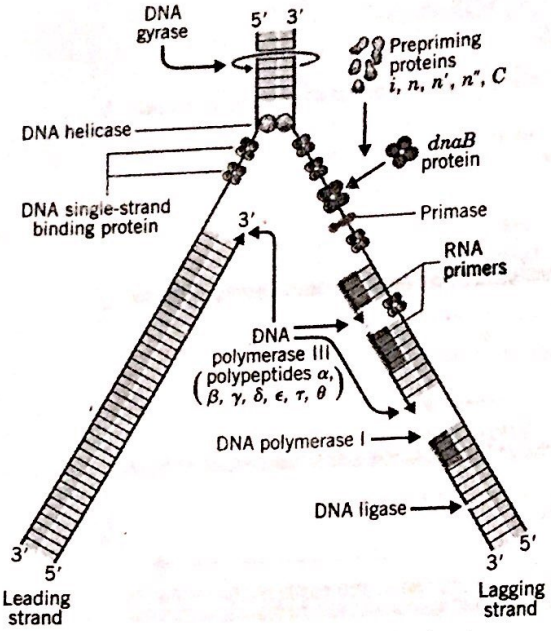

Figure 5.30 Complexity of the *E. coli* replication apparatus. Only those proteins that have been purified (or partially purified) and studied *in vitro* are shown. Other gene-products, such as the products of genes *dnaJ, dnaK, dnaL, dnaP,* and *dnaT,* are known to be required for replication. However, these gene-products have not yet been identified. (After A. Kornberg, *DNA Replication,* Freeman, San Francisco, 1980.)

ning to be identified. Even in prokaryotes, DNA replication requires many different proteins, and the details of how some of these proteins function in DNA replication are still being investigated today. For example, DNA replication in *E. coli* requires at least two dozen different gene-products. Many of these gene-products have been purified and their roles in DNA replication studied *in vitro*. Figure 5.30 shows the involvement of some of these *E. coli* proteins in DNA replication; it is intended to illustrate the complexity of the replication process rather than to illustrate the specific roles of the individual gene-products.

First, the two complementary strands of the parental double helix have to be unwound and separated so that each can serve as a template for the synthesis of a new daughter strand. Unwinding and movement of the replication fork occur ***processively*** with the strands being transiently unwound ahead of the fork as it moves along the chromosome. Three different types of proteins appear to contribute to unwinding the strands of double helices. (1) ***DNA unwinding proteins*** or ***DNA helicases*** are directly involved in catalyzing the unwinding of the double helices. In *E. coli*, two different helicases are involved. One helicase, the product of the *rep* gene, binds to and stimulates separation of the strand that has 3′ to 5′ polarity in the direction of replication fork movement. The other helicase (exact identity still uncertain) binds to and assists unwinding of the strand that has 5′ to 3′ polarity in the direction that the fork is moving. (2) ***DNA single-strand binding proteins (SSBPs)*** bind tightly to single-stranded regions of DNA produced by the action of the helicases and help stabilize the extended single-stranded templates needed for polymerization. The SSBPs bind to DNA as tetramers, and their binding exhibits cooperativity (i.e., the binding of one tetramer stimulates the binding of additional tetramers to adjacent segments of single-stranded DNA). The binding of SSBP to single-stranded DNA tends to hold that DNA in an extended configuration and prevents it from folding back on itself. Single-stranded DNA that is saturated with bound SSBP replicates over 100 times faster than uncomplexed single-stranded DNA *in vitro*. Presumably, uncomplexed single strands of DNA form secondary structures that interfere with the movement of DNA polymerases or other components of the replication complex along the molecule in the normal processive manner. (3) Finally, ***DNA gyrases,*** which catalyze the formation of negative supercoils in DNA (see Fig. 5.36), are essential for replication and are believed to play a key role in the unwinding process. Supercoiling has been proposed to help "drive" the unwinding process, however, we still do not know how this works. Very recently, it has been suggested that DNA gyrase may function by removing positive supercoils that accumulate in front of the replication forks as the helicases unwind the double helices. In any case, DNA gyrases are essential for DNA replication and somehow maintain pre- and postreplicative DNAs in the proper topological structures.

Nascent DNA strands are then initiated by the use of ***RNA primers*** by the mechanism discussed earlier (see Fig. 5.29). Synthesis of the RNA primers is catalyzed by a special class of enzymes called ***primases***. Primase activity requires the formation of a complex of primase and at least six other proteins; this complex is called the ***primosome***. In addition to primase, the primosome contains prepriming proteins tentatively designated ***proteins i, n, n′, and n″*** plus the ***products of genes*** **dnaB** ***and*** **dnaC** (Table 5.4). The primosome carries out the initial priming reaction for the ***leading strand*** (the strand extended continuously in the overall 5′ to 3′ direction) and the repeating priming of the synthesis of "Okazaki fragments" for the ***lagging strand*** (synthesized discontinuously in the overall 3′ to 5′ direction—but 5′ to 3′ at the molecular level, see Fig. 5.28). The functions of the individual proteins in the primosome are still uncertain.

The covalent extension (see Fig. 5.26) of the primed DNA chains during chromosome replication in

E. coli is carried out by DNA polymerase III. Unlike DNA polymerase I of *E. coli* (which is a single polypeptide; see Fig. 5.25), ***DNA polymerase III is a complex enzyme containing seven different polypeptides*** (Fig. 5.31), and all of these polypeptides must be present for proper replicative function. The 5′ to 3′ polymerase activity and the 5′ to 3′ exonuclease activity are both present on the α polypeptide of DNA polymerase III. The 3′ to 5′ proofreading activity (see Fig. 5.27) of polymerase III is present on the ϵ polypeptide. The functions of the other subunits are still uncertain. Subsequent to DNA polymerase III activity at the replication fork, DNA polymerase I catalyzes the removal of the RNA primers by the concerted action of its 5′ to 3′ exonuclease activity and its 5′ to 3′ polymerase activity, and DNA ligase catalyzes covalent closure of the resulting single-stranded "nick" (Fig. 5.30).

Several of the components essential for DNA replication have been identified genetically, that is, *E. coli* strains carrying mutations (heritable changes in the genetic material; see Chapter 11) that result in the

inability to replicate DNA under certain conditions (usually high temperature) have been identified. When these mutations were characterized genetically (see Chapters 7 and 8), they were found to identify a set of genes (designated *dnaA*, *dnaB*, etc.) whose products are required for DNA synthesis *in vivo*. The products of some of these genes are known. For example, *dnaE*, *dnaN*, *dnaX*, and *dnaZ* code for four of the seven subunits (polypeptides) of the complete DNA polymerase III enzyme, and *dnaG* specifies the primase. The products and functions of others (see Table 5.4) are still unknown. Other components of the replication enzymes (e.g., some of the subunits of DNA polymerase III) were discovered by biochemical analyses, and the genes that encode these proteins have still not been identified.