DNA REPLICATION

4.14.1 Introduction

Key Terms

The **replisome** is the multiprotein structure that assembles at the bacterial replicating fork to undertake synthesis of DNA. It contains DNA polymerase and other enzymes.

A **dna mutant** of bacteria is temperature-sensitive; it cannot synthesize DNA at 42°C, but can do so at 37°C.

A **quick-stop mutant** is a type of DNA replication temperature-sensitive mutant (*dna*)in *E. coli* that immediately stops DNA replication when the temperature is increased to 42°C.

A **slow-stop mutant** is a type of DNA replication temperature-sensitive mutant in E. coli that can finish a round of replication at the unpermissive temperature, but cannot start another.

***In vitro* complementation** is a functional assay used to identify components of a process. The reaction is reconstructed using extracts from a mutant cell. Fractions from wild-type cells are then tested for restoration of activity.

Replication of duplex DNA is a complex endeavor involving a conglomerate of enzyme activities. Different activities are involved in the stages of initiation, elongation, and termination.

- Initiation involves recognition of an origin by a complex of proteins. Before DNA synthesis begins, the parental strands must be separated and (transiently) stabilized in the single-stranded state. Then synthesis of daughter strands can be initiated at the replication fork.
- Elongation is undertaken by another complex of proteins. The **replisome** exists only as a protein complex associated with the particular structure that DNA takes at the replication fork. It does not exist as an independent unit (for example, analogous to the ribosome). As the replisome moves along DNA, the parental strands unwind and daughter strands are synthesized.
- At the end of the replicon, joining and/or termination reactions are necessary. Following termination, the duplicate chromosomes must be separated from one another, which requires manipulation of higher-order DNA structure.

Inability to replicate DNA is fatal for a growing cell. Mutants in replication must therefore be obtained as conditional lethals. These are able to accomplish replication under permissive conditions (provided by the normal temperature of incubation), but they are defective under nonpermissive conditions (provided by the higher temperature of 42°C). A comprehensive series of such temperature-sensitive mutants in *E. coli* identifies a set of loci called the *dna* genes. The **dna mutants** distinguish two stages of replication by their behavior when the temperature is raised (545):

© 2004. Virtual Text / www.ergito.com

- The major class of **quick-stop mutants** cease replication immediately on a temperature rise. They are defective in the components of the replication apparatus, typically in the enzymes needed for elongation (but also include defects in the supply of essential precursors).

- The smaller class of **slow-stop mutants** complete the current round of replication, but cannot start another. They are defective in the events involved in initiating a cycle of replication at the origin.

An important assay used to identify the components of the replication apparatus is called ***in vitro* complementation**. An *in vitro* system for replication is prepared from a *dna* mutant and operated under conditions in which the mutant gene product is inactive. Extracts from wild-type cells are tested for their ability to restore activity. The protein coded by the *dna* locus can be purified by identifying the active component in the extract.

Each component of the bacterial replication apparatus is now available for study *in vitro* as a biochemically pure product, and is implicated *in vivo* by mutations in its gene. Eukaryotic replication systems are highly purified, and usually have components analogous to the bacterial proteins, but have not necessarily reached the stage of identification of every single component.

© 2004. Virtual Text / www.ergito.com

References

545. Hirota, Y., Ryter, A., and Jacob, F. (1968). *Thermosensitive mutants of E. coli affected In the processes of DNA synthesis and cellular division*. Cold Spring Harbor Symp. Quant. Biol. 33, 677-693.

This content is available online at http://www.ergito.com/main.jsp?bcs=MBIO.4.14.1

© 2004. Virtual Text / www.ergito.com

DNA REPLICATION

4.14.2 DNA polymerases are the enzymes that make DNA

Key Terms

Replication of duplex DNA takes place by synthesis of two new strands that are complementary to the parental strands. The parental duplex is replaced by two identical daughter duplexes, each of which has one parental strand and one newly synthesized strand. Replication is called semiconservative because the conserved units are the single strands of the parental duplex.

Repair of damaged DNA can take place by repair synthesis, when a strand that has been damaged is excised and replaced by the synthesis of a new stretch. It can also take place by recombination reactions, when the duplex region containing the damaged is replaced by an undamaged region from another copy of the genome.

A **DNA polymerase** is an enzyme that synthesizes a daughter strand(s) of DNA (under direction from a DNA template). Any particular enzyme may be involved in repair or replication (or both).

A **DNA replicase** is a DNA-synthesizing enzyme required specifically for replication.

Key Concepts

- DNA is synthesized in both semiconservative replication and repair reactions.
- A bacterium or eukaryotic cell has several different DNA polymerase enzymes.
- One bacterial DNA polymerase undertakes semiconservative replication; the others are involved in repair reactions.
- Eukaryotic nuclei, mitochondria, and chloroplasts each have a single unique DNA polymerase required for replication, and other DNA polymerases involved in ancillary or repair activities.

There are two basic types of DNA synthesis.

Figure 14.1 shows the result of semiconservative **replication**. The two strands of the parental duplex are separated, and each serves as a template for synthesis of a new strand. The parental duplex is replaced with two daughter duplexes, each of which has one parental strand and one newly synthesized strand.

© 2004. Virtual Text / www.ergito.com

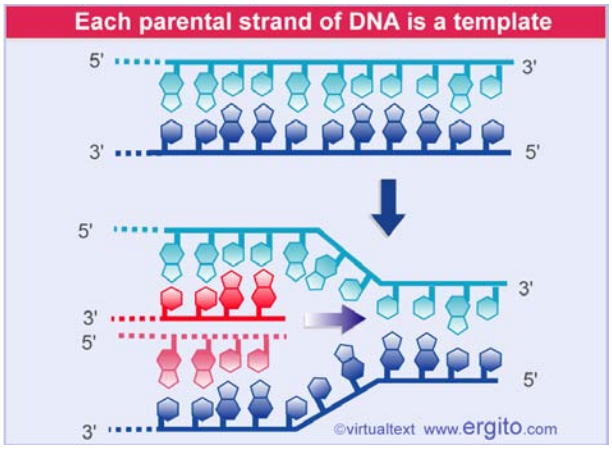


Figure 14.1 Semiconservative replication synthesizes two new strands of DNA.

Figure 14.2 shows the consequences of a **repair** reaction. One strand of DNA has been damaged. It is excised and new material is synthesized to replace it. (Repair synthesis is not the only way to replace damaged DNA; the reactions involved in replacement of damaged sequences are discussed in *Molecular Biology 4.15 Recombination and repair*.)

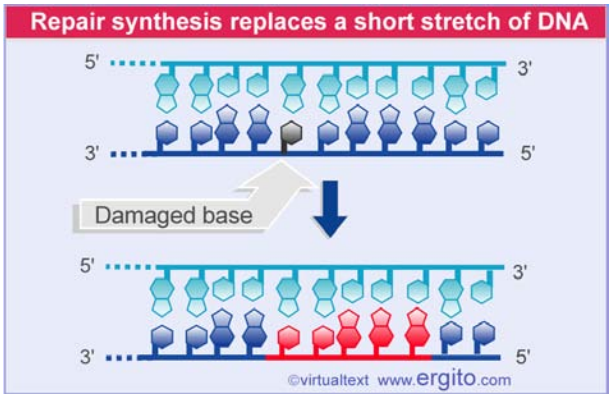


Figure 14.2 Repair synthesis replaces a short stretch of one strand of DNA containing a damaged base.

An enzyme that can synthesize a new DNA strand on a template strand is called a **DNA polymerase**. Both prokaryotic and eukaryotic cells contain multiple DNA polymerase activities. Only some of these enzymes actually undertake replication; sometimes they are called **DNA replicases**. The others are involved in subsidiary roles in replication and/or participate in repair synthesis.

All prokaryotic and eukaryotic DNA polymerases share the same fundamental type of synthetic activity. Each can extend a DNA chain by adding nucleotides one at a time to a 3′–OH end, as illustrated diagrammatically in **Figure 14.3**. The choice of the nucleotide to add to the chain is dictated by base pairing with the template strand.

© 2004. Virtual Text / www.ergito.com

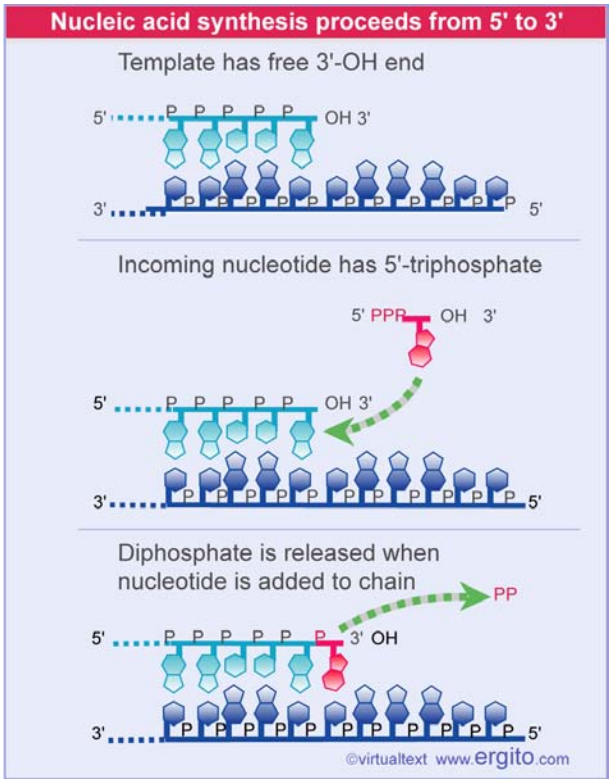


Figure 14.3 DNA synthesis occurs by adding nucleotides to the 3′-OH end of the growing chain, so that the new chain grows in the 5′ → 3′ direction. The precursor for DNA synthesis is a nucleoside triphosphate, which loses the terminal two phosphate groups in the reaction.

Some DNA polymerases function as independent enzymes, but others (most notably the replicases) are incorporated into large protein assemblies. The DNA-synthesizing subunit is only one of several functions of the replicase, which typically contains many other activities concerned with unwinding DNA, initiating new strand synthesis, and so on.

Figure 14.4 summarizes the DNA polymerases that have been characterized in *E. coli*. DNA polymerase III, a multisubunit protein, is the replicase responsible for *de novo* synthesis of new strands of DNA. DNA polymerase I (coded by *polA*) is involved in the repair of damaged DNA and, in a subsidiary role, in semiconservative replication. DNA polymerase II is required to restart a replication fork when its progress is blocked by damage in DNA. DNA polymerases IV and V are involved in allowing replication to bypass certain types of damage.

© 2004. Virtual Text / www.ergito.com

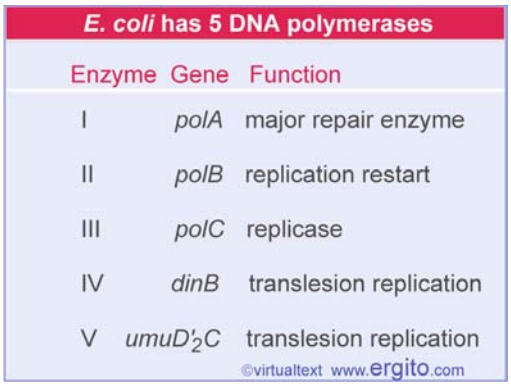


Figure 14.4 Only one DNA polymerase is the replicase. The others participate in repair of damaged DNA, restarting stalled replication forks, or bypassing damage in DNA.

When extracts of *E. coli* are assayed for their ability to synthesize DNA, the predominant enzyme activity is DNA polymerase I. Its activity is so great that it makes it impossible to detect the activities of the enzymes actually responsible for DNA replication! To develop *in vitro* systems in which replication can be followed, extracts are therefore prepared from *polA* mutant cells.

Some phages code for DNA polymerases. They include T4, T5, T7, and SPO1. The enzymes all possess 5′–3′ synthetic activities and 3′–5′ exonuclease proofreading activities (see *Molecular Biology 4.14.3 DNA polymerases have various nuclease activities*). In each case, a mutation in the gene that codes for a single phage polypeptide prevents phage development. Each phage polymerase polypeptide associates with other proteins, of either phage or host origin, to make the intact enzyme.

Several classes of eukaryotic DNA polymerases have been identified. DNA polymerases δ and ε are required for nuclear replication; DNA polymerase α is concerned with "priming" (initiating) replication. Other DNA polymerases are involved in repairing damaged nuclear DNA (β and also ε) or with mitochondrial DNA replication (γ).

Last updated on January 8, 2004

This content is available online at http://www.ergito.com/main.jsp?bcs=MBIO.4.14.2

© 2004. Virtual Text / www.ergito.com

DNA REPLICATION

4.14.3 DNA polymerases have various nuclease activities

Key Terms

Nick translation describes the ability of *E. coli* DNA polymerase I to use a nick as a starting point from which one strand of a duplex DNA can be degraded and replaced by resynthesis of new material; is used to introduce radioactively labeled nucleotides into DNA *in vitro*.

Key Concepts

- DNA polymerase I has a unique 5′–3′ exonuclease activity that can be combined with DNA synthesis to perform nick translation.

Replicases often have nuclease activities as well as the ability to synthesize DNA. A 3′–5′ exonuclease activity is typically used to excise bases that have been added to DNA incorrectly. This provides a "proofreading" error-control system (see *Molecular Biology 4.14.4 DNA polymerases control the fidelity of replication*).

The first DNA-synthesizing enzyme to be characterized was DNA polymerase I, which is a single polypeptide of 103 kD. The chain can be cleaved into two parts by proteolytic treatment. The C-terminal two-thirds of the protein contains the polymerase active site, while the N-terminal third contains the proofreading exonuclease.

The larger cleavage product (68 kD) is called the Klenow fragment. It is used in synthetic reactions *in vitro*. It contains the polymerase and the 3′–5′ exonuclease activities. The active sites are ~30 Å apart in the protein, indicating that there is spatial separation between adding a base and removing one.

The small fragment (35 kD) possesses a 5′–3′ exonucleolytic activity, which excises small groups of nucleotides, up to ~10 bases at a time. This activity is coordinated with the synthetic/proofreading activity. It provides DNA polymerase I with a unique ability to start replication *in vitro* at a nick in DNA. (No other DNA polymerase has this ability.) At a point where a phosphodiester bond has been broken in a double-stranded DNA, the enzyme extends the 3′–OH end. As the new segment of DNA is synthesized, it displaces the existing homologous strand in the duplex.

This process of **nick translation** is illustrated in **Figure 14.5**. The displaced strand is degraded by the 5′–3′ exonucleolytic activity of the enzyme. The properties of the DNA are unaltered, except that a segment of one strand has been replaced with newly synthesized material, and the position of the nick has been moved along the duplex. This is of great practical use; nick translation has been a major technique for introducing radioactively labeled nucleotides into DNA *in vitro*.

© 2004. Virtual Text / www.ergito.com

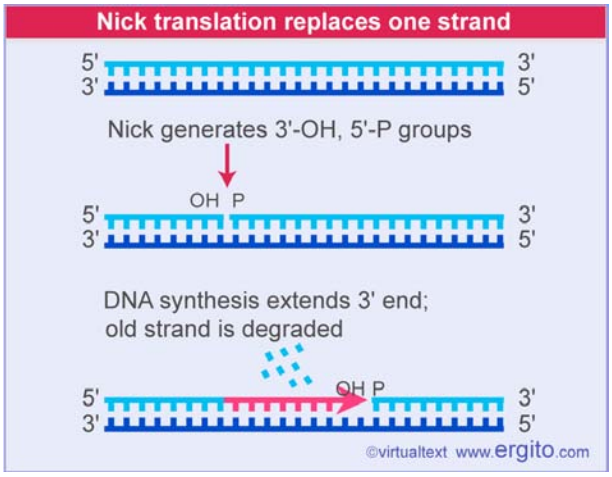


Figure 14.5 Nick translation replaces part of a pre-existing strand of duplex DNA with newly synthesized material.

The 5′–3′ synthetic/3′–5′ exonucleolytic action is probably used *in vivo* mostly for filling in short single-stranded regions in double-stranded DNA. These regions arise during replication, and also when bases that have been damaged are removed from DNA (see *Molecular Biology 4.15 Recombination and repair*). .

This content is available online at http://www.ergito.com/main.jsp?bcs=MBIO.4.14.3

© 2004. Virtual Text / www.ergito.com

DNA REPLICATION

4.14.4 DNA polymerases control the fidelity of replication

Key Terms

Processivity describes the ability of an enzyme to perform multiple catalytic cycles with a single template instead of dissociating after each cycle.

Proofreading refers to any mechanism for correcting errors in protein or nucleic acid synthesis that involves scrutiny of individual units *after* they have been added to the chain.

Key Concepts

- DNA polymerases often have a 3′–5′ exonuclease activity that is used to excise incorrectly paired bases.
- The fidelity of replication is improved by proofreading by a factor of ~100.

The fidelity of replication poses the same sort of problem we have encountered already in considering (for example) the accuracy of translation. It relies on the specificity of base pairing. Yet when we consider the interactions involved in base pairing, we would expect errors to occur with a frequency of $\sim 10^{-3}$ per base pair replicated. The actual rate in bacteria seems to be $\sim 10^{-8}$-10^{-10}. This corresponds to ~1 error per genome per 1000 bacterial replication cycles, or $\sim 10^{-6}$ per gene per generation.

We can divide the errors that DNA polymerase makes during replication into two classes.

- Frameshifts occur when an extra nucleotide is inserted or omitted. Fidelity with regard to frameshifts is affected by the **processivity** of the enzyme: the tendency to remain on a single template rather than to dissociate and reassociate. This is particularly important for the replication of a homopolymeric stretch, for example, a long sequence of dT_n:dA_n, in which "replication slippage" can change the length of the homopolymeric run. As a general rule, increased processivity reduces the likelihood of such events. In multimeric DNA polymerases, processivity is usually increased by a particular subunit that is not needed for catalytic activity *per se*.
- Substitutions occur when the wrong (improperly paired) nucleotide is incorporated. The error level is determined by the efficiency of **proofreading**, in which the enzyme scrutinizes the newly formed base pair and removes the nucleotide if it is mispaired.

All of the bacterial enzymes possess a 3′–5′ exonucleolytic activity that proceeds in the reverse direction from DNA synthesis. This provides the proofreading function

© 2004. Virtual Text / www.ergito.com

illustrated diagrammatically in **Figure 14.6**. In the chain elongation step, a precursor nucleotide enters the position at the end of the growing chain. A bond is formed. The enzyme moves one base pair farther, ready for the next precursor nucleotide to enter. If a mistake has been made, however, the enzyme uses the exonucleolytic activity to excise the last base that was added.

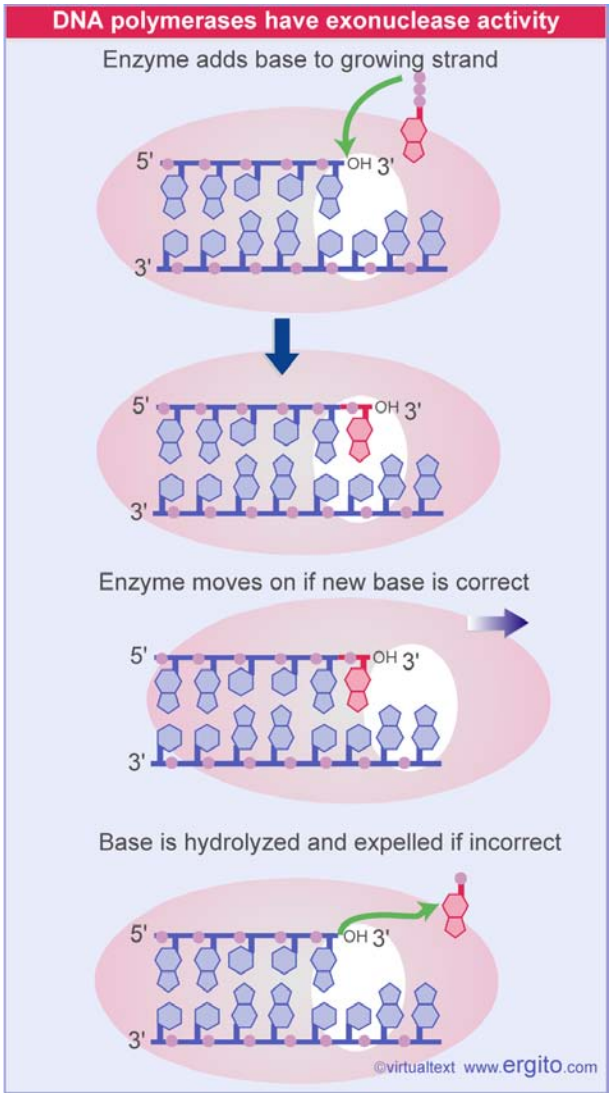


Figure 14.6 Bacterial DNA polymerases scrutinize the base pair at the end of the growing chain and excise the nucleotide added in the case of a misfit.

Different DNA polymerases handle the relationship between the polymerizing and proofreading activities in different ways. In some cases, the activities are part of the same protein subunit, but in others they are contained in different subunits. Each DNA polymerase has a characteristic error rate that is reduced by its proofreading activity. Proofreading typically decreases the error rate in replication from ~10^{-5} to ~10^{-7} per base pair replicated. Systems that recognize errors and correct them following replication then eliminate some of the errors, bringing the overall rate to <10^{-9} per base pair replicated (see *Molecular Biology 4.15.24 Controlling the*

© 2004. Virtual Text / www.ergito.com

direction of mismatch repair).

The replicase activity of DNA polymerase III was originally discovered by a lethal mutation in the *dnaE* locus, which codes for the 130 kD α subunit that possesses the DNA synthetic activity. The 3′–5′ exonucleolytic proofreading activity is found in another subunit, ε, coded by *dnaQ*. The basic role of the ε subunit in controlling the fidelity of replication *in vivo* is demonstrated by the effect of mutations in *dnaQ*: the frequency with which mutations occur in the bacterial strain is increased by $>10^3$-fold.

This content is available online at http://www.ergito.com/main.jsp?bcs=MBIO.4.14.4

© 2004. Virtual Text / www.ergito.com

DNA REPLICATION

4.14.5 DNA polymerases have a common structure

Key Concepts

- Many DNA polymerases have a large cleft composed of three domains that resemble a hand.
- DNA lies across the "palm" in a groove created by the "fingers" and "thumb".

Figure 14.7 shows that all DNA polymerases share some common structural features (3091; 3092). The enzyme structure can be divided into several independent domains, which are described by analogy with a human right hand. DNA binds in a large cleft composed of three domains. The "palm" domain has important conserved sequence motifs that provide the catalytic active site. The "fingers" are involved in positioning the template correctly at the active site. The "thumb" binds the DNA as it exits the enzyme, and is important in processivity. The most important conserved regions of each of these three domains converge to form a continuous surface at the catalytic site. The exonuclease activity resides in an independent domain with its own catalytic site. The N-terminal domain extends into the nuclease domain. DNA polymerases fall into five families based on sequence homologies; the palm is well conserved among them, but the thumb and fingers provide analogous secondary structure elements from different sequences.

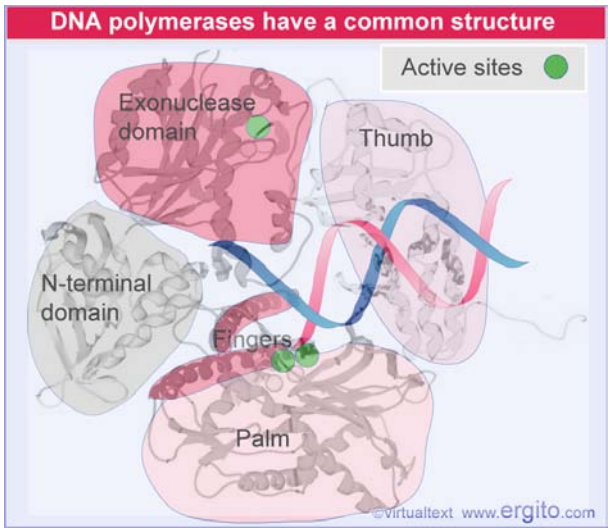


Figure 14.7 The common organization of DNA polymerases has a palm that contains the catalytic site, fingers that position the template, a thumb that binds DNA and is important in processivity, an exonuclease domain with its own active site, and an N-terminal domain.

The catalytic reaction in a DNA polymerase occurs at an active site in which a nucleotide triphosphate pairs with an (unpaired) single strand of DNA. The DNA lies

© 2004. Virtual Text / www.ergito.com

across the palm in a groove that is created by the thumb and fingers. **Figure 14.8** shows the crystal structure of the T7 enzyme complexed with DNA (in the form of a primer annealed to a template strand) and an incoming nucleotide that is about to be added to the primer. The DNA is in the classic B-form duplex up to the last 2 base pairs at the 3′ end of the primer, which are in the more open A-form. A sharp turn in the DNA exposes the template base to the incoming nucleotide. The 3′ end of the primer (to which bases are added) is anchored by the fingers and palm. The DNA is held in position by contacts that are made principally with the phosphodiester backbone (thus enabling the polymerase to function with DNA of any sequence).

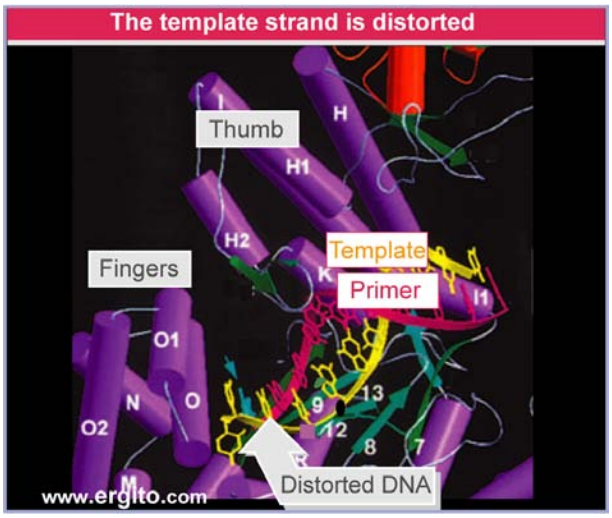


Figure 14.8 The crystal structure of phage T7 DNA polymerase shows that the template strand takes a sharp turn in order to be exposed to the incoming nucleotide. Photograph kindly provided by Charles Richardson and Tom Ellenberger (see 3748).

In structures of DNA polymerases of this family complexed only with DNA (that is, lacking the incoming nucleotide), the orientation of the fingers and thumb relative to the palm is more open, with the O helix (O, O1, O2; see **Figure 14.8**) rotated away from the palm. This suggests that an inward rotation of the O helix occurs to grasp the incoming nucleotide and create the active catalytic site. When a nucleotide binds, the fingers domain rotates 60° toward the palm, with the tops of the fingers moving by 30 Å. The thumb domain also rotates toward the palm by 8°. These changes are cyclical: they are reversed when the nucleotide is incorporated into the DNA chain, which then translocates through the enzyme to recreate an empty site.

The exonuclease activity is responsible for removing mispaired bases. But the catalytic site of the exonuclease domain is distant from the active site of the catalytic domain. The enzyme alternates between polymerizing and editing modes, as determined by a competition between the two active sites for the 3′ primer end of the DNA (3093). Amino acids in the active site contact the incoming base in such a way that the enzyme structure is affected by a mismatched base. When a mismatched base pair occupies the catalytic site, the fingers cannot rotate toward the palm to bind the incoming nucleotide. This leaves the 3′ end free to bind to the active site in the exonuclease domain, which is accomplished by a rotation of the DNA in the enzyme structure (3094).

© 2004. Virtual Text / www.ergito.com

Last updated on 11-20-2002

© 2004. Virtual Text / www.ergito.com

Reviews

3091. Joyce, C. M. and Steitz, T. A. (1994). *Function and structure relationships in DNA polymerases.* Annu. Rev. Biochem. 63, 777-822.

3092. Hubscher, U., Maga, G., and Spadari, S. (2002). *Eukaryotic DNA polymerases.* Annu. Rev. Biochem. 71, 133-163.

3093. Johnson, K. A. (1993). *Conformational coupling in DNA polymerase fidelity.* Annu. Rev. Biochem. 62, 685-713.

References

3094. Shamoo, Y. and Steitz, T. A. (1999). *Building a replisome from interacting pieces: sliding clamp complexed to a peptide from DNA polymerase and a polymerase editing complex.* Cell 99, 155-166.

This content is available online at http://www.ergito.com/main.jsp?bcs=MBIO.4.14.5

© 2004. Virtual Text / www.ergito.com

DNA REPLICATION

4.14.6 DNA synthesis is semidiscontinuous

Key Terms

The **leading strand** of DNA is synthesized continuously in the 5′-3′ direction.

The **lagging strand** of DNA must grow overall in the 3′-5′ direction and is synthesized discontinuously in the form of short fragments (5′-3′) that are later connected covalently.

Okazaki fragments are the short stretches of 1000-2000 bases produced during discontinuous replication; they are later joined into a covalently intact strand.

Semidiscontinuous replication is mode in which one new strand is synthesized continuously while the other is synthesized discontinuously.

Key Concepts

- The DNA replicase advances continuously when it synthesizes the leading strand (5′–3′), but synthesizes the lagging strand by making short fragments that are subsequently joined together.

The antiparallel structure of the two strands of duplex DNA poses a problem for replication. As the replication fork advances, daughter strands must be synthesized on both of the exposed parental single strands. The fork moves in the direction from 5′–3′ on one strand, and in the direction from 3′–5′ on the other strand. Yet nucleic acids are synthesized only from a 5′ end toward a 3′ end. The problem is solved by synthesizing the strand that grows overall from 3′–5′ in a series of short fragments, each actually synthesized in the "backwards" direction, that is, with the customary 5′–3′ polarity.

Consider the region immediately behind the replication fork, as illustrated in **Figure 14.9**. We describe events in terms of the different properties of each of the newly synthesized strands:

© 2004. Virtual Text / www.ergito.com

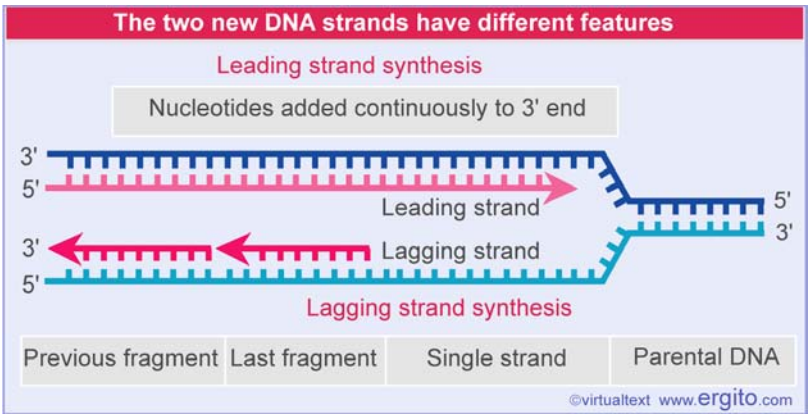


Figure 14.9 The leading strand is synthesized continuously while the lagging strand is synthesized discontinuously.

- On the **leading strand** DNA synthesis can proceed continuously in the 5′ to 3′ direction as the parental duplex is unwound.
- On the **lagging strand** a stretch of single-stranded parental DNA must be exposed, and then a segment is synthesized in the reverse direction (relative to fork movement). A series of these fragments are synthesized, each 5′–3′; then they are joined together to create an intact lagging strand.

Discontinuous replication can be followed by the fate of a very brief label of radioactivity. The label enters newly synthesized DNA in the form of short fragments, sedimenting in the range of 7-11S, corresponding to ~1000-2000 bases in length. These **Okazaki fragments** are found in replicating DNA in both prokaryotes and eukaryotes. After longer periods of incubation, the label enters larger segments of DNA. The transition results from covalent linkages between Okazaki fragments.

(The lagging strand *must* be synthesized in the form of Okazaki fragments. For a long time it was unclear whether the leading strand is synthesized in the same way or is synthesized continuously. All newly synthesized DNA is found as short fragments in *E. coli*. Superficially, this suggests that both strands are synthesized discontinuously. However, it turns out that not all of the fragment population represents *bona fide* Okazaki fragments; some are pseudofragments, generated by breakage in a DNA strand that actually was synthesized as a continuous chain. The source of this breakage is the incorporation of some uracil into DNA in place of thymine. When the uracil is removed by a repair system, the leading strand has breaks until a thymine is inserted.)

So the lagging strand is synthesized discontinuously and the leading strand is synthesized continuously. This is called **semidiscontinuous replication**.

This content is available online at http://www.ergito.com/main.jsp?bcs=MBIO.4.14.6

© 2004. Virtual Text / www.ergito.com

DNA REPLICATION

4.14.7 The ϕX model system shows how single-stranded DNA is generated for replication

Key Terms

A **helicase** is an enzyme that uses energy provided by ATP hydrolysis to separate the strands of a nucleic acid duplex.

The **single-strand binding protein (SSB)** attaches to single-stranded DNA, thereby preventing the DNA from forming a duplex.

Key Concepts

- Replication requires a helicase to separate the strands of DNA using energy provided by hydrolysis of ATP.
- A single-strand binding protein is required to maintain the separated strands.
- The combination of helicase, SSB, and A protein separates a ϕX174 duplex into a single-stranded circle and a single-stranded linear strand.

As the replication fork advances, it unwinds the duplex DNA. One of the template strands is rapidly converted to duplex DNA as the leading daughter strand is synthesized. The other remains single-stranded until a sufficient length has been exposed to initiate synthesis of an Okazaki fragment of the lagging strand in the backward direction. The generation and maintenance of single-stranded DNA is therefore a crucial aspect of replication. Two types of function are needed to convert double-stranded DNA to the single-stranded state:

- A **helicase** is an enzyme that separates the strands of DNA, usually using the hydrolysis of ATP to provide the necessary energy.
- A **single-strand binding protein** (**SSB**) binds to the single-stranded DNA, preventing it from reforming the duplex state. The SSB binds as a monomer, but typically in a cooperative manner in which the binding of additional monomers to the existing complex is enhanced.

Helicases separate the strands of a duplex nucleic acid in a variety of situations, ranging from strand separation at the growing point of a replication fork to catalyzing migration of Holliday (recombination) junctions along DNA. There are 12 different helicases in *E. coli*. A helicase is generally multimeric. A common form of helicase is a hexamer (1195). This typically translocates along DNA by using its multimeric structure to provide multiple DNA-binding sites.

Figure 14.10 shows a generalized schematic model for the action of a hexameric helicase. It is likely to have one conformation that binds to duplex DNA and another that binds to single-stranded DNA. Alternation between them drives the motor that

© 2004. Virtual Text / www.ergito.com

melts the duplex, and requires ATP hydrolysis – typically 1 ATP is hydrolyzed for each base pair that is unwound (1194). A helicase usually initiates unwinding at a single-stranded region adjacent to a duplex, and may function with a particular polarity, preferring single-stranded DNA with a 3′ end (3′–5′ helicase) or with a 5′ end (5′–3′ helicase).

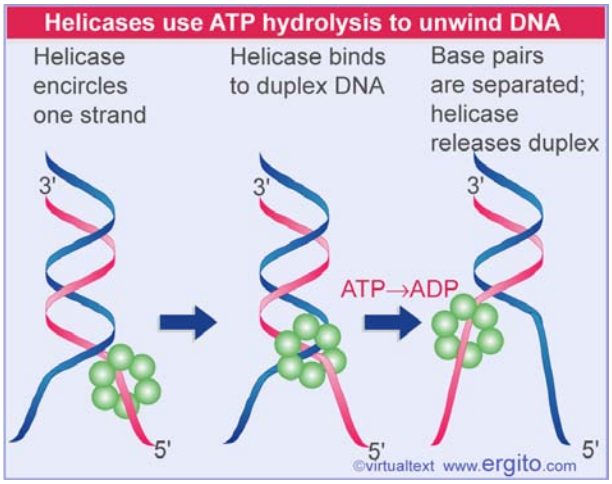


Figure 14.10 A hexameric helicase moves along one strand of DNA. It probably changes conformation when it binds to the duplex, uses ATP hydrolysis to separate the strands, and then returns to the conformation it has when bound only to a single strand.

The conversion of ϕX174 double-stranded DNA into individual single strands illustrates the features of the strand separation process. **Figure 14.11** shows that a single strand is peeled off the circular strand, resembling the rolling circle described previously in **Figure 13.16**. The reaction can occur in the absence of DNA synthesis when the appropriate 3 proteins are provided *in vitro*.

© 2004. Virtual Text / www.ergito.com

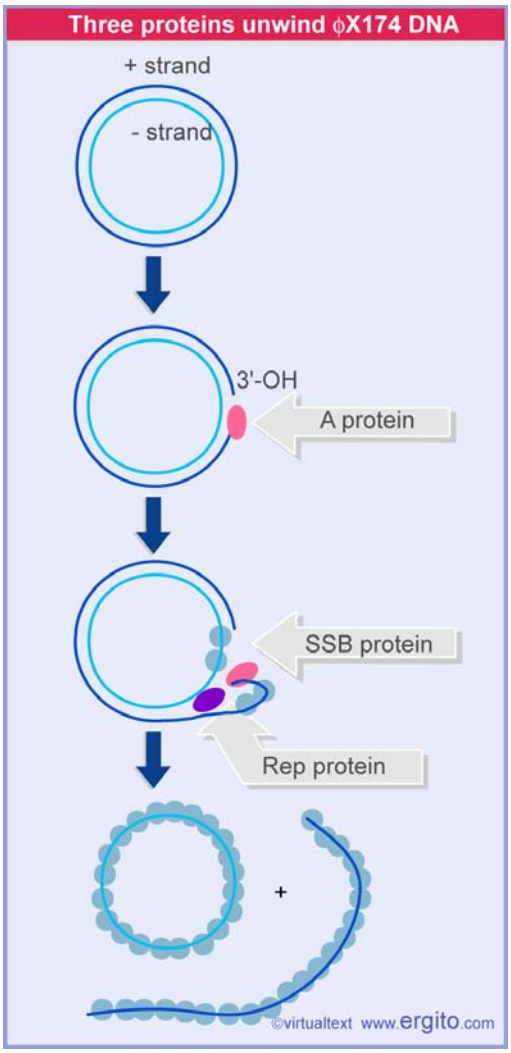


Figure 14.11 ϕ X174 DNA can be separated into single strands by the combined effects of 3 functions: nicking with A protein, unwinding by Rep, and single-strand stabilization by SSB.

The phage A protein nicks the viral (+) strand at the origin of replication. In the presence of 2 host proteins, Rep and SSB, and ATP, the nicked DNA unwinds. The Rep protein provides a helicase that separates the strands; the SSB traps them in single-stranded form. The *E. coli* SSB is a tetramer of 74 kD that binds cooperatively to single-stranded DNA.

The significance of the cooperative mode of binding is that the binding of one protein molecule makes it much easier for another to bind. So once the binding reaction has started on a particular DNA molecule, it is rapidly extended until all of the single-stranded DNA is covered with the SSB protein. Note that this protein is not a DNA-unwinding protein; its function is to stabilize DNA that is already in the single-stranded condition.

Under normal circumstances *in vivo*, the unwinding, coating, and replication

© 2004. Virtual Text / www.ergito.com

reactions proceed in tandem. The SSB binds to DNA as the replication fork advances, keeping the two parental strands separate so that they are in the appropriate condition to act as templates. SSB is needed in stoichiometric amounts at the replication fork. It is required for more than one stage of replication; *ssb* mutants have a quick-stop phenotype, and are defective in repair and recombination as well as in replication. (Some phages use different SSB proteins, notably T4; this shows that there may be specific interactions between components of the replication apparatus and the SSB; see *Molecular Biology 4.14.14 Phage T4 provides its own replication apparatus*).

© 2004. Virtual Text / www.ergito.com

References

1194. Dillingham, M. S., Wigley, D. B., and Webb, M. R. (2000). *Demonstration of unidirectional single-stranded DNA translocation by PcrA helicase: measurement of step size and translocation speed.* Biochemistry 39, 205-212.

1195. Singleton, M. R., Sawaya, M. R., Ellenberger, T., and Wigley, D. B. (2000). *Crystal structure of T7 gene 4 ring helicase indicates a mechanism for sequential hydrolysis of nucleotides.* Cell 101, 589-600.

This content is available online at http://www.ergito.com/main.jsp?bcs=MBIO.4.14.7

© 2004. Virtual Text / www.ergito.com

DNA REPLICATION

4.14.8 Priming is required to start DNA synthesis

Key Terms

A **primer** is a short sequence (often of RNA) that is paired with one strand of DNA and provides a free 3′-OH end at which a DNA polymerase starts synthesis of a deoxyribonucleotide chain.

The **primase** is a type of RNA polymerase that synthesizes short segments of RNA that will be used as primers for DNA replication.

Key Concepts

- All DNA polymerases require a 3′–OH priming end to initiate DNA synthesis.
- The priming end can be provided by an RNA primer, a nick in DNA, or a priming protein.
- For DNA replication, a special RNA polymerase called a primase synthesizes an RNA chain that provides the priming end.
- *E. coli* has two types of priming reaction, which occur at the bacterial origin (*oriC*) and the ϕX174 origin.
- Priming of replication on double-stranded DNA always requires a replicase, SSB, and primase.
- DnaB is the helicase that unwinds DNA for replication in *E. coli*.

A common feature of all DNA polymerases is that they cannot initiate synthesis of a chain of DNA *de novo*. **Figure 14.12** shows the features required for initiation. Synthesis of the new strand can only start from a pre-existing 3′–OH end; and the template strand must be converted to a single-stranded condition.

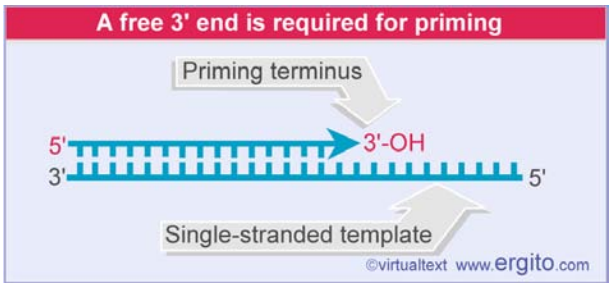


Figure 14.12 A DNA polymerase requires a 3′-OH end to initiate replication.

The 3′–OH end is called a **primer**. The primer can take various forms. Types of priming reaction are summarized in **Figure 14.13**:

© 2004. Virtual Text / www.ergito.com

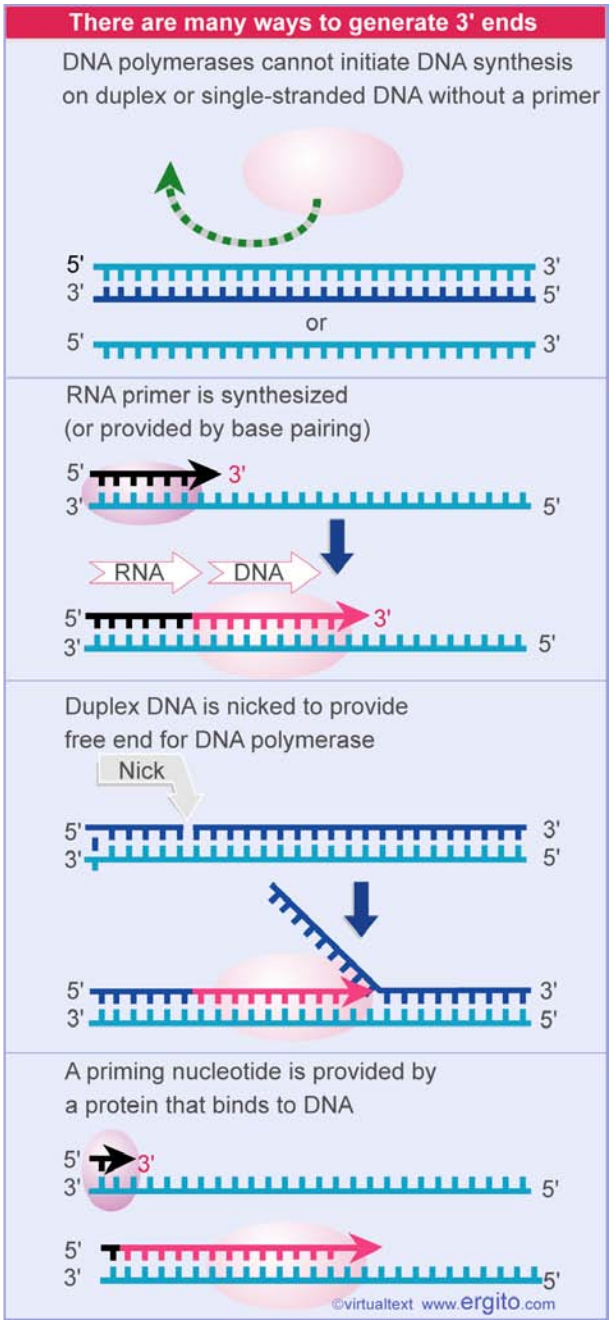


Figure 14.13 There are several methods for providing the free 3′-OH end that DNA polymerases require to initiate DNA synthesis.

- A sequence of RNA is synthesized on the template, so that the free 3′–OH end of the RNA chain is extended by the DNA polymerase. This is commonly used in replication of cellular DNA, and by some viruses (see **Figure 13.40** in *Molecular Biology 4.13.23 The ColE1 compatibility system is controlled by an RNA regulator*).

© 2004. Virtual Text / www.ergito.com

- A preformed RNA pairs with the template, allowing its 3′–OH end to be used to prime DNA synthesis. This mechanism is used by retroviruses to prime reverse transcription of RNA (see **Figure 17.6** in *Molecular Biology 4.17.4 Viral DNA is generated by reverse transcription*).

- A primer terminus is generated within duplex DNA. The most common mechanism is the introduction of a nick, as used to initiate rolling circle replication (see **Figure 13.16**). In this case, the pre-existing strand is displaced by new synthesis. (Note the difference from nick translation shown in **Figure 14.5**, in which DNA polymerase I simultaneously synthesizes and degrades DNA from a nick.)

- A protein primes the reaction directly by presenting a nucleotide to the DNA polymerase. This reaction is used by certain viruses (see **Figure 13.15** in *Molecular Biology 4.13.8 The ends of linear DNA are a problem for replication*).

Priming activity is required to provide 3′–OH ends to start off the DNA chains on both the leading and lagging strands. The leading strand requires only one such initiation event, which occurs at the origin. But there must be a series of initiation events on the lagging strand, since each Okazaki fragment requires its own start *de novo*. Each Okazaki fragment starts with a primer sequence of RNA, ~10 bases long, that provides the 3′–OH end for extension by DNA polymerase.

A **primase** is required to catalyze the actual priming reaction. This is provided by a special RNA polymerase activity, the product of the *dnaG* gene. The enzyme is a single polypeptide of 60 kD (much smaller than RNA polymerase). The primase is an RNA polymerase that is used only under specific circumstances, that is, to synthesize short stretches of RNA that are used as primers for DNA synthesis. DnaG primase associates transiently with the replication complex, and typically synthesizes an 11-12 base primer. Primers start with the sequence pppAG, opposite the sequence 3′–GTC-5′ in the template.

(Some systems use alternatives to the DnaG primase. In the examples of the two phages M13 and G4, which were used for early work on replication, an interesting difference emerged. G4 priming uses DnaG, but M13 priming uses bacterial RNA polymerase. These phages have another unusual feature, which is that the site of priming is indicated by a region of secondary structure.)

There are two types of priming reaction in *E. coli*.

- The *oriC* system, named for the bacterial origin, basically involves the association of the DnaG primase with the protein complex at the replication fork.

- The ϕX system, named for phage ϕX174, requires an initiation complex consisting of additional components, called the primosome (see *Molecular Biology 4.14.17 The primosome is needed to restart replication*).

Sometimes replicons are referred to as being of the ϕX or *oriC* type.

© 2004. Virtual Text / www.ergito.com

The types of activities involved in the initiation reaction are summarized in **Figure 14.14**. Although other replicons in *E. coli* may have alternatives for some of these particular proteins, the same general types of activity are required in every case. A helicase is required to generate single strands, a single-strand binding protein is required to maintain the single-stranded state, and the primase synthesizes the RNA primer.

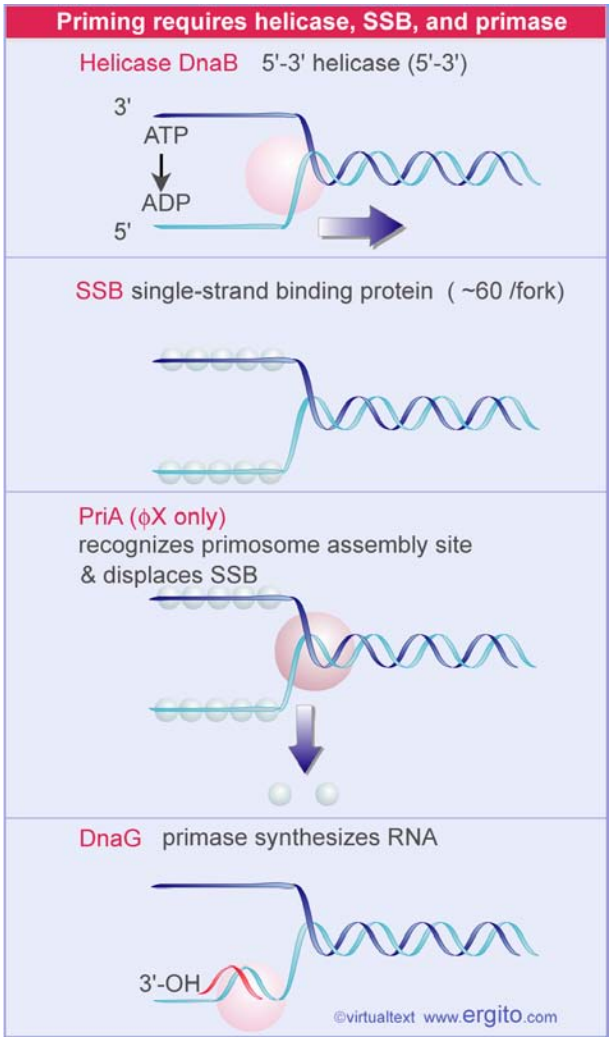


Figure 14.14 Initiation requires several enzymatic activities, including helicases, single-strand binding proteins, and synthesis of the primer.

DnaB is the central component in both ϕX and *oriC* replicons. It provides the 5′–3′ helicase activity that unwinds DNA. Energy for the reaction is provided by cleavage of ATP. Basically DnaB is the active component of the growing point. In *oriC* replicons, DnaB is initially loaded at the origin as part of a large complex (see *Molecular Biology 4.14.15 Creating the replication forks at an origin*). It forms the growing point at which the DNA strands are separated as the replication fork advances. It is part of the DNA polymerase complex and interacts with the DnaG primase to initiate synthesis of each Okazaki fragment on the lagging strand.

© 2004. Virtual Text / www.ergito.com

Last updated on 4-18-2000

This content is available online at http://www.ergito.com/main.jsp?bcs=MBIO.4.14.8

© 2004. Virtual Text / www.ergito.com

DNA REPLICATION

4.14.9 Coordinating synthesis of the lagging and leading strands

Key Concepts

- Different enzyme units are required to synthesize the leading and lagging strands.
- In *E. coli* both these units contain the same catalytic subunit (DnaE).
- In other organisms, different catalytic subunits may be required for each strand.

Each new DNA strand is synthesized by an individual catalytic unit. **Figure 14.15** shows that the behavior of these two units is different because the new DNA strands are growing in opposite directions. One enzyme unit is moving with the unwinding point and synthesizing the leading strand continuously. The other unit is moving "backwards," relative to the DNA, along the exposed single strand. Only short segments of template are exposed at any one time. When synthesis of one Okazaki fragment is completed, synthesis of the next Okazaki fragment is required to start at a new location approximately in the vicinity of the growing point for the leading strand. This requires a translocation relative to the DNA of the enzyme unit that is synthesizing the lagging strand.

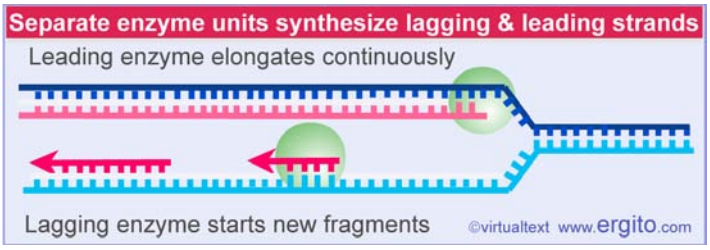


Figure 14.15 Leading and lagging strand polymerases move apart.
This is a static version of an interactive figure; see http://www.ergito.com/main.jsp?bcs=MBIO.4.14.9 to view properly.

The term "enzyme unit" avoids the issue of whether the DNA polymerase that synthesizes the leading strand is the same type of enzyme as the DNA polymerase that synthesizes the lagging strand. In the case that we know best, *E. coli*, there is only a single type of DNA polymerase catalytic subunit used in replication, the DnaE protein. The active replicase is a dimer, and each half of the dimer contains DnaE as the catalytic subunit, supported by other proteins (which differ between the leading and lagging strands).

The use of a single type of catalytic subunit, however, may be atypical. In the bacterium *B. subtilis*, there are two different catalytic subunits (2184). PolC is the homologue to *E. coli*'s DnaE, and is responsible for synthesizing the leading strand. A related protein, $DnaE_{BS}$, is the catalytic subunit that synthesizes the lagging strand. Eukaryotic DNA polymerases have the same general structure, with different enzyme

© 2004. Virtual Text / www.ergito.com

units synthesizing the leading and lagging strands, but it is not clear whether the same or different types of catalytic subunits are used (see *Molecular Biology 4.14.13 Separate eukaryotic DNA polymerases undertake initiation and elongation*).

A major problem of the semidiscontinuous mode of replication follows from the use of different enzyme units to synthesize each new DNA strand: how is synthesis of the lagging strand coordinated with synthesis of the leading strand? As the replisome moves along DNA, unwinding the parental strands, one enzyme unit elongates the leading strand. Periodically the primosome activity initiates an Okazaki fragment on the lagging strand, and the other enzyme unit must then move in the reverse direction to synthesize DNA. **Figure 14.16** proposes two types of model for what happens to this enzyme unit when it completes synthesis of an Okazaki fragment. The same complex may be reutilized for synthesis of successive Okazaki fragments. Or the complex might dissociate from the template, so that a new complex must be assembled to elongate the next Okazaki fragment. We see in the *Molecular Biology 4.14.11 The clamp controls association of core enzyme with DNA* that the first model applies.

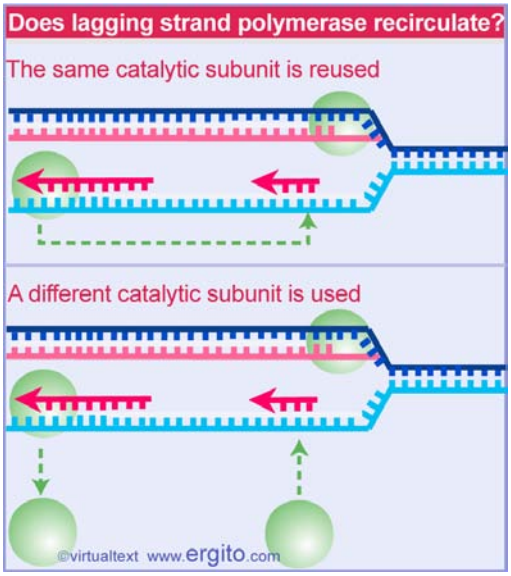


Figure 14.16 The upper model for the action of lagging strand polymerase is that when an enzyme unit completes one Okazaki fragment, it moves to a new position to synthesize the next fragment. The lower model is that the lagging strand polymerase dissociates when it completes an Okazaki fragment, and a new enzyme unit associates with DNA to synthesize the next Okazaki fragment.

Last updated on 12-6-2001

© 2004. Virtual Text / www.ergito.com

References

2184. Dervyn, E., Suski, C., Daniel, R., Bruand, C., Chapuis, J., Errington, J., Janniere, L., and Ehrlich, S. D. (2001). *Two essential DNA polymerases at the bacterial replication fork.* Science 294, 1716-1719.

This content is available online at http://www.ergito.com/main.jsp?bcs=MBIO.4.14.9

© 2004. Virtual Text / www.ergito.com

DNA REPLICATION

4.14.10 DNA polymerase holoenzyme has 3 subcomplexes

Key Terms

The **clamp loader** is a 5 subunit protein complex which is responsible for loading the β clamp on to DNA at the replication fork.

Key Concepts

- The *E. coli* replicase DNA polymerase III is a 900 kD complex with a dimeric structure.
- Each monomeric unit has a catalytic core, a dimerization subunit, and a processivity component.
- A clamp loader places the processivity subunits on DNA, and they form a circular clamp around the nucleic acid.
- One catalytic core is associated with each template strand.

We can now relate the subunit structure of *E. coli* DNA polymerase III to the activities required for DNA synthesis and propose a model for its action. The holoenzyme is a complex of 900 kD that contains 10 proteins organized into four types of subcomplex:

- There are two copies of the catalytic core. Each catalytic core contains the α subunit (the DNA polymerase activity), ε subunit (3′–5′ proofreading exonuclease), and θ subunit (stimulates exonuclease).
- There are two copies of the dimerizing subunit, τ, which link the two catalytic cores together (2185).
- There are two copies of the *clamp*, which is responsible for holding catalytic cores on to their template strands. Each clamp consists of a homodimer of β subunits that binds around the DNA and ensures processivity.
- The γ complex is a group of 5 proteins, the **Clamp loader**, that places the clamp on DNA (2185).

A model for the assembly of DNA polymerase III is shown in **Figure 14.17**. The holoenzyme assembles on DNA in three stages:

© 2004. Virtual Text / www.ergito.com

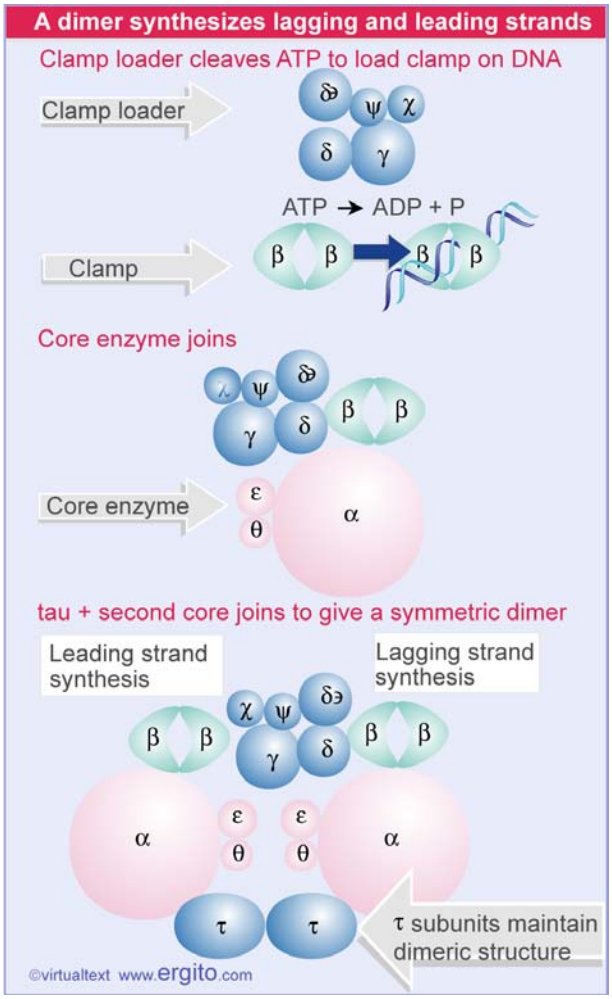


Figure 14.17 DNA polymerase III holoenzyme assembles in stages, generating an enzyme complex that synthesizes the DNA of both new strands.

- First the clamp loader uses hydrolysis of ATP to bind β subunits to a template-primer complex.
- Binding to DNA changes the conformation of the site on β that binds to the clamp loader, and as a result it now has a high affinity for the core polymerase. This enables core polymerase to bind, and this is the means by which the core polymerase is brought to DNA.
- A τ dimer binds to the core polymerase, and provides a dimerization function that binds a second core polymerase (associated with another β clamp). The holoenzyme is asymmetric, because it has only 1 clamp loader. The clamp loader is responsible for adding a pair of β dimers to each parental strand of DNA.

Each of the core complexes of the holoenzyme synthesizes one of the new strands of DNA. Because the clamp loader is also needed for unloading the β complex from DNA, the two cores have different abilities to dissociate from DNA. This corresponds to the need to synthesize a continuous leading strand (where polymerase

© 2004. Virtual Text / www.ergito.com

remains associated with the template) and a discontinuous lagging strand (where polymerase repetitively dissociates and reassociates). The clamp loader is associated with the core polymerase that synthesizes the lagging strand, and plays a key role in the ability to synthesize individual Okazaki fragments (2186).

Last updated on 5-14-2002

© 2004. Virtual Text / www.ergito.com

References

2185. Studwell-Vaughan, P. S. and O'Donnell, M. (1991). *Constitution of the twin polymerase of DNA polymerase III holoenzyme.* J. Biol. Chem. 266, 19833-19841.

2186. Stukenberg, P. T., Studwell-Vaughan, P. S., and O'Donnell, M. (1991). *Mechanism of the sliding beta-clamp of DNA polymerase III holoenzyme.* J. Biol. Chem. 266, 11328-11334.

This content is available online at http://www.ergito.com/main.jsp?bcs=MBIO.4.14.10

© 2004. Virtual Text / www.ergito.com

DNA REPLICATION

4.14.11 The clamp controls association of core enzyme with DNA

Key Concepts

- The core on the leading strand is processive because its clamp keeps it on the DNA.
- The clamp associated with the core on the lagging strand dissociates at the end of each Okazaki fragment and reassembles for the next fragment.
- The helicase DnaB is responsible for interacting with the primase DnaG to initiate each Okazaki fragment.

The β dimer makes the holoenzyme highly processive. β is strongly bound to DNA, but can slide along a duplex molecule. The crystal structure of β shows that it forms a ring-shaped dimer (2187). The model in **Figure 14.18** shows the β-ring in relationship to a DNA double helix. The ring has an external diameter of 80 Å and an internal cavity of 35 Å, almost twice the diameter of the DNA double helix (20 Å). The space between the protein ring and the DNA is filled by water. Each of the β subunits has three globular domains with similar organization (although their sequences are different). As a result, the dimer has 6-fold symmetry, reflected in 12 α-helices that line the inside of the ring.

© 2004. Virtual Text / www.ergito.com

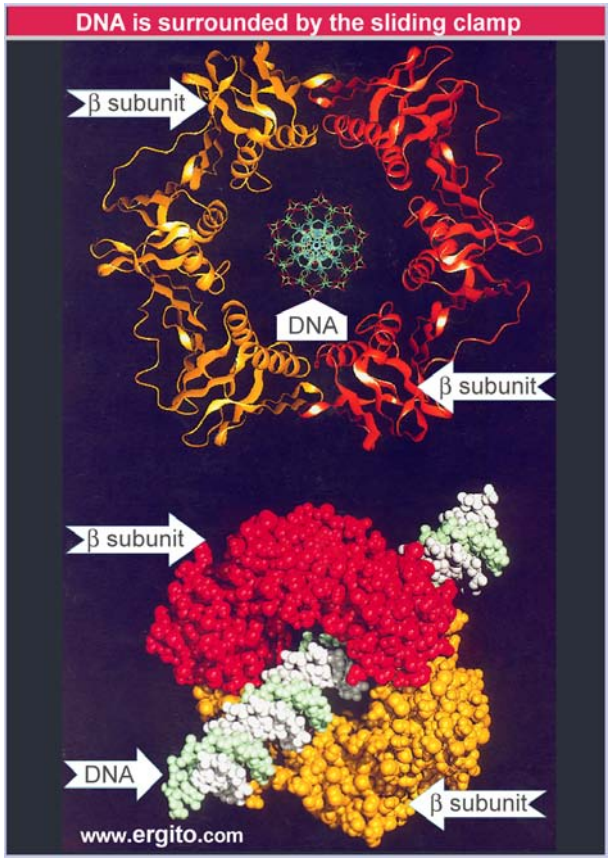


Figure 14.18 The β subunit of DNA polymerase III holoenzyme consists of a head to tail dimer (the two subunits are shown in red and orange) that forms a ring completely surrounding a DNA duplex (shown in the center). Photograph kindly provided by John Kuriyan.

The dimer surrounds the duplex, providing the "sliding clamp" that allows the holoenzyme to slide along DNA. The structure explains the high processivity – there is no way for the enzyme to fall off! The α-helices on the inside have some positive charges that may interact with the DNA via the intermediate water molecules. Because the protein clamp does not directly contact the DNA, it may be able to "ice-skate" along the DNA, making and breaking contacts via the water molecules.

How does the clamp get on to the DNA? Because the clamp is a circle of subunits surrounding DNA, its assembly or removal requires the use of an energy-dependent process by the clamp loader. The γ clamp loader is a pentameric circular structure that binds an open form of the β ring preparatory to loading it on to DNA (2188). In effect, the ring is opened at one of the interfaces between the two β subunits by the δ subunit of the clamp loader. The clamp loader uses hydrolysis of ATP to provide the energy to open the ring of the clamp and insert DNA into its central cavity.

The relationship between the β clamp and the γ clamp loader is a paradigm for similar systems used by DNA replicases ranging from bacteriophages to animal cells. The clamp is a heteromer (sometimes a dimer, sometimes a trimer) that forms a ring around DNA with a set of 12 α-helices forming 6-fold symmetry for the structure as

© 2004. Virtual Text / www.ergito.com

a whole. The clamp loader has some subunits that hydrolyze ATP to provide energy for the reaction (for review see 3441).

The basic principle that is established by the dimeric polymerase model is that, while one polymerase subunit synthesizes the leading strand continuously, the other cyclically initiates and terminates the Okazaki fragments of the lagging strand within a large single-stranded loop formed by its template strand. **Figure 14.19** draws a generic model for the operation of such a replicase. The replication fork is created by a helicase, typically forming a hexameric ring, that translocates in the 5′–3′ direction on the template for the lagging strand. The helicase is connected to two DNA polymerase catalytic subunits, each of which is associated with a sliding clamp.

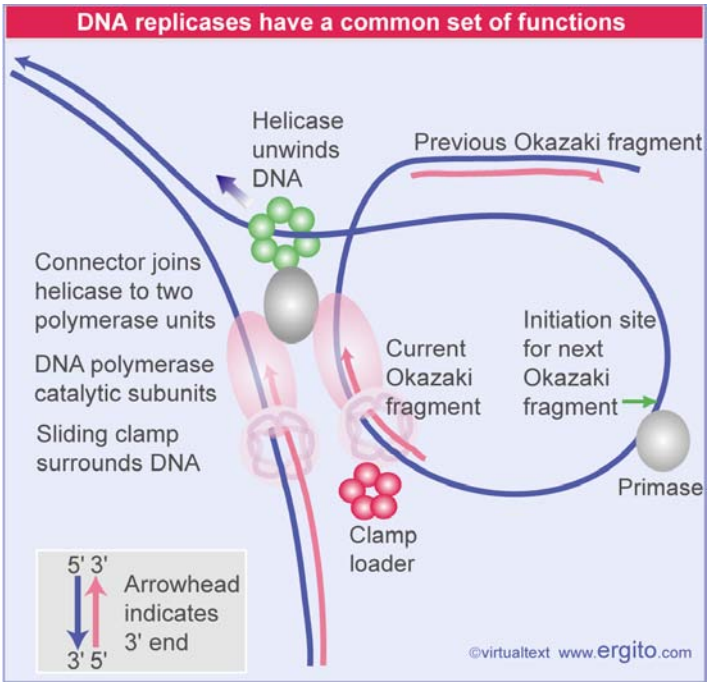


Figure 14.19 The helicase creating the replication fork is connected to two DNA polymerase catalytic subunits, each of which is held on to DNA by a sliding clamp. The polymerase that synthesizes the leading strand moves continuously. The polymerase that synthesizes the lagging strand dissociates at the end of an Okazaki fragment and then reassociates with a primer in the single-stranded template loop to synthesize the next fragment.

We can describe this model for DNA polymerase III in terms of the individual components of the enzyme complex, as illustrated in **Figure 14.20**. A catalytic core is associated with each template strand of DNA. The holoenzyme moves continuously along the template for the leading strand; the template for the lagging strand is "pulled through," creating a loop in the DNA. DnaB creates the unwinding point, and translocates along the DNA in the "forward" direction (for review see 2279)

© 2004. Virtual Text / www.ergito.com

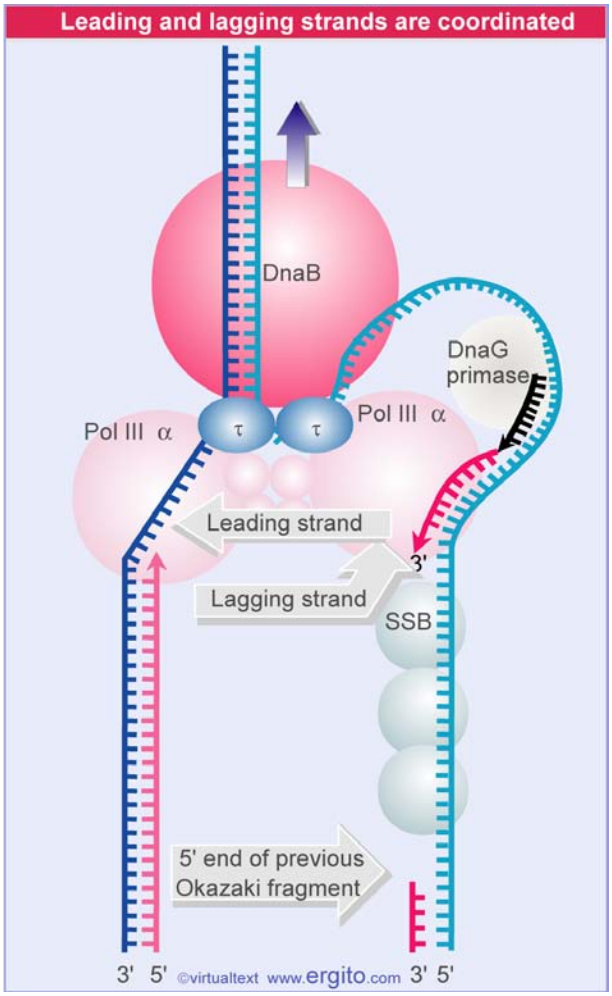


Figure 14.20 Each catalytic core of Pol III synthesizes a daughter strand. DnaB is responsible for forward movement at the replication fork.

DnaB contacts the τ subunit(s) of the clamp loader. This establishes a direct connection between the helicase-primase complex and the catalytic cores. This link has two effects. One is to increase the speed of DNA synthesis by increasing the rate of movement by DNA polymerase core by 10×. The second is to prevent the leading strand polymerase from falling off, that is, to increase its processivity.

Synthesis of the leading strand creates a loop of single-stranded DNA that provides the template for lagging strand synthesis, and this loop becomes larger as the unwinding point advances. After initiation of an Okazaki fragment, the lagging strand core complex pulls the single-stranded template through the β clamp while synthesizing the new strand. The single-stranded template must extend for the length of at least one Okazaki fragment before the lagging polymerase completes one fragment and is ready to begin the next.

What happens to the loop when the Okazaki fragment is completed? **Figure 14.21** suggests that the core complex dissociates when it completes synthesis of each fragment, releasing the loop. The core complex then associates with a β clamp to

© 2004. Virtual Text / www.ergito.com

initiate the next Okazaki fragment. Probably a new β clamp will already be present at the next initiation site, and the β clamp that has lost its core complex will dissociate from the template (with the assistance of the clamp loader complex) to be used again. So the lagging strand polymerase will probably transfer from one β clamp to the next in each cycle, without dissociating from the replicating complex.

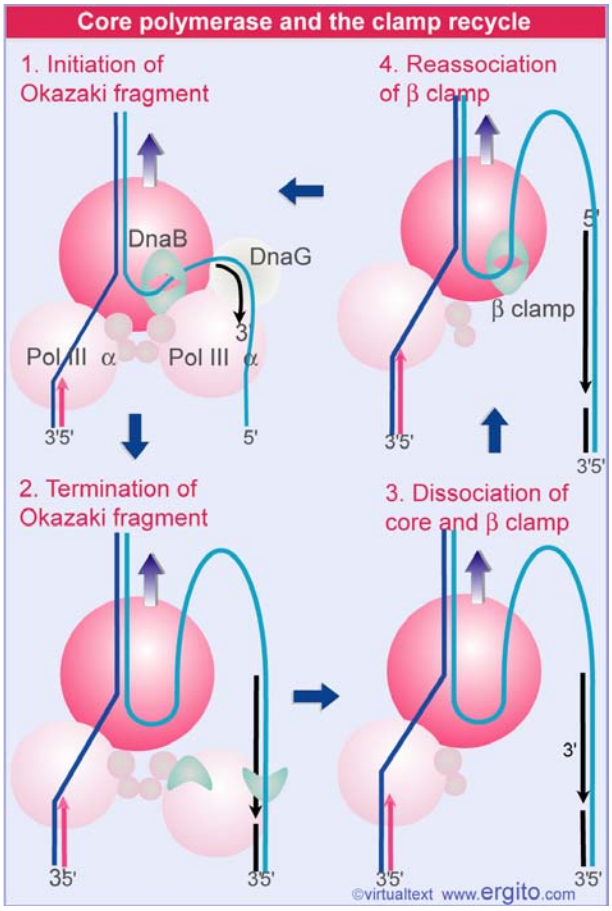


Figure 14.21 Core polymerase and the β clamp dissociate at completion of Okazaki fragment synthesis and reassociate at the beginning.

What is responsible for recognizing the sites for initiating synthesis of Okazaki fragments? In *oriC* replicons, the connection between priming and the replication fork is provided by the dual properties of DnaB: it is the helicase that propels the replication fork; and it interacts with the DnaG primase at an appropriate site. Following primer synthesis, the primase is released. The length of the priming RNA is limited to 8-14 bases. Apparently DNA polymerase III is responsible for displacing the primase.

Last updated on 11-20-2002

© 2004. Virtual Text / www.ergito.com

Reviews

2279. Benkovic, S. J., Valentine, A. M., and Salinas, F. (2001). *Replisome-mediated DNA replication.* Annu. Rev. Biochem. 70, 181-208.

3441. Davey, M. J., Jeruzalmi, D., Kuriyan, J., and O'Donnell, M. (2002). *Motors and switches: AAA+ machines within the replisome.* Nat. Rev. Mol. Cell Biol. 3, 826-835.

References

2187. Kong, X. P., Onrust, R., O'Donnell, M., and Kuriyan, J. (1992). *Three-dimensional structure of the beta subunit ofE. coli DNA polymerase III holoenzyme: a sliding DNA clamp.* Cell 69, 425-437.

2188. Jeruzalmi, D., O'Donnell, M., and Kuriyan, J. (2001). *Crystal structure of the processivity clamp loader gamma (gamma) complex of E. coli DNA polymerase III.* Cell 106, 429-441.

This content is available online at http://www.ergito.com/main.jsp?bcs=MBIO.4.14.11

© 2004. Virtual Text / www.ergito.com

DNA REPLICATION

4.14.12 Okazaki fragments are linked by ligase

Key Terms

DNA ligase makes a bond between an adjacent 3′-OH and 5′-phosphate end where there is a nick in one strand of duplex DNA.

Key Concepts

- Each Okazaki fragment starts with a primer and stops before the next fragment.
- DNA polymerase I removes the primer and replaces it with DNA in an action that resembles nick translation.
- DNA ligase makes the bond that connects the 3′ end of one Okazaki fragment to the 5′ beginning of the next fragment.

We can now expand our view of the actions involved in joining Okazaki fragments, as illustrated in **Figure 14.22**. The complete order of events is uncertain, but must involve synthesis of RNA primer, its extension with DNA, removal of the RNA primer, its replacement by a stretch of DNA, and the covalent linking of adjacent Okazaki fragments.

© 2004. Virtual Text / www.ergito.com

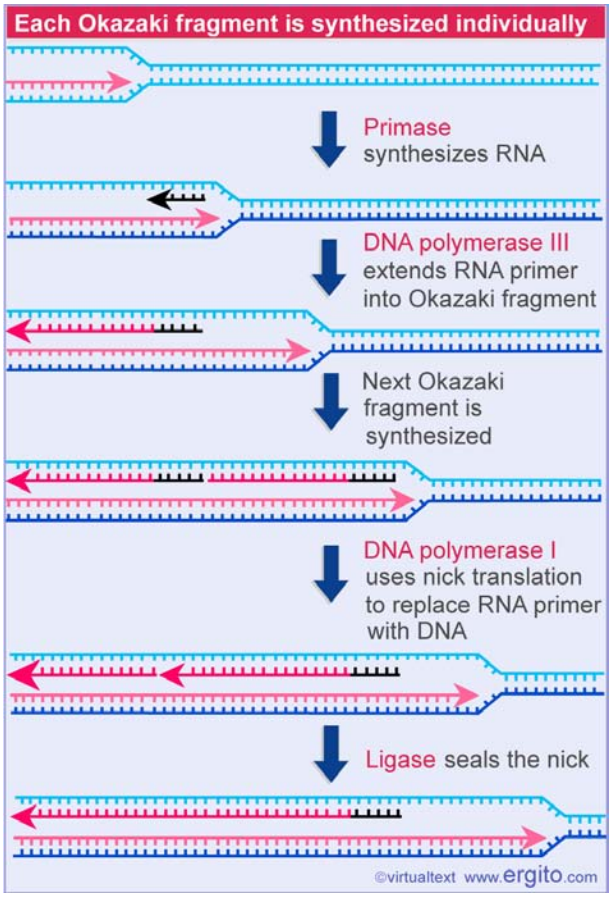


Figure 14.22 Synthesis of Okazaki fragments requires priming, extension, removal of RNA, gap filling, and nick ligation.

The figure suggests that synthesis of an Okazaki fragment terminates just before the start of the RNA primer of the preceding fragment. When the primer is removed, there will be a gap. The gap is filled by DNA polymerase I; *polA* mutants fail to join their Okazaki fragments properly. The 5′–3′ exonuclease activity removes the RNA primer while simultaneously replacing it with a DNA sequence extended from the 3′–OH end of the next Okazaki fragment. This is equivalent to nick translation, except that the new DNA replaces a stretch of RNA rather than a segment of DNA. In mammalian systems (where the DNA polymerase does not have a 5′–3′ exonuclease activity), Okazaki fragments are removed by a two-step process. First RNAase HI (an enzyme that is specific for a DNA-RNA hybrid substrate) makes an endonucleolytic cleavage; then a 5′–3′ exonuclease called FEN1 removes the RNA.

Once the RNA has been removed and replaced, the adjacent Okazaki fragments must be linked together. The 3′–OH end of one fragment is adjacent to the 5′–phosphate end of the previous fragment. The responsibility for sealing this nick lies with the enzyme **DNA ligase**. Ligases are present in both prokaryotes and eukaryotes. Unconnected fragments persist in lig^- mutants, because they fail to join Okazaki fragments together.

The *E. coli* and T4 ligases share the property of sealing nicks that have 3′–OH and 5

© 2004. Virtual Text / www.ergito.com

′ –phosphate termini, as illustrated in **Figure 14.23**. Both enzymes undertake a two-step reaction, involving an enzyme-AMP complex. (The *E. coli* and T4 enzymes use different cofactors. The *E. coli* enzyme uses NAD [nicotinamide adenine dinucleotide] as a cofactor, while the T4 enzyme uses ATP.) The AMP of the enzyme complex becomes attached to the 5′ –phosphate of the nick; and then a phosphodiester bond is formed with the 3′ –OH terminus of the nick, releasing the enzyme and the AMP.

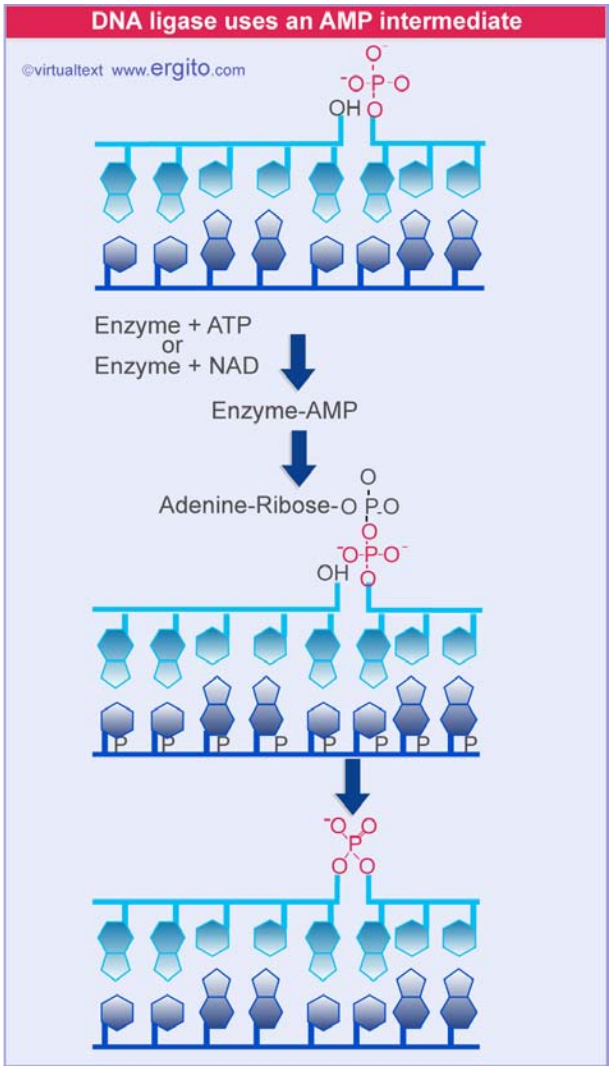


Figure 14.23 DNA ligase seals nicks between adjacent nucleotides by employing an enzyme-AMP intermediate.

This content is available online at http://www.ergito.com/main.jsp?bcs=MBIO.4.14.12

© 2004. Virtual Text / www.ergito.com

DNA REPLICATION

4.14.13 Separate eukaryotic DNA polymerases undertake initiation and elongation

Key Concepts

- A replication fork has 1 complex of DNA polymerase α/primase and 2 complexes of DNA polymerase δ and/or ε.
- The DNA polymerase α/primase complex initiates the synthesis of both DNA strands.
- DNA polymerase δ elongates the leading strand and a second DNA polymerase δ or DNA polymerase elongates the lagging strand

Eukaryotic cells have a large number of DNA polymerases (for review see 3092). They can be broadly divided into those required for semiconservative replication and those involved in synthesizing material to repair damaged DNA. Nuclear DNA replication requires DNA polymerases α, δ, and ε, and mitochondrial replication requires DNA polymerase γ. All the other enzymes are concerned with synthesizing stretches of new DNA to replace damaged material. **Figure 14.24** shows that all of the nuclear replicases are large heterotetrameric enzymes. In each case, one of the subunits has the responsibility for catalysis, and the others are concerned with ancillary functions, such as priming, processivity, or proofreading. These enzymes all replicate DNA with high fidelity, as does the slightly less complex mitochondrial enzyme. The repair polymerases have much simpler structures, often consisting of a single monomeric subunit (although it may function in the context of a complex of other repair enzymes). Of the enzymes involved in repair, only DNA polymerase β has a fidelity approaching the replicases: all of the others have much greater error rates (for review see 3097).

© 2004. Virtual Text / www.ergito.com

DNA polymerases undertake replication or repair		
DNA polymerase	Function	Structure
	High fidelity replicases	
α	Nuclear replication	350 kD tetramer
δ	"	250 kD tetramer
ε	"	350 kD tetramer
γ	Mitochondrial replication	200 kD dimer
	High fidelity repair	
β	Base excision repair	39 kD monomer
	Low fidelity repair	
ζ	Thymine dimer bypass	heteromer
η	Base damage repair	monomer
ι	Required in meiosis	monomer
κ	Deletion and base substitution	monomer

©virtualtext www.ergito.com

Figure 14.24 Eukaryotic cells have many DNA polymerases. The replicative enzymes operate with high fidelity. Except for the β enzyme, the repair enzymes all have low fidelity. Replicative enzymes have large structures, with separate subunits for different activities. Repair enzymes have much simpler structures.

Each of the three nuclear DNA replicases has a different function:

- DNA polymerase α initiates the synthesis of new strands.
- DNA polymerase δ elongates the leading strand.
- DNA polymerase ε may be involved in lagging strand synthesis, but also has other roles.

DNA polymerase α is unusual because it has the ability to initiate a new strand. It is used to initiate both the leading and lagging strands. The enzyme exists as a complex consisting of a 180 kD catalytic subunit, associated with the B subunit that appears necessary for assembly, and two smaller proteins that provide a primase activity. Reflecting its dual capacity to prime and extend chains, it is sometimes called pol α/primase.

The pol α/primase enzyme binds to the initiation complex at the origin and synthesizes a short strand consisting of ~10 bases of RNA followed by 20-30 bases of DNA (sometimes called iDNA). Then it is replaced by an enzyme that will extend the chain. On the leading strand, this is DNA polymerase δ. This event is called the pol switch. It involves interactions among several components of the initiation complex.

© 2004. Virtual Text / www.ergito.com

DNA polymerase δ is a highly processive enzyme that continuously synthesizes the leading strand. Its processivity results from its interaction with two other proteins, RF-C and PCNA.

The roles of RF-C and PCNA are analogous to the *E. coli* γ clamp loader and β processivity unit (see *Molecular Biology 4.14.11 The clamp controls association of core enzyme with DNA*). RF-C is a clamp loader that catalyzes the loading of PCNA on to DNA. It binds to the 3′ end of the iDNA and uses ATP-hydrolysis to open the ring of PCNA so that it can encircle the DNA (3095). The processivity of DNA polymerase δ is maintained by PCNA, which tethers DNA polymerase δ to the template. (PCNA is called proliferating cell nuclear antigen for historical reasons.) The crystal structure of PCNA closely resembles the *E. coli* β subunit: a trimer forms a ring that surrounds the DNA. Although the sequence and subunit organization are different from the dimeric β clamp, the function is likely to be similar.

We are less certain about events on the lagging strand. One possibility is that DNA polymerase δ also elongates the lagging strand. It has the capability to dimerize, which suggests a model analogous to the behavior of *E. coli* replicase (see *Molecular Biology 4.14.10 DNA polymerase holoenzyme has 3 subcomplexes*) (3096). However, there are some indications that DNA polymerase ε may elongate the lagging strand (2189; 2208), although it also has been identified with other roles.

A general model suggests that a replication fork contains 1 complex of DNA polymerase α/primase and two other DNA polymerase complexes. One is DNA polymerase δ and the other is either a second DNA polymerase δ or may possibly be a DNA polymerase ε. The two complexes of DNA polymerase δ/ ε behave in the same way as the two complexes of DNA polymerase III in the *E. coli* replisome: one synthesizes the leading strand, and the other synthesizes Okazaki fragments on the lagging strand. The exonuclease MF1 removes the RNA primers of Okazaki fragments. The enzyme DNA ligase I is specifically required to seal the nicks between the completed Okazaki fragments.

Last updated on 11-20-2002

© 2004. Virtual Text / www.ergito.com

Reviews

3092. Hubscher, U., Maga, G., and Spadari, S. (2002). *Eukaryotic DNA polymerases.* Annu. Rev. Biochem. 71, 133-163.

3097. Goodman, M. F. (2002). *Error-prone repair DNA polymerases in prokaryotes and eukaryotes.* Annu. Rev. Biochem. 71, 17-50.

References

2189. Karthikeyan, R., Vonarx, E. J., Straffon, A. F., Simon, M., Faye, G., and Kunz, B. A. (2000). *Evidence from mutational specificity studies that yeast DNA polymerases delta and epsilon replicate different DNA strands at an intracellular replication fork.* J. Mol. Biol. 299, 405-419.

2208. Waga, S., Masuda, T., Takisawa, H., and Sugino, A. (2001). *DNA polymerase epsilon is required for coordinated and efficient chromosomal DNA replication in Xenopus egg extracts.* Proc. Natl. Acad. Sci. USA 98, 4978-4983.

3095. Shiomi, Y., Usukura, J., Masamura, Y., Takeyasu, K., Nakayama, Y., Obuse, C., Yoshikawa, H., and Tsurimoto, T. (2000). *ATP-dependent structural change of the eukaryotic clamp-loader protein, replication factor C.* Proc. Natl. Acad. Sci. USA 97, 14127-14132.

3096. Zuo, S., Bermudez, V., Zhang, G., Kelman, Z., and Hurwitz, J. (2000). *Structure and activity associated with multiple forms of S. pombe DNA polymerase delta.* J. Biol. Chem. 275, 5153-5162.

This content is available online at http://www.ergito.com/main.jsp?bcs=MBIO.4.14.13

© 2004. Virtual Text / www.ergito.com

DNA REPLICATION

4.14.14 Phage T4 provides its own replication apparatus

Key Concepts

- Phage T4 provides its own replication apparatus, which consists of DNA polymerase, the gene *32* SSB, a helicase, a primase, and accessory proteins that increase speed and processivity.

When phage T4 takes over an *E. coli* cell, it provides several functions of its own that either replace or augment the host functions. The phage places little reliance on expression of host functions. The degradation of host DNA is important in releasing nucleotides that are reused in the synthesis of phage DNA. (The phage DNA differs in base composition from cellular DNA in using hydroxymethylcytosine instead of the customary cytosine.)

The phage-coded functions concerned with DNA synthesis in the infected cell can be identified by mutations that impede the production of mature phages. Essential phage functions are identified by conditional lethal mutations, which fall into three phenotypic classes:

- Those in which there is no DNA synthesis at all identify genes whose products either are components of the replication apparatus or are involved in the provision of precursors (especially the hydroxymethylcytosine).
- Those in which the onset of DNA synthesis is delayed are concerned with the initiation of replication.
- Those in which DNA synthesis starts but then is arrested include regulatory functions, the DNA ligase, and some of the enzymes concerned with host DNA degradation.
- There are also nonessential genes concerned with replication; for example, including those involved in glucosylating the hydroxymethylcytosine in the DNA.

Synthesis of T4 DNA is catalyzed by a multienzyme aggregate assembled from the products of a small group of essential genes.

The gene *32* protein (gp32) is a highly cooperative single-strand binding protein, needed in stoichiometric amounts. It was the first example of its type to be characterized. The geometry of the T4 replication fork may specifically require the phage-coded protein, since the *E. coli* SSB cannot substitute. The gp32 forms a complex with the T4 DNA polymerase; this interaction could be important in constructing the replication fork.

© 2004. Virtual Text / www.ergito.com

The T4 system uses an RNA priming event that is similar to that of its host. With single-stranded T4 DNA as template, the gene *41* and *61* products act together to synthesize short primers. Their behavior is analogous to that of DnaB and DnaG in *E. coli*. The gene *41* protein is the counterpart to DnaB. It is a hexameric helicase that uses hydrolysis of GTP to provide the energy to unwind DNA (3100). The p41/p61 complex moves processively in the 5′–3′ direction in lagging strand synthesis, periodically initiating Okazaki fragments. Another protein, the product of gene *59*, loads the p41/p61 complex on to DNA; it is required to displace the p32 protein in order to allow the helicase to assemble on DNA (3099).

The gene *61* protein is needed in much smaller amounts than most of the T4 replication proteins. There are as few as 10 copies of gp61 per cell. (This impeded its characterization. It is required in such small amounts that originally it was missed as a necessary component, because enough was present as a contaminant of the gp32 preparation!) Gene *61* protein has the primase activity, analogous to DnaG of *E. coli*. The primase recognizes the template sequence 3′–TTG-5′ and synthesizes pentaribonucleotide primers that have the general sequence pppApCpNpNpNp. If the complete replication apparatus is present, these primers are extended into DNA chains.

The gene *43* DNA polymerase has the usual 5′–3′ synthetic activity, associated with a 3′–5′ exonuclease proofreading activity. It catalyzes DNA synthesis and removes the primers. When T4 DNA polymerase uses a single-stranded DNA as template, its rate of progress is uneven. The enzyme moves rapidly through single-stranded regions, but proceeds much more slowly through regions that have a base-paired intrastrand secondary structure. The accessory proteins assist the DNA polymerase in passing these roadblocks, and maintaining its speed.

The remaining three proteins are referred to as "polymerase accessory proteins". They increase the affinity of the DNA polymerase for the DNA, and also its processivity and speed. The gene *45* product is a trimer that acts as a sliding clamp. The structure of the trimer is similar to that of the E. coli β dimer, forming a circle around DNA that holds the DNA polymerase subunit more tightly on the template.

The products of genes *44* and *62* form a tight complex, which has ATPase activity. They are the equivalent of the γ δ clamp loader complex, and their role is to load p45 on to DNA. Four molecules of ATP are hydrolyzed in loading the p45 clamp and the p43 DNA polymerase on to DNA.

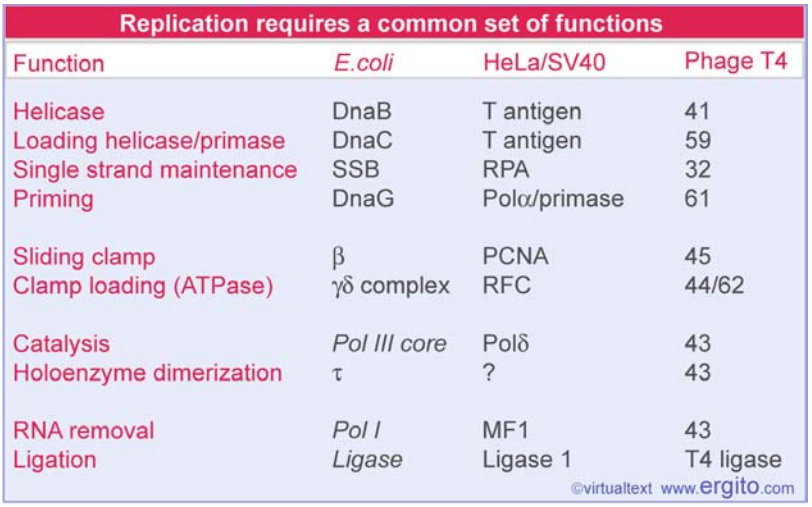
The overall structure of the replisome is similar to that of *E. coli*. It consists of two coupled holoenzyme complexes, one synthesizing the leading strand and the other synthesizing the lagging strand. In this case, the dimerization involves a direct interaction between the p43 DNA polymerase subunits, and p32 plays a role in coordinating the actions of the two DNA polymerase units (3101).

We have dealt with DNA replication so far solely in terms of the progression of the replication fork. The need for other functions is shown by the DNA-delay and DNA-arrest mutants. The four genes of the DNA-delay mutants include *39*, *52*, and *60*, which code for the three subunits of T4 topoisomerase II, an activity needed for removing supercoils in the template (see *Molecular Biology 4.15.13 Topoisomerases relax or introduce supercoils in DNA*). The essential role of this enzyme suggests

© 2004. Virtual Text / www.ergito.com

that T4 DNA does not remain in a linear form, but becomes topologically constrained during some stage of replication. The topoisomerase could be needed to allow rotation of DNA ahead of the replication fork.

Comparison of the T4 apparatus with the *E. coli* apparatus suggests that DNA replication poses a set of problems that are solved in analogous ways in different systems. We may now compare the enzymatic and structural activities found at the replication fork in *E. coli*, T4, and HeLa (human) cells. **Figure 14.25** summarizes the functions and assigns them to individual proteins. We can interpret the known properties of replication complex proteins in terms of similar functions, involving the unwinding, priming, catalytic, and sealing reactions. The components of each system interact in restricted ways, as shown by the fact that phage T4 requires its own helicase, primase, clamp, etc., and the bacterial proteins cannot substitute for their phage counterparts.



Replication requires a common set of functions

Function	*E.coli*	HeLa/SV40	Phage T4
Helicase	DnaB	T antigen	41
Loading helicase/primase	DnaC	T antigen	59
Single strand maintenance	SSB	RPA	32
Priming	DnaG	Polα/primase	61
Sliding clamp	β	PCNA	45
Clamp loading (ATPase)	γδ complex	RFC	44/62
Catalysis	*Pol III core*	Polδ	43
Holoenzyme dimerization	τ	?	43
RNA removal	*Pol I*	MF1	43
Ligation	*Ligase*	Ligase 1	T4 ligase

©virtualtext www.ergito.com

Figure 14.25 Similar functions are required at all replication forks.

© 2004. Virtual Text / www.ergito.com

References

3099. Ishmael, F. T., Alley, S. C., and Benkovic, S. J. (2002). *Assembly of the bacteriophage T4 helicase: architecture and stoichiometry of the gp41-gp59 complex.* J. Biol. Chem. 277, 20555-20562.

3100. Schrock, R. D. and Alberts, B. (1996). *Processivity of the gene 41 DNA helicase at the bacteriophage T4 DNA replication fork.* J. Biol. Chem. 271, 16678-16682.

3101. Salinas, F., and Benkovic, S. J. (2000). *Characterization of bacteriophage T4-coordinated leading- and lagging-strand synthesis on a minicircle substrate.* Proc. Natl. Acad. Sci. USA 97, 7196-7201.

This content is available online at http://www.ergito.com/main.jsp?bcs=MBIO.4.14.14

© 2004. Virtual Text / www.ergito.com

DNA REPLICATION

4.14.15 Creating the replication forks at an origin

Key Concepts

- Initiation at *oriC* requires the sequential assembly of a large protein complex.
- DnaA binds to short repeated sequences and forms an oligomeric complex that melts DNA.
- 6 DnaC monomers bind each hexamer of DnaB and this complex binds to the origin.
- A hexamer of DnaB forms the replication fork. Gyrase and SSB are also required.

Starting a cycle of replication of duplex DNA requires several successive activities:

- The two strands of DNA must suffer their initial separation. This is in effect a melting reaction over a short region.
- An unwinding point begins to move along the DNA; this marks the generation of the replication fork, which continues to move during elongation.
- The first nucleotides of the new chain must be synthesized into the primer. This action is required once for the leading strand, but is repeated at the start of each Okazaki fragment on the lagging strand.

Some events that are required for initiation therefore occur uniquely at the origin; others recur with the initiation of each Okazaki fragment during the elongation phase.

Plasmids carrying the *E. coli oriC* sequence have been used to develop a cell-free system for replication. Initiation of replication at *oriC in vitro* starts with formation of a complex that requires six proteins: DnaA, DnaB, DnaC, HU, Gyrase, and SSB. Of the six proteins involved in prepriming, DnaA draws our attention as the only one uniquely involved in initiation vis-À-vis elongation. DnaB/DnaC provides the "engine" of initiation at the origin.

The first stage in complex formation is binding to *oriC* by DnaA protein (3102). The reaction involves action at two types of sequences: 9 bp and 13 bp repeats. Together the 9 bp and 13 bp repeats define the limits of the 245 bp minimal origin, as indicated in **Figure 14.26**. An origin is activated by the sequence of events summarized in **Figure 14.27**, in which binding of DnaA is succeeded by association with the other proteins.

© 2004. Virtual Text / www.ergito.com

Figure 14.26 The minimal origin is defined by the distance between the outside members of the 13-mer and 9-mer repeats.

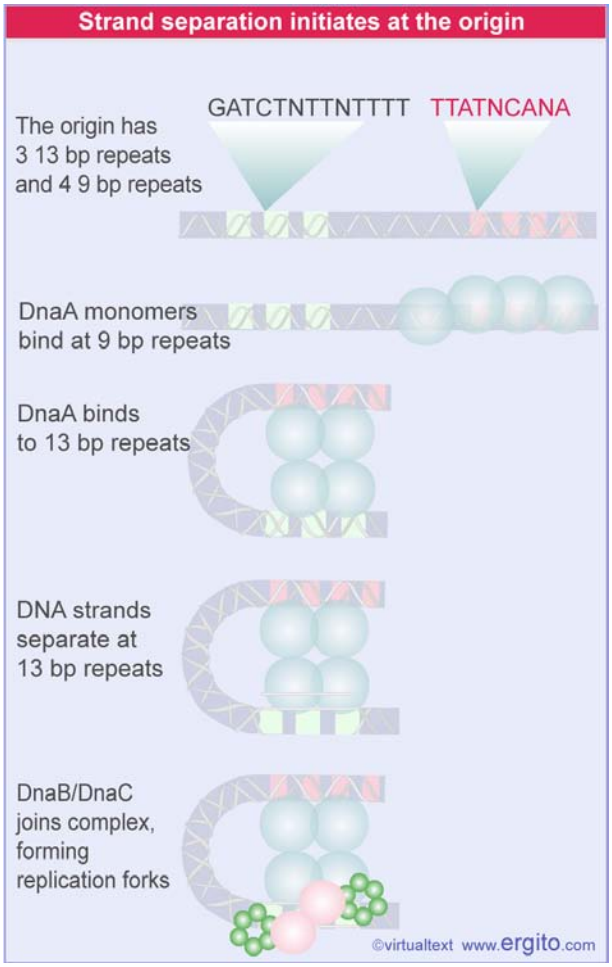


Figure 14.27 Prepriming involves formation of a complex by sequential association of proteins, leading to the separation of DNA strands.

The four 9 bp consensus sequences on the right side of *oriC* provide the initial binding sites for DnaA. It binds cooperatively to form a central core around which *oriC* DNA is wrapped. Then DnaA acts at three A-T-rich 13 bp tandem repeats located in the left side of *oriC*. In the presence of ATP, DnaA melts the DNA strands at each of these sites to form an open complex (3103; 3104). All three 13 bp repeats must be opened for the reaction to proceed to the next stage.

Altogether, 2-4 monomers of DnaA bind at the origin, and they recruit 2

© 2004. Virtual Text / www.ergito.com

"prepriming" complexes of DnaB-DnaC to bind, so that there is one for each of the two (bidirectional) replication forks. Each DnaB-DnaC complex consists of 6 DnaC monomers bound to a hexamer of DnaB. Each DnaB·DnaC complex transfers a hexamer of DnaB to an opposite strand of DNA. DnaC hydrolyzes ATP in order to release DnaB (3106).

The prepriming complex generates a protein aggregate of 480 kD, corresponding to a sphere of radius 6 nm (3107). The formation of a complex at *oriC* is detectable in the form of the large protein blob visualized in **Figure 14.28**. When replication begins, a replication bubble becomes visible next to the blob.

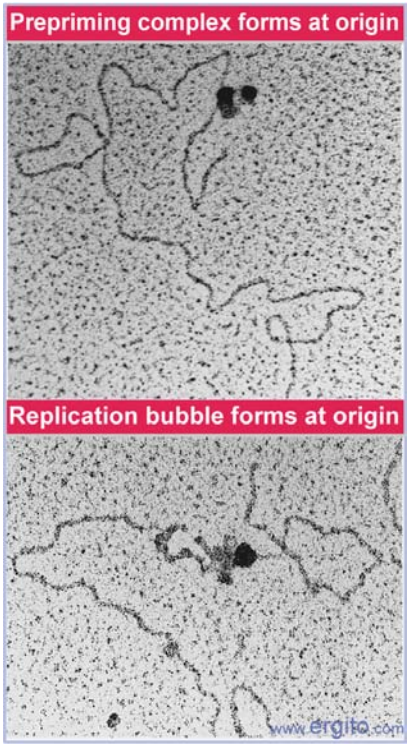


Figure 14.28 The complex at *oriC* can be detected by electron microscopy. Both complexes were visualized with antibodies against DnaB protein. Photographs kindly provided by Barbara Funnell.

The region of strand separation in the open complex is large enough for both DnaB hexamers to bind, initiating the two replication forks. As DnaB binds, it displaces DnaA from the 13 bp repeats, and extends the length of the open region. Then it uses its helicase activity to extend the region of unwinding. Each DnaB activates a DnaG primase, in one case to initiate the leading strand, and in the other to initiate the first Okazaki fragment of the lagging strand.

Two further proteins are required to support the unwinding reaction. Gyrase provides a swivel that allows one strand to rotate around the other (a reaction discussed in more detail in *Molecular Biology 4.15.15 Gyrase functions by coil inversion*); without this reaction, unwinding would generate torsional strain in the DNA. The protein SSB stabilizes the single-stranded DNA as it is formed. The length of duplex

© 2004. Virtual Text / www.ergito.com

DNA that usually is unwound to initiate replication is probably <60 bp.

The protein HU is a general DNA-binding protein in *E. coli* (see *Molecular Biology 4.18 Rearrangement of DNA*). Its presence is not absolutely required to initiate replication *in vitro*, but it stimulates the reaction. HU has the capacity to bend DNA, and is involved in building the structure that leads to formation of the open complex.

Input of energy in the form of ATP is required at several stages for the prepriming reaction. It is required for unwinding DNA. The helicase action of DnaB depends on ATP hydrolysis; and the swivel action of gyrase requires ATP hydrolysis. ATP is also needed for the action of primase and to activate DNA polymerase III.

Following generation of a replication fork as indicated in **Figure 14.27**, the priming reaction occurs to generate a leading strand. We know that synthesis of RNA is used for the priming event, but the details of the reaction are not known. Some mutations in *dnaA* can be suppressed by mutations in RNA polymerase, which suggests that DnaA could be involved in an initiation step requiring RNA synthesis *in vivo*.

RNA polymerase could be required to read into the origin from adjacent transcription units; by terminating at sites in the origin, it could provide the 3′–OH ends that prime DNA polymerase III. (An example is provided by the use of D loops at mitochondrial origins, as discussed in *Molecular Biology 4.13.7 D loops maintain mitochondrial origins* .) Alternatively, the act of transcription could be associated with a structural change that assists initiation. This latter idea is supported by observations that transcription does not have to proceed into the origin; it is effective up to 200 bp away from the origin, and can use either strand of DNA as template *in vitro*. The transcriptional event is inversely related to the requirement for supercoiling *in vitro*, which suggests that it acts by changing the local DNA structure so as to aid melting of DNA.

Last updated on 11-25-2002

© 2004. Virtual Text / www.ergito.com

References

3102. Fuller, R. S., Funnell, B. E., and Kornberg, A. (1984). *The dnaA protein complex with the E. coli chromosomal replication origin (oriC) and other DNA sites.* Cell 38, 889-900.

3103. Sekimizu, K, Bramhill, D., and Kornberg, A. (1987). *ATP activates dnaA protein in initiating replication of plasmids bearing the origin of the E. coli chromosome.* Cell 50, 259-265.

3104. Bramhill, D. and Kornberg, A. (1988). *Duplex opening by dnaA protein at novel sequences in initiation of replication at the origin of the E. coli chromosome.* Cell 52, 743-755.

3106. Wahle, E., Lasken, R. S., and Kornberg, A. (1989). *The dnaB-dnaC replication protein complex of Escherichia coli. II. Role of the complex in mobilizing dnaB functions.* J. Biol. Chem. 264, 2469-2475.

3107. Funnell, B. E. and Baker, T. A. (1987). *In vitro assembly of a prepriming complex at the origin of the E. coli chromosome.* J. Biol. Chem. 262, 10327-10334.

This content is available online at http://www.ergito.com/main.jsp?bcs=MBIO.4.14.15

© 2004. Virtual Text / www.ergito.com

DNA REPLICATION

4.14.16 Common events in priming replication at the origin

Key Concepts

- The general principle of bacterial initiation is that the origin is initially recognized by a protein that forms a large complex with DNA.
- A short region of A·T-rich DNA is melted.
- DnaB is bound to the complex and creates the replication fork.

Another system for investigating interactions at the origin is provided by phage lambda. A map of the region is shown in **Figure 14.29**. Initiation of replication at the lambda origin requires "activation" by transcription starting from P_R. As with the events at *oriC*, this does not necessarily imply that the RNA provides a primer for the leading strand. Analogies between the systems suggest that RNA synthesis could be involved in promoting some structural change in the region.

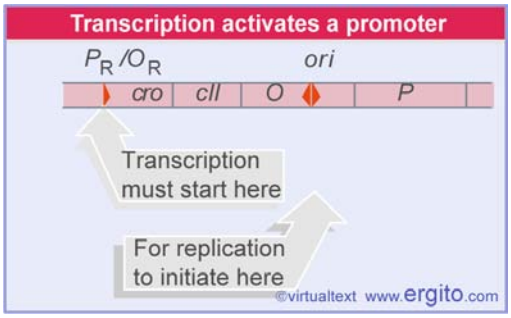


Figure 14.29 Transcription initiating at P_R is required to activate the origin of lambda DNA.

Initiation requires the products of phage genes *O* and *P*, as well as several host functions. The phage O protein binds to the lambda origin; the phage P protein interacts with the O protein and with the bacterial proteins. The origin lies within gene *O*, so the protein acts close to its site of synthesis.

Variants of the phage called λ*dv* consist of shorter genomes that carry all the information needed to replicate, but lack infective functions. λ*dv* DNA survives in the bacterium as a plasmid, and can be replicated *in vitro* by a system consisting of the phage-coded proteins O and P together with bacterial replication functions.

Lambda proteins O and P form a complex together with DnaB at the lambda origin, *ori* λ. The origin consists of two regions; as illustrated in **Figure 14.30**, a series of four binding sites for the O protein is adjacent to an A·T-rich region.

© 2004. Virtual Text / www.ergito.com

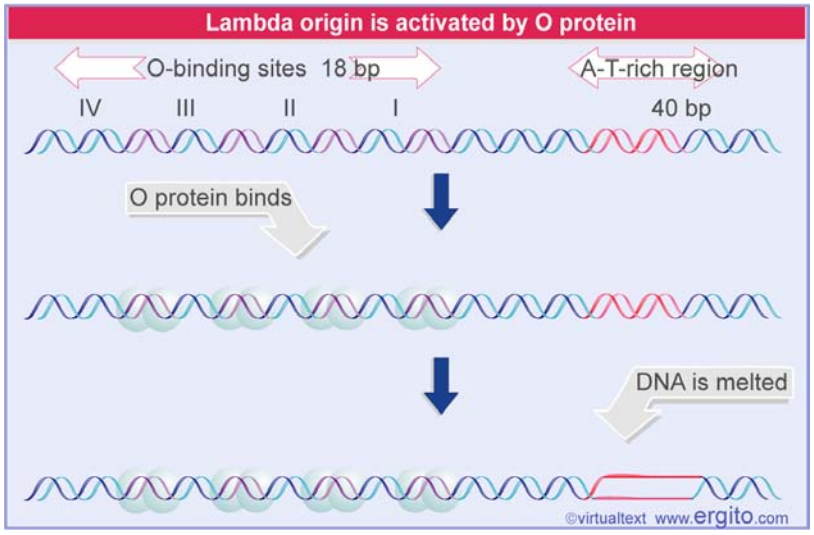


Figure 14.30 The lambda origin for replication comprises two regions. Early events are catalyzed by O protein, which binds to a series of 4 sites; then DNA is melted in the adjacent A-T-rich region. Although the DNA is drawn as a straight duplex, it is actually bent at the origin.

The first stage in initiation is the binding of O to generate a roughly spherical structure of diameter ~11 nm, sometimes called the O-some. The O-some contains ~100 bp or 60 kD of DNA. There are four 18 bp binding sites for O protein, which is ~34 kD. Each site is palindromic, and probably binds a symmetrical O dimer. The DNA sequences of the O-binding sites appear to be bent, and binding of O protein induces further bending.

If the DNA is supercoiled, binding of O protein causes a structural change in the origin. The A·T-rich region immediately adjacent to the O-binding sites becomes susceptible to S1 nuclease, an enzyme that specifically recognizes unpaired DNA. This suggests that a melting reaction occurs next to the complex of O proteins.

The role of the O protein is analogous to that of DnaA at *oriC*: it prepares the origin for binding of DnaB. Lambda provides its own protein, P, which substitutes for DnaC, and brings DnaB to the origin. When lambda P protein and bacterial DnaB proteins are added, the complex becomes larger and asymmetrical. It includes more DNA (a total of ~160 bp) as well as extra proteins. The λ P protein has a special role: it inhibits the helicase action of DnaB. Replication fork movement is triggered when P protein is released from the complex. Priming and DNA synthesis follow.

Some proteins are essential for replication without being directly involved in DNA synthesis as such. Interesting examples are provided by the DnaK and DnaJ proteins. DnaK is a chaperone, related to a common stress protein of eukaryotes. Its ability to interact with other proteins in a conformation-dependent manner plays a role in many cellular activities, including replication. The role of DnaK/DnaJ may be to disassemble the pre-priming complex; by causing the release of P protein, they allow replication to begin.

The initiation reactions at *oriC* and ori λ are similar. The same stages are involved, and rely upon overlapping components. The first step is recognition of the origin by a

© 2004. Virtual Text / www.ergito.com

protein that binds to form a complex with the DNA, DnaA for *oriC* and O protein for *ori* λ. A short region of A·T-rich DNA is melted. Then DnaB is loaded; this requires different functions at *oriC* and *ori* λ (and yet other proteins are required for this stage at other origins). When the helicase DnaB joins the complex, a replication fork is created. Finally an RNA primer is synthesized, after which replication begins.

The use of *oriC* and *ori* λ provides a general model for activation of origins. A similar series of events occurs at the origin of the virus SV40 in mammalian cells. Two hexamers of T antigen, a protein coded by the virus, bind to a series of repeated sites in DNA. In the presence of ATP, changes in DNA structure occur, culminating in a melting reaction. In the case of SV40, the melted region is rather short and is not A·T-rich, but it has an unusual composition in which one strand consists almost exclusively of pyrimidines and the other of purines. Near this site is another essential region, consisting of A·T base pairs, at which the DNA is bent; it is underwound by the binding of T antigen. An interesting difference from the prokaryotic systems is that T antigen itself possesses the helicase activity needed to extend unwinding, so that an equivalent for DnaB is not needed.

This content is available online at http://www.ergito.com/main.jsp?bcs=MBIO.4.14.16

© 2004. Virtual Text / www.ergito.com

DNA REPLICATION

4.14.17 The primosome is needed to restart replication

Key Terms

The **primosome** describes the complex of proteins involved in the priming action that initiates replication on ϕX-type origins. It is also involved in restarting stalled replication forks.

Key Concepts

- Initiation of ϕX replication requires the primosome complex to displace SSB from the origin.
- A replication fork stalls when it arrives at damaged DNA.
- After the damage has been repaired, the primosome is required to reinitiate replication.
- The Tus protein binds to *ter* sites and stops DnaB from unwinding DNA, which causes replication to terminate.

Early work on replication made extensive use of phage ϕX174, and led to the discovery of a complex system for priming. ϕX174 DNA is not by itself a substrate for the replication apparatus, because the naked DNA does not provide a suitable template. But once the single-stranded form has been coated with SSB, replication can proceed. A **primosome** assembles at a unique site on the single-stranded DNA, called the assembly site (*pas*). The *pas* is the equivalent of an origin for synthesis of the complementary strand of ϕX174. The primosome consists of six proteins: PriA, PriB, PriC, DnaT, DnaB, and DnaC. The key event in localizing the primosome is the ability of PriA to displace SSB from single-stranded DNA.

Although the primosome forms initially at the *pas* on ϕX174 DNA, primers are initiated at a variety of sites. PriA translocates along the DNA, displacing SSB, to reach additional sites at which priming occurs. As in *oriC* replicons, DnaB plays a key role in unwinding and priming in ϕX replicons. The role of PriA is to load DnaB to form a replication fork.

It has always been puzzling that ϕX origins should use a complex structure that is not required to replicate the bacterial chromosome. Why does the bacterium provide this complex?

The answer is provided by the fate of stalled replication forks. **Figure 14.31** compares an advancing replication fork with what happens when there is damage to a base in the DNA or a nick in one strand. In either case, DNA synthesis is halted, and the replication fork is either stalled or disrupted. It is not clear whether the components of the fork remain associated with the DNA or disassemble. Replication fork stalling appears to be quite common; estimates for the frequency in *E. coli*

© 2004. Virtual Text / www.ergito.com

suggest that 18-50% of bacteria encounter a problem during a replication cycle (for review see 3442).

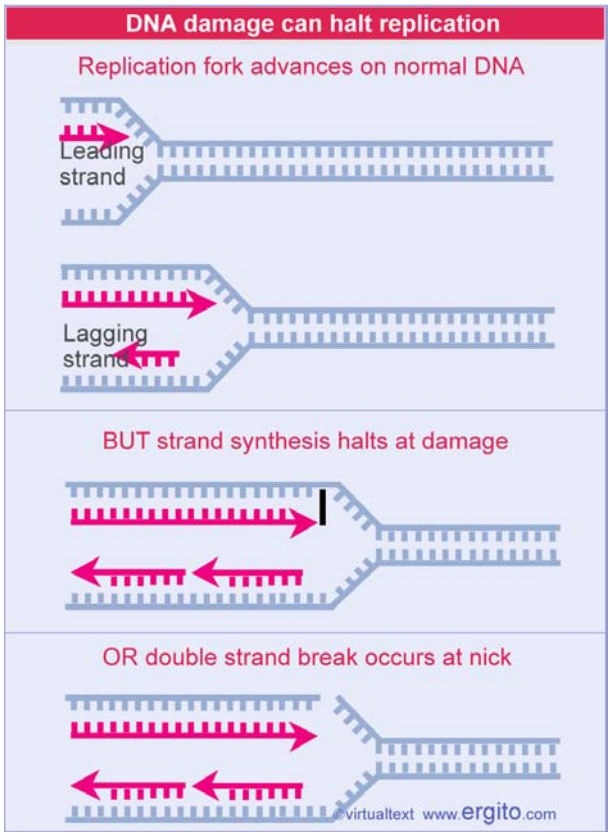


Figure 14.31 Replication is halted by a damaged base or nick in DNA.

The situation is rescued by a recombination event that excises and replaces the damage or provides a new duplex to replace the region containing the double-strand break (958; see **Figure 15.48** in *Molecular Biology 4.15.25 Recombination-repair systems in E. coli*). The principle of the repair event is to use the built in redundancy of information between the two DNA strands. **Figure 14.32** shows the key events in such a repair event. Basically, information from the undamaged DNA daughter duplex is used to repair the damaged sequence. This creates a typical recombination-junction that is resolved by the same systems that perform homologous recombination. In fact, one view is that the major importance of these systems for the cell is in repairing damaged DNA at stalled replication forks (for review see 2287).

© 2004. Virtual Text / www.ergito.com

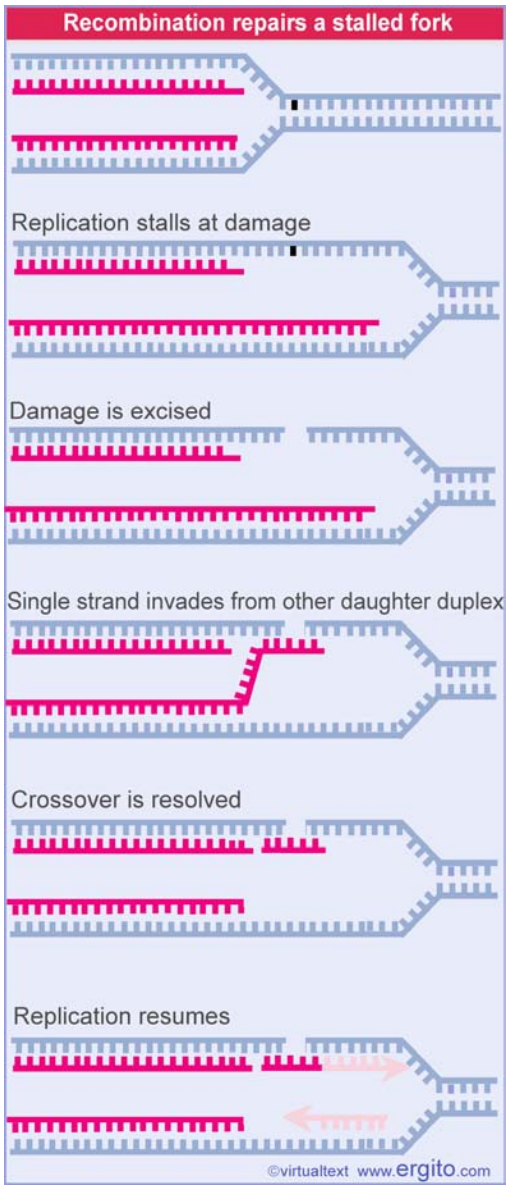


Figure 14.32 When replication halts at damaged DNA, the damaged sequence is excised, and the complementary (newly synthesized) strand of the other daughter duplex crosses over to repair the gap. Replication can now resume, and the gaps are filled in.

After this the damage has been repaired, the replication fork must be restarted. **Figure 14.33** shows that this may be accomplished by assembly of the primosome, which in effect reloads DnaB so that helicase action can continue (957; 956).

© 2004. Virtual Text / www.ergito.com

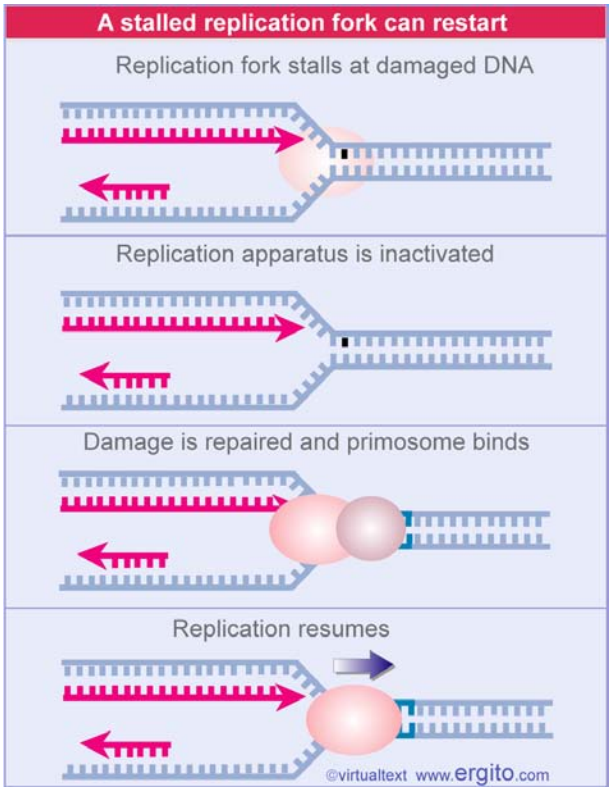


Figure 14.33 The primosome is required to restart a stalled replication fork after the DNA has been repaired.
This is a static version of an interactive figure; see http://www.ergito.com/main.jsp?bcs=MBIO.4.14.17 to view properly.

Replication fork reactivation is a common (and therefore important) reaction. It may be required in most chromosomal replication cycles. It is impeded by mutations in either the retrieval systems that replace the damaged DNA or in the components of the primosome.

Replication forks must stop and disassemble at the termination of replication. How is this accomplished?

Sequences that stop movement of replication forks have been identified in the form of the *ter* elements of the *E. coli* chromosome (see **Figure 13.7**) or equivalent sequences in some plasmids. The common feature of these elements is a 23 bp consensus sequence that provides the binding site for the product of the *tus* gene, a 36 kD protein that is necessary for termination. Tus binds to the consensus sequence, where it provides a contra-helicase activity and stops DnaB from unwinding DNA. The leading strand continues to be synthesized right up to the *ter* element, while the nearest lagging strand is initiated 50-100 bp before reaching *ter*.

The result of this inhibition is to halt movement of the replication fork and (presumably) to cause disassembly of the replication apparatus. **Figure 14.34** reminds us that Tus stops the movement of a replication fork in only one direction. The crystal structure of a Tus-ter complex shows that the Tus protein binds to DNA

© 2004. Virtual Text / www.ergito.com

asymmetrically; α-helices of the protein protrude around the double helix at the end that blocks the replication fork. Presumably a fork proceeding in the opposite direction can displace Tus and thus continue. A difficulty in understanding the function of the system *in vivo* is that it appears to be dispensable, since mutations in the *ter* sites or in *tus* are not lethal.

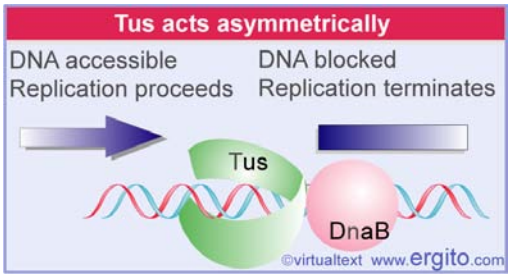


Figure 14.34 Tus binds to *ter* asymmetrically and blocks replication in only one direction.

Last updated on 5-15-2002

© 2004. Virtual Text / www.ergito.com

Reviews

957. Kuzminov, A. (1995). *Collapse and repair of replication forks in E. coli.* Mol. Microbiol. 16, 373-384.

2287. Cox, M. M. (2001). *Recombinational DNA repair of damaged replication forks in E. coli: questions.* Annu. Rev. Genet. 35, 53-82.

3442. McGlynn, P. and Lloyd, R. G. (2002). *Recombinational repair and restart of damaged replication forks.* Nat. Rev. Mol. Cell Biol. 3, 859-870.

References

956. Cox, M. M., Goodman, M. F., Kreuzer, K. N., Sherratt, D. J., Sandler, S. J., and Marians, K. J. (2000). *The importance of repairing stalled replication forks.* Nature 404, 37-41.

958. Seigneur, M., Bidnenko, V., Ehrlich, S. D., and Michel, B. (1998). *RuvAB acts at arrested replication forks.* Cell 95, 419-430.

This content is available online at http://www.ergito.com/main.jsp?bcs=MBIO.4.14.17

© 2004. Virtual Text / www.ergito.com

DNA REPLICATION

4.14.18 Does methylation at the origin regulate initiation?

Key Terms

Hemimethylated DNA is methylated on one strand of a target sequence that has a cytosine on each strand.

Key Concepts

- *oriC* contains 11 GATC / CTAG repeats that are methylated on adenine on both strands.
- Replication generates hemimethylated DNA, which cannot initiate replication.
- There is a 13 min delay before the GATC / CTAG repeats are remethylated.

What feature of a bacterial (or plasmid) origin ensures that it is used to initiate replication only once per cycle? Is initiation associated with some change that marks the origin so that a replicated origin can be distinguished from a nonreplicated origin?

Some sequences that are used for this purpose are included in the origin. *oriC* contains 11 copies of the sequence GATC / CTAG, which is a target for methylation at the N_6 position of adenine by the Dam methylase. The reaction is illustrated in **Figure 14.35**.

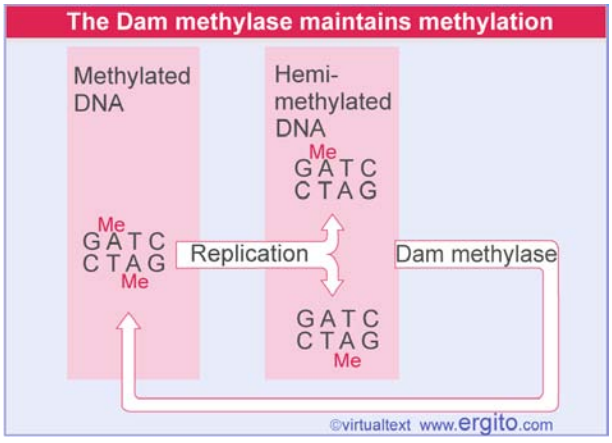


Figure 14.35 Replication of methylated DNA gives hemimethylated DNA, which maintains its state at GATC sites until the Dam methylase restores the fully methylated condition.

© 2004. Virtual Text / www.ergito.com

Before replication, the palindromic target site is methylated on the adenines of each strand. Replication inserts the normal (nonmodified) bases into the daughter strands, generating **hemimethylated** DNA, in which one strand is methylated and one strand is unmethylated. So the replication event converts Dam target sites from fully methylated to hemimethylated condition.

What is the consequence for replication? The ability of a plasmid relying upon *oriC* to replicate in *dam*$^{-}$*E. coli* depends on its state of methylation. If the plasmid is methylated, it undergoes a single round of replication, and then the hemimethylated products accumulate, as described in **Figure 14.36**. So a hemimethylated origin cannot be used to initiate a replication cycle.

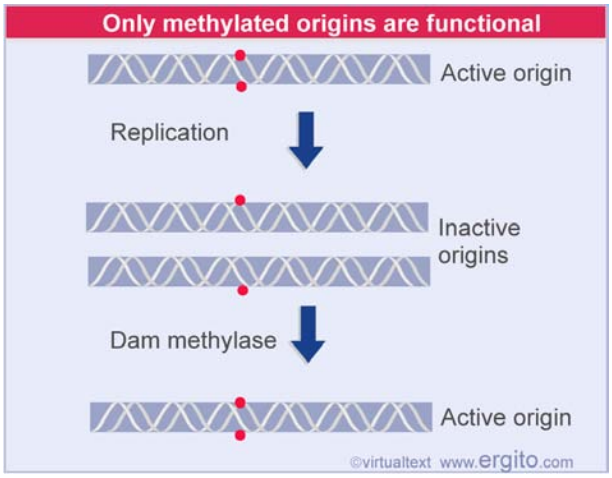


Figure 14.36 Only fully methylated origins can initiate replication; hemimethylated daughter origins cannot be used again until they have been restored to the fully methylated state.

Two explanations suggest themselves. Initiation may require full methylation of the Dam target sites in the origin. Or initiation may be inhibited by hemimethylation of these sites. The latter seems to be the case, because an origin of nonmethylated DNA can function effectively.

So hemimethylated origins cannot initiate again until the Dam methylase has converted them into fully methylated origins. The GATC sites at the origin remain hemimethylated for ~13 minutes after replication (3111). This long period is unusual, because at typical GATC sites elsewhere in the genome, remethylation begins immediately (<1.5 min) following replication. One other region behaves like *oriC*; the promoter of the *dnaA* gene also shows a delay before remethylation begins.

While it is hemimethylated, the *dnaA* promoter is repressed, which causes a reduction in the level of DnaA protein. So the origin itself is inert, and production of the crucial initiator protein is repressed, during this period.

© 2004. Virtual Text / www.ergito.com

References

3111. Campbell, J. L. and Kleckner, N. (1990). *E. coli oriC and the dnaA gene promoter are sequestered from dam methyltransferase following the passage of the chromosomal replication fork.* Cell 62, 967-979.

This content is available online at http://www.ergito.com/main.jsp?bcs=MBIO.4.14.18

© 2004. Virtual Text / www.ergito.com

DNA REPLICATION

4.14.19 Origins may be sequestered after replication

Key Concepts

- SeqA binds to hemimethylated DNA and is required for delaying rereplication.
- SeqA may interact with DnaA.
- While the origins are hemimethylated, they bind to the cell membrane, and may be unavailable to methylases.
- The nature of the connection between the origin and the membrane is still unclear.

What is responsible for the delay in remethylation at *oriC* and *dnaA*? The most likely explanation is that these regions are sequestered in a form in which they are inaccessible to the Dam methylase.

A circuit responsible for controlling reuse of origins is identified by mutations in the gene *seqA*. The mutants reduce the delay in remethylation at both *oriC* and *dnaA*. As a result, they initiate DNA replication too soon, thereby accumulating an excessive number of origins (3108). This suggests that *seqA* is part of a negative regulatory circuit that prevents origins from being remethylated (3110). SeqA binds to hemimethylated DNA more strongly than to fully methylated DNA. It may initiate binding when the DNA becomes hemimethylated, and then its continued presence prevents formation of an open complex at the origin. SeqA does not have specificity for the *oriC* sequence, and it seems likely that this is conferred by DnaA protein, which would explain genetic interactions between *seqA* and *dnaA* (3109).

Hemimethylation of the GATC sequences in the origin is required for its association with the cell membrane *in vitro*. Hemimethylated *oriC* DNA binds to the membranes, but DNA that is fully methylated does not bind. One possibility is that membrane association is involved in controlling the activity of the origin. This function could be separate from any role that the membrane plays in segregation (see **Figure 13.26**). Association with the membrane could prevent reinitiation from occurring prematurely, either indirectly because the origins are sequestered or directly because some component at the membrane inhibits the reaction.

The properties of the membrane fraction suggest that it includes components that regulate replication. An inhibitor is found in this fraction that competes with DnaA protein. This inhibitor can prevent initiation of replication only if it is added to an *in vitro* system before DnaA protein. This suggests the model of **Figure 14.37**, in which the inhibitor specifically recognizes hemimethylated DNA and prevents DnaA from binding. When the DNA is remethylated, the inhibitor is released, and DnaA now is free to initiate replication. If the inhibitor is associated with the membrane, then association and dissociation of DNA with the membrane may be involved in the control of replication.

© 2004. Virtual Text / www.ergito.com

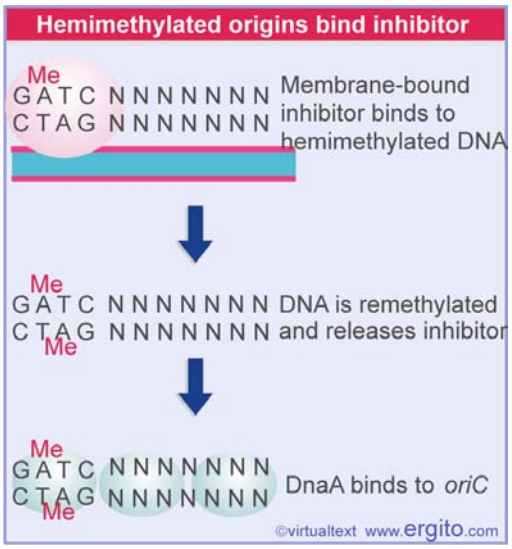


Figure 14.37 A membrane-bound inhibitor binds to hemimethylated DNA at the origin, and may function by preventing the binding of DnaA. It is released when the DNA is remethylated.

The full scope of the system used to control reinitiation is not clear, but several mechanisms may be involved: physical sequestration of the origin; delay in remethylation; inhibition of DnaA binding; repression of *dnaA* transcription. It is not immediately obvious which of these events cause the others, and whether their effects on initiation are direct or indirect.

We still have to come to grips with the central issue of which feature has the basic responsibility for timing. One possibility is that attachment to the membrane occurs at initiation, and that assembly of some large structure is required to release the DNA. The period of sequestration appears to increase with the length of the cell cycle, which suggests that it directly reflects the clock that controls reinitiation.

As the only member of the replication apparatus uniquely required at the origin, DnaA has attracted much attention. DnaA is a target for several regulatory systems. It may be that no one of these systems by itself is adequate to control frequency of initiation, but the combination achieves the desired result. Some mutations in *dnaA* render replication asynchronous, suggesting that DnaA could be the "titrator" or "clock" that measures the number of origins relative to cell mass. Overproduction of DnaA yields conflicting results, varying from no effect to causing initiation to take place at reduced mass.

It has been difficult to identify the protein component(s) that mediate membrane-attachment. A hint that this is a function of DnaA is provided by its response to phospholipids. Phospholipids promote the exchange of ATP with ADP bound to DnaA. We do not know what role this plays in controlling the activity of DnaA (which requires ATP), but the reaction implies that DnaA is likely to interact with the membrane. This would imply that more than one event is involved in associating with the membrane. Perhaps a hemimethylated origin is bound by the membrane-associated inhibitor, but when the origin becomes fully methylated, the inhibitor is displaced by DnaA associated with the membrane.

© 2004. Virtual Text / www.ergito.com

If DnaA is the initiator that triggers a replication cycle, the key event will be its accumulation at the origin to a critical level. There are no cyclic variations in the overall concentration or expression of DnaA, which suggests that local events must be responsible. To be active in initiating replication, DnaA must be in the ATP-bound form. DnaA has a weak intrinsic activity that converts the ATP to ADP. This activity is enhanced by the β subunit of DNA polymerase III. When the replicase is incorporated into the replication complex, this interaction causes hydrolysis of the ATP bound to DnaA, thereby inactivating DnaA, and preventing it from starting another replication cycle (3112; 3113). This reaction has been called *RIDA* (regulatory inactivation of DnaA). It is enhanced by a protein called Hda (3114). We do not yet know what controls the timing of the reactivation of DnaA.

Another factor that controls availability of DnaA at the origin is the competition for binding it to other sites on DNA. In particular, a locus called *dat* has a large concentration of DnaA-binding sites. It binds about 8× more DnaA than the origin. Deletion of *dat* causes initiation to occur more frequently (3115). This significantly reduces the amount of DnaA available to the origin, but we do not yet understand exactly what role this may play in controlling the timing of initiation.

© 2004. Virtual Text / www.ergito.com

References

3108. Lu, M., Campbell, J. L., Boye, E., and Kleckner, N. (1994). *SeqA: a negative modulator of replication initiation in E. coli.* Cell 77, 413-426.

3109. Slater, S., Wold, S., Lu, M., Boye, E., Skarstad, K., and 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.

3110. Wold, S., Boye, S., Wold, S., Slater, S., Kleckner, N., and Skarstad, K. (1998). *Effects of purified SeqA protein on oriC-dependent DNA replication in vitro.* EMBO J. 17, 4158-4165.

3112. Katayama, T., Kurokawa, K., Crooke, E., and Sekimizu, K. (1998). *The initiator function of DnaA protein is negatively regulated by the sliding clamp of the E. coli chromosomal replicase.* Cell 94, 61-71.

3113. Kurokawa, K., Nishida, S., Emoto, A., Sekimizu, K., and Katayama, T. (1999). *Replication cycle-coordinated change of the adenine nucleotide-bound forms of DnaA protein in Escherichia coli.* EMBO J. 18, 6642-6652.

3114. Kato , J. and Katayama, T. (2001). *Hda, a novel DnaA-related protein, regulates the replication cycle in Escherichia coli.* EMBO J. 20, 4253-4262.

3115. Kitagawa, R., Ozaki, T., Moriya, S., and Ogawa, T. (1998). *Negative control of replication initiation by a novel chromosomal locus exhibiting exceptional affinity for E. coli DnaA protein.* Genes Dev. 12, 3032-3043.

This content is available online at http://www.ergito.com/main.jsp?bcs=MBIO.4.14.19

© 2004. Virtual Text / www.ergito.com

DNA REPLICATION

4.14.20 Licensing factor controls eukaryotic rereplication

Key Terms

A **licensing factor** is something in the nucleus that is necessary for replication, and is inactivated or destroyed after one round of replication. New licensing factors must be provided for further rounds of replication to occur.

Key Concepts

- Licensing factor is necessary for initiation of replication at each origin.
- It is present in the nucleus prior to replication, but is inactivated or destroyed by replication.
- Initiation of another replication cycle becomes possible only after licensing factor reenters the nucleus after mitosis.

A eukaryotic genome is divided into multiple replicons, and the origin in each replicon is activated once and only once in a single division cycle. This could be achieved by providing some rate-limiting component that functions only once at an origin or by the presence of a repressor that prevents rereplication at origins that have been used. The critical questions about the nature of this regulatory system are how the system determines whether any particular origin has been replicated, and what protein components are involved.

Insights into the nature of the protein components have been provided by using a system in which a substrate DNA undergoes only one cycle of replication. *Xenopus* eggs have all the components needed to replicate DNA – in the first few hours after fertilization they undertake 11 division cycles without new gene expression – and they can replicate the DNA in a nucleus that is injected into the egg. **Figure 14.38** summarizes the features of this system.

© 2004. Virtual Text / www.ergito.com

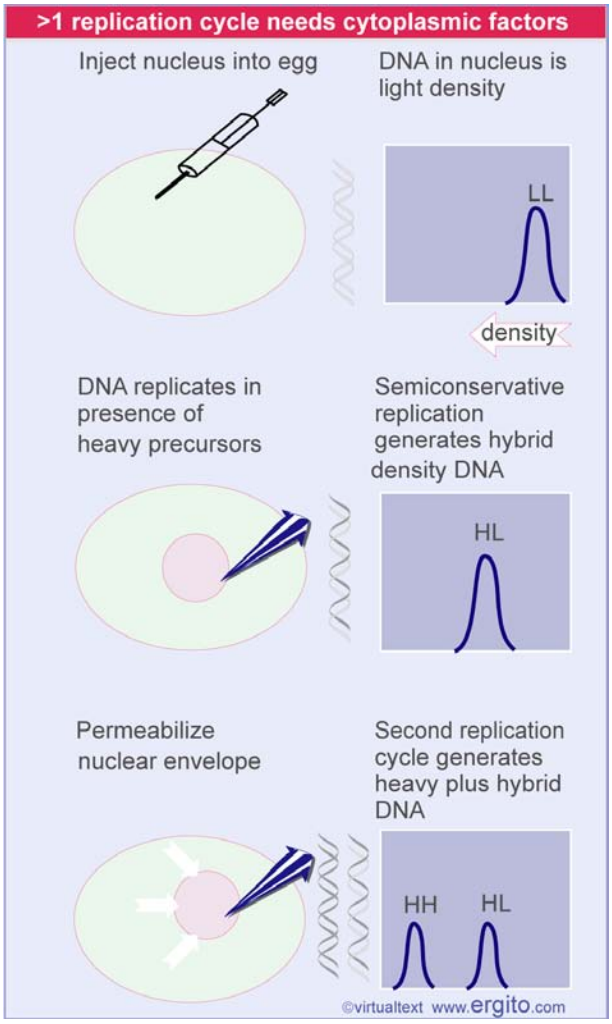


Figure 14.38 A nucleus injected into a *Xenopus* egg can replicate only once unless the nuclear membrane is permeabilized to allow subsequent replication cycles.

When a sperm or interphase nucleus is injected into the egg, its DNA is replicated only once (this can be followed by use of a density label, just like the original experiment that characterized semiconservative replication, shown previously in **Figure 1.12**). If protein synthesis is blocked in the egg, the membrane around the injected material remains intact, and the DNA cannot replicate again. However, in the presence of protein synthesis, the nuclear membrane breaks down just as it would for a normal cell division, and in this case subsequent replication cycles can occur. The same result can be achieved by using agents that permeabilize the nuclear membrane. This suggests that the nucleus contains a protein(s) needed for replication that is used up in some way by a replication cycle; although more of the protein is present in the egg cytoplasm, it can only enter the nucleus if the nuclear membrane breaks down. The system can in principle be taken further by developing an *in vitro* extract that supports nuclear replication, thus allowing the components of the extract to be isolated, and the relevant factors identified.

Figure 14.39 explains the control of reinitiation by proposing that this protein is a

© 2004. Virtual Text / www.ergito.com

licensing factor (3116). It is present in the nucleus prior to replication. One round of replication either inactivates or destroys the factor, and another round cannot occur until further factor is provided. Factor in the cytoplasm can gain access to the nuclear material only at the subsequent mitosis when the nuclear envelope breaks down. This regulatory system achieves two purposes. By removing a necessary component after replication, it prevents more than one cycle of replication from occurring. And it provides a feedback loop that makes the initiation of replication dependent on passing through cell division.

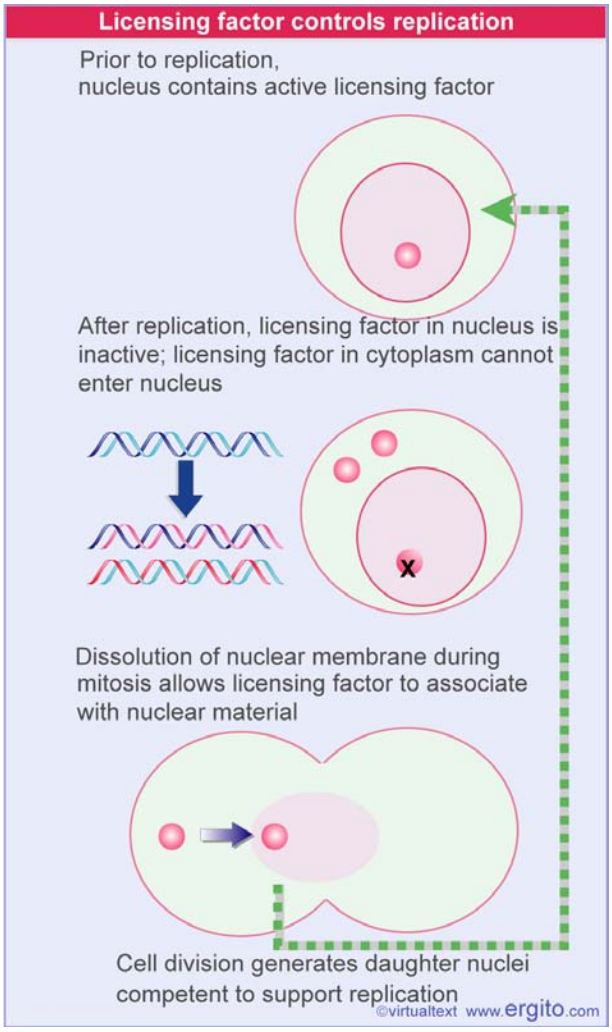


Figure 14.39 Licensing factor in the nucleus is inactivated after replication. A new supply of licensing factor can enter only when the nuclear membrane breaks down at mitosis.

© 2004. Virtual Text / www.ergito.com

References

3116. Blow, J. J. and Laskey, R. A. (1988). *A role for the nuclear envelope in controlling DNA replication within the cell cycle.* Nature 332, 546-548.

This content is available online at http://www.ergito.com/main.jsp?bcs=MBIO.4.14.20

© 2004. Virtual Text / www.ergito.com

DNA REPLICATION

4.14.21 Licensing factor consists of MCM proteins

Key Terms

The **prereplication complex** is a protein-DNA complex at the origin in *S. cerevisiae* that is required for DNA replication. The complex contains the ORC complex, Cdc6, and the MCM proteins.

The **postreplication complex** is a protein-DNA complex in *S. cerevisiae* that consists of the ORC complex bound to the origin.

Key Concepts

- The ORC is a protein complex that is associated with yeast origins throughout the cell cycle.
- Cdc6 protein is an unstable protein that is synthesized only in G1.
- Cdc6 binds to ORC and allows MCM proteins to bind.
- When replication is initiated, Cdc6 and MCM proteins are displaced. The degradation of Cdc6 prevents reinitiation.
- Some MCM proteins are in the nucleus throughout the cycle, but others may enter only after mitosis.

The key event in controlling replication is the behavior of the ORC complex at the origin (for review see 2503; 3119; 3089). Recall that ORC is a 400 kD complex that binds to the *S. cerevisiae ARS* sequence (see *Molecular Biology 4.13.6 Replication origins can be isolated in yeast*). The origin (*ARS*) consists of the A consensus sequence and three B elements (see **Figure 13.10**). The ORC complex of 6 proteins (all of which are coded by essential genes) binds to the A and adjacent B1 element. ATP is required for the binding, but is not hydrolyzed until some later stage. The transcription factor ABF1 binds to the B3 element; this assists initiation, but it is the events that occur at the A and B1 elements that actually cause initiation. Most origins are localized in regions between genes, which suggests that it may be important for the local chromatin structure to be in a nontranscribed condition.

The striking feature is that ORC remains bound at the origin through the entire cell cycle. However, changes occur in the pattern of protection of DNA as a result of the binding of other proteins to the ORC-origin complex (2502; 2501; 3118). **Figure 14.40** summarizes the cycle of events at the origin.

© 2004. Virtual Text / www.ergito.com

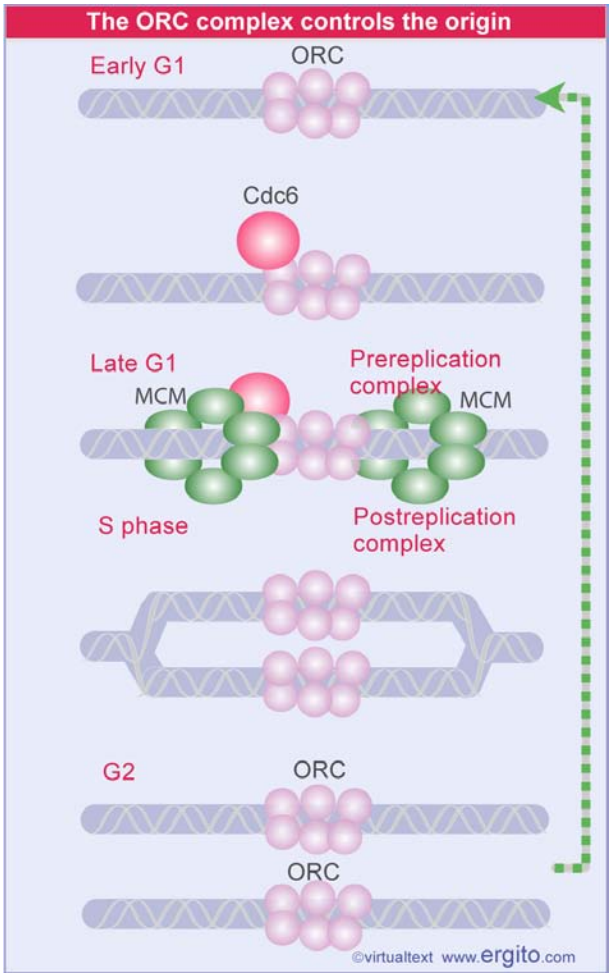


Figure 14.40 Proteins at the origin control susceptibility to initiation.

At the end of the cell cycle, ORC is bound to A/B1, and generates a pattern of protection *in vivo* that is similar to that found when it binds to free DNA *in vitro*. Basically the region across A-B1 is protected against DNAase, but there is a hypersensitive site in the center of B1.

During G1, this pattern changes, most strikingly by the loss of the hypersensitive site. This is due to the binding of Cdc6 protein to the ORC. In yeast, Cdc6 is a highly unstable protein (half-life <5 minutes). It is synthesized during G1, and typically binds to the ORC between the exit from mitosis and late G1. Its rapid degradation means that no protein is available later in the cycle. In mammalian cells it is controlled differently; it is phosphorylated during S phase, and as a result is exported from the nucleus. Cdc6 provides the connection between ORC and a complex of proteins that is involved in licensing and initiation. Cdc6 has an ATPase activity that is required for it to support initiation (3090).

An insight into the system that controls availability of licensing factor is provided by certain mutants in yeast. Mutations in the licensing factor itself could prevent initiation of replication. This is how mutations behave in *MCM2, 3, 5*. Mutations in

© 2004. Virtual Text / www.ergito.com

the system that inactivates licensing factor after the start of replication should allow the accumulation of excess quantities of DNA, because the continued presence of licensing factor allows rereplication to occur. Such mutations are found in genes that code for components of the ubiquitination system that is responsible for degrading certain proteins. This suggests that licensing factor may be destroyed after the start of a replication cycle.

The proteins MCM2, 3, 5 are required for replication and enter the nucleus only during mitosis. Homologues are found in animal cells, where MCM3 is bound to chromosomal material before replication, but is released after replication. The animal cell MCM2,3,5 complex remains in the nucleus throughout the cell cycle, suggesting that it may be one component of the licensing factor. Another component, able to enter only at mitosis, may be necessary for MCM2,3,5 to associate with chromosomal material.

In yeast, the presence of Cdc6 at the origin allows MCM proteins to bind to the complex. Their presence is necessary for initiation to occur at the origin. The origin therefore enters S phase in the condition of a **prereplication complex**, containing ORC, Cdc6, and MCM proteins. When initiation occurs, Cdc6 and MCM are displaced, returning the origin to the state of the **postreplication complex**, which contains only ORC. Because Cdc6 is rapidly degraded during S phase, it is not available to support reloading of MCM proteins, and so the origin cannot be used for a second cycle of initiation during the S phase.

If Cdc6 is made available to bind to the origin during G2 (by ectopic expression), MCM proteins do not bind until the following G1, suggesting that there is a secondary mechanism to ensure that they associate with origins only at the right time. This could be another part of licensing control. At least in *S. cerevisiae*, this control does not seem to be exercised at the level of nuclear entry, but this could be a difference between yeasts and animal cells. We discuss how the cell cycle control system regulates initiation (and reinitiation) of replication in *Molecular Biology 6.29 Cell cycle and growth regulation.*

The MCM2-7 proteins form a 6-member ring-shaped complex around DNA (2504). Some of the ORC proteins have similarities to replication proteins that load DNA polymerase on to DNA. It is possible that ORC uses hydrolysis of ATP to load the MCM ring on to DNA. In *Xenopus* extracts, replication can be initiated if ORC is removed after it has loaded Cdc6 and MCM proteins (2505). This shows that the major role of ORC is to identify the origin to the Cdc6 and MCM proteins that control initiation and licensing.

The MCM proteins are required for elongation as well as for initiation, and continue to function at the replication fork. Their exact role in elongation is not clear, but one possibility is that they contribute to the helicase activity that unwinds DNA. Another possibility is that they act as an advance guard that acts on chromatin in order to allow a helicase to act on DNA.

Last updated on 11-20-2002

© 2004. Virtual Text / www.ergito.com

Reviews

2503. Bell, S. P. (2002). *The origin recognition complex: from simple origins to complex functions.* Genes Dev. 16, 659-672.

3089. Bell, S. P. and Dutta, A. (2002). *DNA replication in eukaryotic cells.* Annu. Rev. Biochem. 71, 333-374.

3119. Dutta, A. and Bell, S. P. (1997). *Initiation of DNA replication in eukaryotic cells.* Annu. Rev. Cell Dev. Biol. 13, 293-332.

References

2501. Liang, C. and Stillman, B. (1997). *Persistent initiation of DNA replication and chromatin-bound MCM proteins during the cell cycle in cdc6 mutants.* Genes Dev. 11, 3375-3386.

2502. Diffley, J. F., Cocker, J. H., Dowell, S. J., and Rowley, A. (1994). *Two steps in the assembly of complexes at yeast replication origins in vitro.* Cell 78, 303-316.

2504. Perkins, G. and Diffley, J. F. (1998). *Nucleotide-dependent prereplicative complex assembly by Cdc6p, a homolog of eukaryotic and prokaryotic clamp-loaders.* Mol. Cell 2, 23-32.

2505. Rowles, A., Tada, S., and Blow, J. J. (1999). *Changes in association of the Xenopus origin recognition complex with chromatin on licensing of replication origins.* J. Cell. Sci. 112 , 2011-2018.

3090. Weinreich, M., Liang, C., and Stillman, B. (1999). *The Cdc6p nucleotide-binding motif is required for loading mcm proteins onto chromatin.* Proc. Natl. Acad. Sci. USA 96, 441-446.

3118. Bell, S. P. and Stillman, B. (1992). *ATP-dependent recognition of eukaryotic origins of DNA replication by a multiprotein complex.* Nature 357, 128-134.

This content is available online at http://www.ergito.com/main.jsp?bcs=MBIO.4.14.21

© 2004. Virtual Text / www.ergito.com

DNA REPLICATION

4.14.22 Summary

DNA synthesis occurs by semidiscontinuous replication, in which the leading strand of DNA growing 5′–3′ is extended continuously, but the lagging strand that grows overall in the opposite 3′–5′ direction is made as short Okazaki fragments, each synthesized 5′–3′. The leading strand and each Okazaki fragment of the lagging strand initiate with an RNA primer that is extended by DNA polymerase. Bacteria and eukaryotes each possess more than one DNA polymerase activity. DNA polymerase III synthesizes both lagging and leading strands in *E. coli*. Many proteins are required for DNA polymerase III action and several constitute part of the replisome within which it functions.

The replisome contains an asymmetric dimer of DNA polymerase III; each new DNA strand is synthesized by a different core complex containing a catalytic (α) subunit. Processivity of the core complex is maintained by the β clamp, which forms a ring round DNA. The clamp is loaded on to DNA by the clamp loader complex. Clamp/clamp loader pairs with similar structural features are widely found in both prokaryotic and eukaryotic replication systems.

The looping model for the replication fork proposes that, as one half of the dimer advances to synthesize the leading strand, the other half of the dimer pulls DNA through as a single loop that provides the template for the lagging strand. The transition from completion of one Okazaki fragment to the start of the next requires the lagging strand catalytic subunit to dissociate from DNA and then to reattach to a β clamp at the priming site for the next Okazaki fragment.

DnaB provides the helicase activity at a replication fork; this depends on ATP cleavage. DnaB may function by itself in *oriC* replicons to provide primosome activity by interacting periodically with DnaG, which provides the primase that synthesizes RNA.

Phage T4 codes for a replication apparatus consisting of 7 proteins: DNA polymerase, helicase, single-strand binding protein, priming activities, and accessory proteins. Similar functions are required in other replication systems, including a HeLa cell system that replicates SV40 DNA. Different enzymes, DNA polymerase α and DNA polymerase δ, initiate and elongate the new strands of DNA.

The ϕX priming event also requires DnaB, DnaC, and DnaT. PriA is the component that defines the primosome assembly site (*pas*) for ϕX replicons; it displaces SSB from DNA in an action that involves cleavage of ATP. PriB and PriC are additional components of the primosome. The importance of the primosome for the bacterial cell is that it is used to restart replication at forks that stall when they encounter damaged DNA.

The common mode of origin activation involves an initial limited melting of the double helix, followed by more general unwinding to create single strands. Several proteins act sequentially at the *E. coli* origin. Replication is initiated at *oriC* in *E. coli*

© 2004. Virtual Text / www.ergito.com

when DnaA binds to a series of 9 bp repeats. This is followed by binding to a series of 13 bp repeats, where it uses hydrolysis of ATP to generate the energy to separate the DNA strands. The pre-priming complex of DnaC-DnaB displaces DnaA. DnaC is released in a reaction that depends on ATP hydrolysis; DnaB is joined by the replicase enzyme, and replication is initiated by two forks that set out in opposite directions. Similar events occur at the lambda origin, where phage proteins O and P are the counterparts of bacterial proteins DnaA and DnaC, respectively. In SV40 replication, several of these activities are combined in the functions of T antigen.

The availability of DnaA at the origin is an important component of the system that determines when replication cycles should initiate. Following initiation of replication, DnaA hydrolyzes its ATP under the stimulus of the β sliding clamp, generating an inactive form of the protein. Also, *oriC* must compete with the *dat* site for binding DnaA.

Several sites that are methylated by the Dam methylase are present in the *E. coli* origin, including those of the 13-mer binding sites for DnaA. The origin remains hemimethylated and is in a sequestered state for ~10 minutes following initiation of a replication cycle. During this period it is associated with the membrane, and reinitiation of replication is repressed. The protein SeqA is involved in sequestration and may interact with DnaA.

After cell division, nuclei of eukaryotic cells have a licensing factor that is needed to initiate replication. Its destruction after initiation of replication prevents further replication cycles from occurring in yeast. Licensing factor cannot be imported into the nucleus from the cytoplasm, and can be replaced only when the nuclear membrane breaks down during mitosis.

The origin is recognized by the ORC proteins, which in yeast remain bound throughout the cell cycle. The protein Cdc6 is available only at S phase. In yeast it is synthesized during S phase and rapidly degraded. In animal cells it is synthesized continuously, but is exported from the nucleus during S phase. The presence of Cdc6 allows the MCM proteins to bind to the origin. The MCM proteins are required for initiation. The action of Cdc6 and the MCM proteins provides the licensing function.

This content is available online at http://www.ergito.com/main.jsp?bcs=MBIO.4.14.22

© 2004. Virtual Text / www.ergito.com