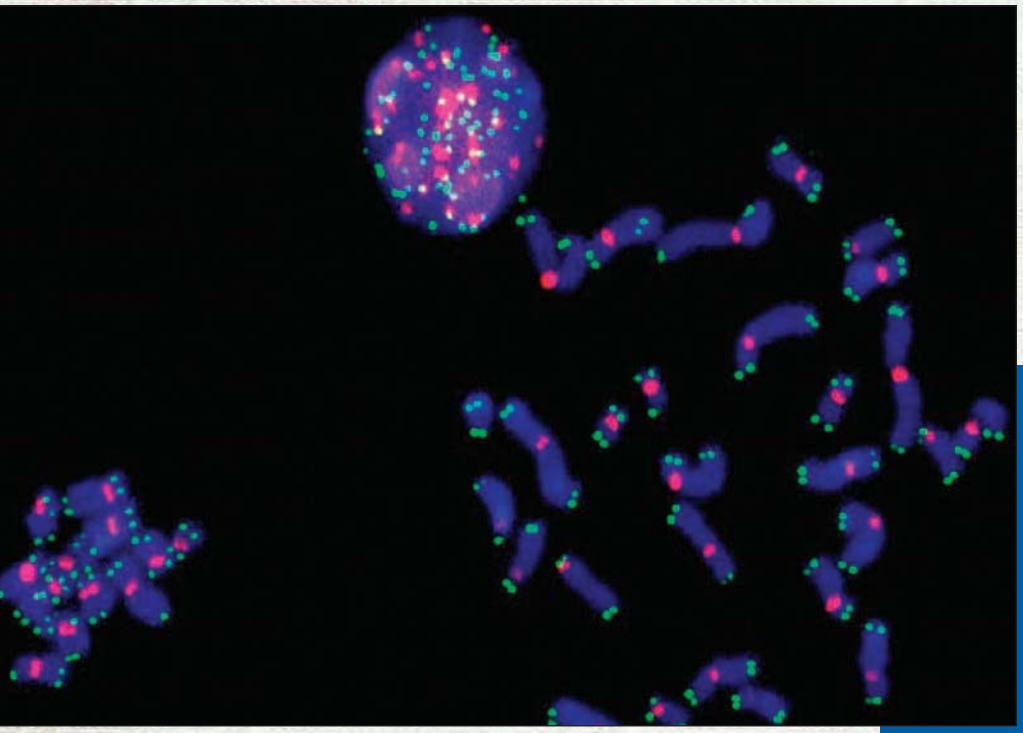
CHAPTER 21

DNA Replication II: Detailed Mechanism



Telomeres in human chromosomes. The telomeres are stained green and the centromeres are stained pink. *Cal Harley/Geron Corporation & Peter Rabinovitch, Univ. of Washington.*

We learned in Chapter 20 that DNA replication is at least semidiscontinuous and requires the synthesis of primers before DNA synthesis can begin. We have also learned about some of the major proteins involved in DNA replication in *E. coli.* Thus, we know that DNA replication is complex and involves more than just a DNA polymerase. This chapter presents a close look at the mechanism of this process in *E. coli* and in eukaryotes. We will look at the three stages of replication—initiation, elongation, and termination—in a variety of systems.

21.1 Initiation

As we have seen, initiation of DNA replication means primer synthesis. Different organisms use different mechanisms to make primers; even different phages that infect *E. coli* (coliphages) use quite different primer synthesis strategies. The coliphages were convenient tools to probe *E. coli* DNA replication because they are so simple that they have to rely primarily on host proteins to replicate their DNAs.

Priming in *E. coli*

As mentioned in Chapter 20, the first example of coliphage primer synthesis was found by accident in M13 phage, when this phage was discovered to use the host RNA polymerase as its **primase** (primer-synthesizing enzyme). But *E. coli* and its other phages do not use the host RNA polymerase as a primase. Instead, they employ a primase called **DnaG,** which is the product of the *E. coli dnaG* gene. Arthur Kornberg noted that *E. coli* and most of its phages need at least one more protein (**DnaB,** a DNA helicase introduced in Chapter 20) to form primers, at least on the lagging strand.

Arthur Kornberg and colleagues discovered the importance of DnaB with an assay in which single-stranded ϕX174 phage DNA (without SSB) is converted to double-stranded form. Synthesis of the second strand of phage DNA required primer synthesis, then DNA replication. The DNA replication part used pol III holoenzyme, so the other required proteins should be the ones needed for primer synthesis. Kornberg and colleagues found that three proteins: DnaG (the primase), DnaB, and pol III holoenzyme were required in this assay. Thus, DnaG and DnaB were apparently needed for primer synthesis. Kornberg coined the term **primosome** to refer to the collection of proteins needed to make primers for a given replicating DNA. Usually this is just two proteins, DnaG and DnaB, although other proteins may be needed to assemble the primosome.

The *E. coli* primosome is mobile and can repeatedly synthesize primers as it moves around the uncoated circular ϕX174 phage DNA. As such, it is also well suited for the repetitious task of priming Okazaki fragments on at least the lagging strand of *E. coli* DNA. This contrasts with the activity of RNA polymerase or primase alone, which prime DNA synthesis at only one spot—the origin of replication.

Two different general approaches were used to identify the important components of the *E. coli* DNA replication system, with DNA from phages ϕX174 and G4 as model substrates. The first approach was a combination genetic–biochemical one, the strategy of which was to isolate mutants with defects in their ability to replicate phage DNA, then to complement extracts from these mutants with proteins from wild-type cells. The mutant extracts were incapable of replicating the phage DNA in vitro unless the right wild-type protein was added. Using this system as an assay, the protein can be highly purified and then characterized. The second approach was the classical biochemical one: Purify all of the components needed and then add them all back together to reconstitute the replication system in vitro.

The Origin of Replication in *E. coli* Before we discuss priming further, let us consider the unique site at which DNA replication begins in ***E. coli: oriC*.** An origin of replication is a DNA site at which DNA replication begins and which is essential for proper replication to occur. We can locate the place where replication begins by several means, but how do we know how much of the DNA around the initiation site is essential for replication to begin? One way is to clone a DNA fragment, including the initiation site, into a plasmid that lacks its own origin of replication but has an antibiotic resistance gene. Then we can use the antibiotic to select for autonomously replicating plasmids. Any cell that replicates in the presence of the antibiotic must have a plasmid with a functional origin. Once we have such an *oriC* plasmid, we can begin trimming and mutating the DNA fragment containing *oriC* to find the minimal effective DNA sequence. The minimal origin in *E. coli* is 245 bp long. Some features of the origins are conserved in bacteria, and the spacing between them is also conserved.

Figure 21.1 illustrates the steps in initiation at *oriC*. The origin includes four 9-mers with the consensus sequence TTATCCACA. Two of these are in one orientation, and two are in the opposite orientation. DNase foot-printing shows that these 9-mers are binding sites for the ***dnaA* product (DnaA).** These 9-mers are therefore sometimes called ***dnaA* boxes.** DnaA appears to facilitate the binding of DnaB to the origin.

DnaA helps DnaB bind at the origin by stimulating the melting of three 13-mer repeats at the left end of *oriC* to form an **open complex.** This is analogous to the open promoter complex we discussed in Chapter 6. DnaB can then bind to the melted DNA region. Another protein, DnaC, binds to DnaB and helps deliver it to the origin.

The evidence also strongly suggests that DnaA directly assists the binding of DnaB. Here is one line of evidence that points in this direction. A *dnaA* box resides in the stem of a hairpin stem loop in a plasmid called R6K. When DnaA binds to this DNA, DnaB (with the help of DnaC) can also bind. Here, no DNA melting appears to occur, so we infer that DnaA directly affects binding between DNA and DnaB.

At least two other factors participate in open complex formation at *oriC*. The first of these is RNA polymerase. This enzyme does not serve as a primase, as it does in M13 phage replication, but it still serves an essential function. We know RNA polymerase action is required, because rifampicin blocks primosome assembly. The role of RNA polymerase seems to be to synthesize a short piece of RNA that creates an R loop (Chapter 14). The R loop can be adjacent to *oriC,* rather than within it. The second factor is

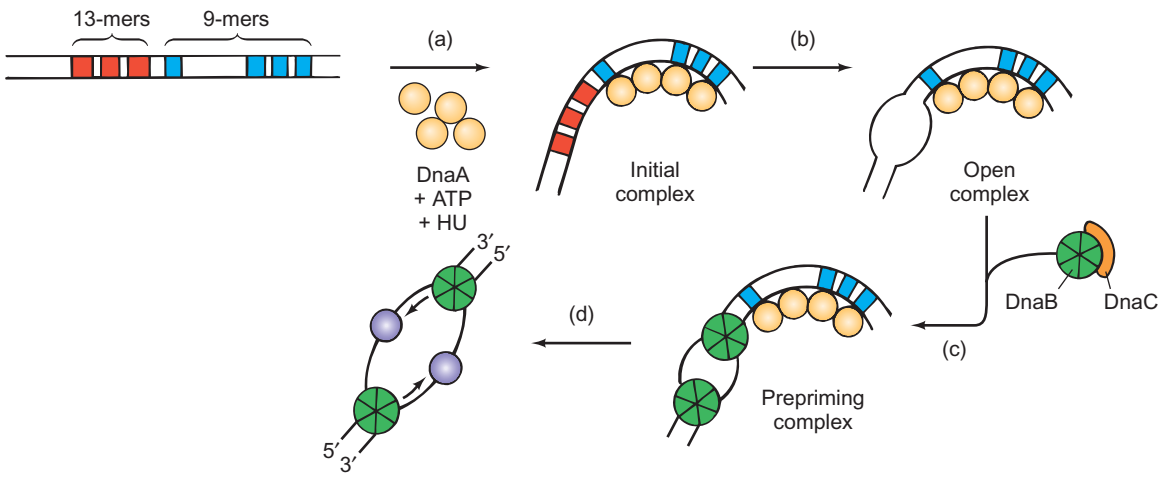


Figure 21.1 Priming at *oriC*. **(a)** Formation of the initial complex. First, DnaA (yellow) binds ATP and forms a multimer. Along with the HU protein, the DnaA/ATP complex binds to the DNA, encompassing the four 9-mers. In all, this complex covers about 200 bp. HU protein probably induces the bend in the DNA pictured here. **(b)** Formation of the open complex. The binding of DnaA, along with the bending induced by HU protein, apparently destabilizes the adjacent 13-mer repeats and causes local DNA melting there. This allows the binding of DnaB protein to the melted region. **(c)** Formation of the prepriming complex. DnaC binds to the DnaB protein and helps deliver it to the DNA. **(d)** Priming. Finally, primase (purple) binds to the prepriming complex and converts it to the primosome, which can make primers to initiate DNA replication. Primers are represented by arrows. (*Source:* Adapted from *DNA Replication,* 2/e, (plate 15) by Arthur Kornberg and Tania Baker.)

HU protein. This is a small basic DNA-binding protein that can induce bending in double-stranded DNA. This bending, together with the R loop, presumably destabilizes the DNA double helix and facilitates melting of the DNA to form the open complex.

Finally, DnaB stimulates the binding of the primase (DnaG), completing the primosome. Priming can now occur, so DNA replication can get started. The primosome remains with the replication machinery, or **replisome,** as it carries out elongation, and serves at least two functions. First, it must operate repeatedly in priming Okazaki fragment synthesis to build the lagging strand. Second, DnaB serves as the helicase that unwinds DNA to provide templates for both the leading and lagging strands. To accomplish this task, DnaB moves in the 5′→3′ direction on the lagging strand template—the same direction in which the replicating fork is moving. This anchors the primosome to the lagging strand template, where it is needed for priming Okazaki fragment synthesis.

SUMMARY Primer synthesis in *E. coli* requires a primosome composed of the DNA helicase, DnaB, and the primase, DnaG. Primosome assembly at the origin of replication, *oriC,* occurs as follows: DnaA binds to *oriC* at sites called *dnaA* boxes and cooperates with RNA polymerase and HU protein in melting a DNA region adjacent to the leftmost *dnaA* box. DnaB then binds to the open complex and facilitates binding of the primase to complete the primosome. The primosome remains with the replisome, repeatedly priming Okazaki fragment synthesis, at least on the lagging strand. DnaB also has a helicase activity that unwinds the DNA as the replisome progresses.

Priming in Eukaryotes

Eukaryotic replication is considerably more complex than the bacterial replication we have just studied. One complicating factor is the much bigger size of eukaryotic genomes. This, coupled with the slower movement of eukaryotic replicating forks, means that each chromosome must have multiple origins. Otherwise, replication would not finish within the time allotted—the S phase of the cell cycle—which can be as short as a few minutes. Because of this multiplicity and other factors, identification of eukaryotic origins of replication has lagged considerably behind similar work in prokaryotes. However, when molecular biologists face a complex problem, they frequently resort to simpler systems such as viruses to give them clues about the viruses' more complex hosts. Scientists followed this strategy to identify the origin of replication in the simple monkey virus SV40 as early as 1972. Let us begin our study of eukaryotic origins of replication there, then move on to origins in yeast.

The Origin of Replication in SV40 Two research groups, one headed by Norman Salzman, the other by Daniel

Nathans, identified the SV40 origin of replication in 1972 and showed that DNA replication proceeded bidirectionally from this origin. Salzman's strategy was to use *Eco*RI to cleave replicating SV40 DNA molecules at a unique site. (Although this enzyme had only a short time before been discovered and characterized, Salzman knew that SV40 DNA contained only a single *Eco*RI site.) After cutting the replicating SV40 DNA with *Eco*RI, Salzman and colleagues visualized the molecules by electron microscopy. They observed only a single replicating bubble, which indicated a single origin of replication. Furthermore, as they followed the growth of this bubble, they found that it grew at both ends, showing that both replicating forks were moving away from the single origin. This analysis revealed that the origin lies 33% of the genome length from the *Eco*RI site. But which direction from the *Eco*RI site? Because the SV40 DNA is circular, and these pictures contain no other markers besides the single *Eco*RI site, we cannot tell. But Nathans used another restriction enzyme (*Hin*dII), and his results, combined with these, placed the origin at a site overlapping the SV40 control region, adjacent to the GC boxes and the 72-bp repeat enhancer we discussed in Chapters 10 and 12 (Figure 21.2).

The minimal *ori* sequence (the *ori* core) is 64 bp long and includes several essential elements (1) four pentamers (5′-GAGGC-3′), which are the binding site for **large T antigen,** the major product of the viral early region; (2) a 15-bp palindrome, which is the earliest region melted during DNA replication; and (3) a 17-bp region consisting only of A–T pairs, which probably facilitates melting of the nearby palindrome region.

Other elements surrounding the *ori* core also participate in initiation. These include two additional large T antigen-binding sites, and the GC boxes to the left of the *ori* core. The GC boxes provide about a 10-fold stimulation of initiation of replication. If the number of GC boxes is reduced, or if they are moved only 180 bp away from *ori,* this stimulation is reduced or eliminated. This effect is somewhat akin to the participation of RNA polymerase in initiation at *ori*C in *E. coli.* One difference: At the SV40 *ori,* no transcription need occur; binding of the transcription factor Sp1 to the GC boxes is sufficient to stimulate initiation of replication.

Once large T antigen binds at the SV40 *ori,* its DNA helicase activity unwinds the DNA and prepares the way for primer synthesis. Just as in bacteria, eukaryotic primers are made of RNA. The primase in eukaryotic cells associates with DNA polymerase α, and this also serves as the primase for SV40 replication.

Figure 21.2 Location of the SV40 *ori* in the transcription control region. The core *ori* sequence (green) encompasses part of the early region TATA box and the cluster of early transcription initiation sites. Pink arrows denote bidirectional replication away from the replication initiation site. Black arrows denote transcription initiation sites.

SUMMARY The SV40 origin of replication is adjacent to the viral transcription control region. Initiation of replication depends on the viral large T antigen, which binds to a region within the 64-bp *ori* core, and at two adjacent sites, and exercises a helicase activity, which opens up a replication bubble within the *ori* core. Priming is carried out by a primase associated with the host DNA polymerase α.

The Origin of Replication in Yeast So far, yeast has provided most of our information about eukaryotic origins of replication. This is not surprising, because yeasts are among the simplest eukaryotes, and they lend themselves well to genetic analysis. As a result, yeast genetics are well understood. As early as 1979, C.L. Hsiao and J. Carbon discovered a yeast DNA sequence that could replicate independently of the yeast chromosomes, suggesting that it contains an origin of replication. This DNA fragment contained the yeast $ARG4^+$ gene. Cloned into a plasmid, it transformed $arg4^-$ yeast cells to $ARG4^+$, as demonstrated by their growth on medium lacking arginine. Any yeast cells that grew must have incorporated the $ARG4^+$ gene of the plasmid and, furthermore, must be propagating that gene somehow. One way to propagate the gene would be by incorporating it into the host chromosomes by recombination, but that was known to occur with a low frequency—about 10^{-6}–10^{-7}. Hsiao and Carbon obtained $ARG4^+$ cells at a much higher frequency—about 10^{-4}. Furthermore, shuttling the plasmid back and forth between yeast and *E. coli* caused no change in the plasmid structure, whereas recombination with the yeast genome would have changed it noticeably. Thus, these investigators concluded that the yeast DNA fragment they had cloned in the plasmid probably contained an origin of replication. Also in 1979, R.W. Davis and colleagues performed a similar study with a plasmid containing a yeast DNA fragment that converted trp^- yeast cells to TRP^+. They named the 850-bp yeast fragment **autonomously replicating sequence 1,** or **ARS1.**

Although these early studies were suggestive, they failed to establish that DNA replication actually begins in the ARS sequences. To demonstrate that ARS1 really does have this key characteristic of an origin of replication, Bonita Brewer and Walton Fangman used two-dimensional electrophoresis to detect the site of replication initiation in a plasmid bearing ARS1. This technique depends on the fact that circular and branched DNAs migrate more slowly than linear DNAs of the same size during gel electrophoresis, especially at high voltage or high agarose concentration.

Brewer and Fangman prepared a yeast plasmid bearing ARS1 as the only origin of replication. They allowed this plasmid to replicate in synchronized yeast cells and then isolated replication intermediates (RIs). They linearized these RIs with a restriction endonuclease, then electrophoresed them in the first dimension under conditions (low voltage and low agarose concentration) that separate DNA molecules roughly according to their sizes. Then they electrophoresed the DNAs in the second dimension using higher voltage and agarose concentrations that cause retardation of branched and circular molecules. Finally, they Southern blotted the DNAs in the gel and probed the blot with a labeled plasmid-specific DNA.

Figure 21.3 shows an idealized version of the behavior of various branched and circular RIs of a hypothetical 1-kb fragment. Simple Y's (panel a) begin as essentially linear 1-kb fragments with a tiny Y at their right ends; these would behave almost like linear 1-kb fragments. As the fork moves from right to left, the Y grows larger and the mobility of the fragment in the second (vertical) dimension slows. Then, as the Y grows even larger, the fragment begins to look more and more like a linear 2-kb fragment, with just a short stem on the Y. This is represented by the horizontal linear form with a short vertical stem in panel (a). Because these forms resemble linear shapes more and more as the fork moves, their mobility increases correspondingly, until the fork has nearly reached the end of the fragment. At this point, they have a shape and mobility that is almost like a true linear 2-kb fragment. This behavior gives rise to an arc-shaped pattern, where the apex of the arc corresponds to a Y that is half-replicated, at which point it is least like a linear molecule.

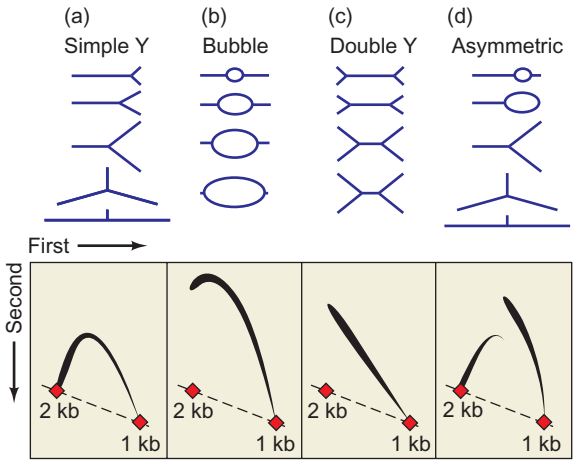


Figure 21.3 Theoretical behaviors of various types of replication intermediates on two-dimensional gel electrophoresis. The top parts of panels **a–d** are cartoons showing the shapes of growing simple Y's, bubbles, double Y's, and asymmetric bubbles that convert to simple Y's as replication progresses. The bottom parts of each panel are cartoons that depict the expected deviation of the changing mobilities of each type of growing RI from the mobilities of linear forms growing progressively from 1 to 2 kb (dashed lines). (*Source*: Adapted from Brewer, B.J. and W.L. Fangman, The localization of replication origins on ARS plasmids in *S. cerevisiae*. *Cell* 51:464, 1987.)

Figure 21.3b shows what to expect for a bubble-shaped fragment. Again, we begin with a 1-kb linear fragment, but this time with a tiny bubble right in the middle. As the bubble grows larger, the mobility of the fragment slows more and more, yielding the arc shown at the bottom of the panel. Panel (c) shows the behavior of a double Y, where the RI becomes progressively more branched as the two forks approach the center of the fragment. Accordingly, the mobility of the RI decreases almost linearly. Finally, panel (d) shows what happens to a bubble that is asymmetrically placed in the fragment. It begins as a bubble, but then, when one fork passes the restriction site at the right end of the fragment, it converts to a Y. The mobilities of the RIs reflect this discontinuity: The curve begins like that of a bubble, then abruptly changes to that of a Y, with an obvious discontinuity showing exactly when the fork passed the restriction site and converted the bubble to a Y.

This kind of behavior is especially valuable in mapping the origin of replication. In panel (d), for example, we can see that the discontinuity occurs in the middle of the curve, when the mobility in the first dimension was that of a 1.5-kb fragment. This tells us that the arms of the Y are each 500 bp long. Assuming that the two forks are moving at an equal rate, we can conclude that the origin of replication was 250 bp from the right end of the fragment.

Now let us see how this works in practice. Brewer and Fangman chose restriction enzymes that would cleave the plasmid with its ARS1 just once, but in locations that would be especially informative if the origin of replication really lies within ARS1. Figure 21.4 shows the locations of the two restriction sites, at top, and the experimental results, at bottom. The first thing to notice about the autoradiographs is that they are simple and correspond to the patterns we have seen in Figure 21.3. This means that there is a single origin of replication; otherwise, there would have been a mixture of different kinds of RIs, and the results would have been more complex.

The predicted origin within ARS1 lies adjacent to a *Bgl*II site (B, in panel a). Thus, if the RI is cleaved with this enzyme, it should yield double-Y RIs. Indeed, as we see in the lower part of panel (a), the autoradiograph is nearly linear—just as we expect for a double-Y RI. Panel (b) shows that a *Pvu*I site (P) lies almost halfway around the plasmid from the predicted origin. Therefore, cleaving with *Pvu*I should yield the bubble-shaped RI shown at the top of panel (b). The autoradiograph at the bottom of panel (b) shows that Brewer and Fangman observed the discontinuity expected for a bubble-shaped RI that converts at the very end to a very large single Y, as one fork reaches the *Pvu*I site, then perhaps to a very asymmetric double Y as the fork passes that site. Both of these results place the origin of replication adjacent to the *Bgl*II site, just where we expect it if ARS1 contains the origin.

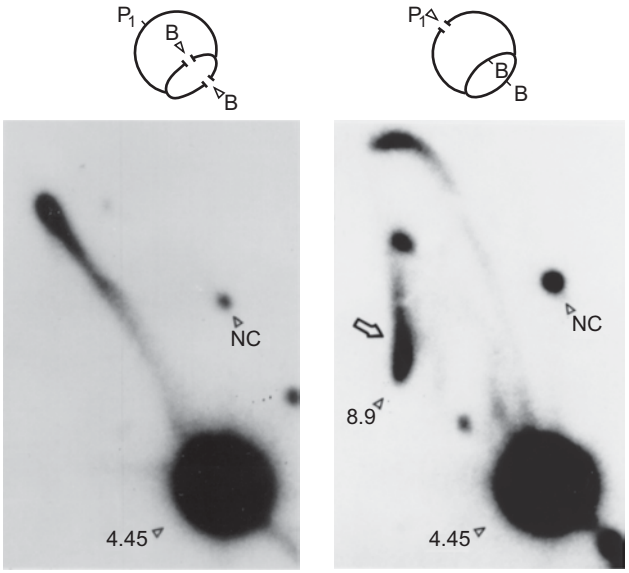


Figure 21.4 Locating the origin of replication in ARS1. (a) Results of cleaving 2-μm plasmid with *Bgl*II. Top: cartoon showing the shape expected when an RI is cut with *Bgl*II, assuming the origin lies adjacent to the *Bgl*II site within ARS1. The bubble contains DNA that has already replicated, so there are two copies of the *Bgl*II site (arrowheads labeled B), both of which are cut to yield the double-Y intermediate depicted. Bottom: experimental results showing the straight curve expected of double-Y intermediates. **(b)** Results of cleaving the plasmid with *Pvu*I. Top: cartoon showing the shape expected when an RI is cut with *Pvu*I, assuming the origin lies almost across the circle from the *Pvu*I site within ARS1. Bottom: experimental results showing the rising arc, with a discontinuity near the end. This is what we expect for a bubble-shaped RI that converts to a nearly linear Y as one of the replication forks passes a *Pvu*I site. Both of these results confirm the expectations for an origin of replication within ARS1. NC denotes nicked circles. The large open arrow points to large Y's or very asymmetric double Y's that result when a replicating fork passes a *Pvu*I site. Numbers refer to sizes in kb. (*Source:* Brewer, B.J. and W.L. Fangman, The localization of replication origins on ARS plasmids in *S. cerevisiae. Cell* 51 (6 Nov 1987) f. 8, p. 469. Reprinted by permission of Elsevier Science.)

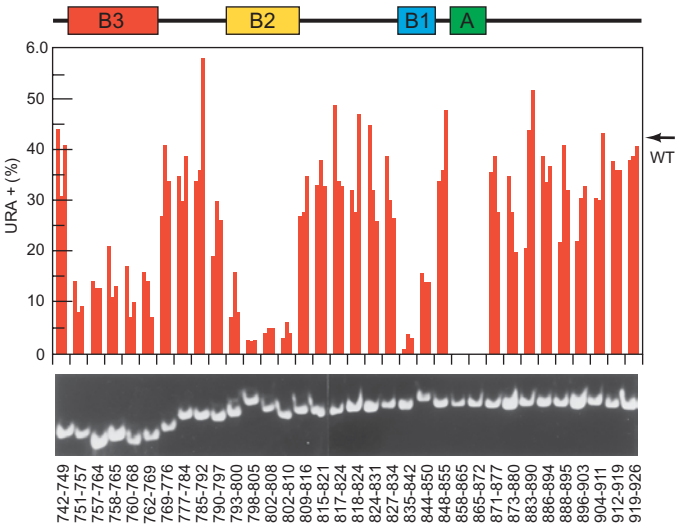


Figure 21.5 Linker scanning analysis of ARS1. Marahrens and Stillman substituted linkers throughout an ARS1 sequence within a plasmid bearing a yeast centromere and the *URA3* selectable marker. To test for replication efficiency of the mutants, they grew them for 14 generations in nonselective medium, then tested them for growth on selective (uracil-free) medium. The vertical bars show the results of three independent determinations for each mutant plasmid. Results are presented as a percentage of the yeast cells that retained the plasmid (as assayed by their ability to grow). Note that even the wild-type plasmid was retained with only 43% efficiency in nonselective medium (arrow at right). Four important regions (A, B1, B2, and B3) were identified. The regions that were mutated are identified by base number at bottom. The stained gel at bottom shows the electrophoretic mobility of each mutant plasmid. Note the altered mobility of the B3 mutant plasmids, which suggests altered bending. (*Source:* From Marahrens, Y. and B. Stillman, A yeast chromosomal origin of DNA replication defined by multiple functional elements. *Science* 255 (14 Feb 1992) f. 2, p. 819. Copyright © AAAS. Reprinted with permission from AAAS.)

York Marahrens and Bruce Stillman performed linker scanning experiments to define the important regions within ARS1. They constructed a plasmid very similar to the one used by Brewer and Fangman, containing (1) ARS1 in a 185-bp DNA sequence; (2) a yeast centromere; and (3) a selectable marker—*URA3*—which confers on *ura3-52* yeast cells the ability to grow in uracil-free medium. Then they performed linker scanning (Chapter 10) by systematically substituting an 8-bp *Xho*I linker for the normal DNA at sites spanning the ARS1 region. They transformed yeast cells with each of the linker scanning mutants and selected for transformed cells with uracil-free medium. Some of the transformants containing mutant ARS1 sequences grew more slowly than those containing wild-type ARS1 sequences. Because the centromere in each plasmid ensured proper segregation of the plasmid, the most likely explanation for poor growth was poor replication due to mutation of ARS1.

To check this hypothesis, Marahrens and Stillman grew all the transformants in a nonselective medium containing uracil for 14 generations, then challenged them again with a uracil-free medium to see which ones had not maintained the plasmid well. The mutations in these unstable plasmids presumably interfered with ARS1 function. Figure 21.5 shows the results. Four regions of ARS1 appear to be important. These were named A, B1, B2, and B3 in order of decreasing effect on plasmid stability. Element A is 15 bp long, and contains an 11-bp ARS consensus sequence:

$$5'\text{-}{}^{\mathrm{T}}_{\mathrm{A}}\mathrm{TTTA}^{\mathrm{TA}}_{\mathrm{CG}}\mathrm{TTT}^{\mathrm{T}}_{\mathrm{A}}\text{-}3'$$

When it was mutated, *all* ARS1 activity was lost. The other regions had a less drastic effect, especially in selective medium. However, mutations in B3 had an apparent effect on the bending of the plasmid, as assayed by gel electrophoresis. The stained gel below the bar graph shows increased electrophoretic mobility of the mutants in the B3 region. Marahrens and Stillman interpreted this as altered bending of the ARS1 in the presence of the replicating machinery.

The existence of four important regions within ARS1 raises the question whether these are also *sufficient* for ARS function. To find out, Marahrens and Stillman constructed a synthetic ARS1 with wild-type versions of all four regions, spaced just as in the wild-type ARS1, but with random sequences in between. A plasmid bearing this synthetic ARS1 was almost as stable under nonselective conditions as one bearing a wild-type ARS1. Thus, the four DNA elements defined by linker scanning are sufficient for ARS1 activity. Finally, these workers replaced the wild-type 15-bp region A with the 11-bp ARS consensus sequence. This reduced plasmid stability dramatically, suggesting that the other 4 bp in region A are also important for ARS activity.

SUMMARY The yeast origins of replication are contained within autonomously replicating sequences (ARSs) that are composed of four important regions (A, B1, B2, and B3). Region A is 15 bp long and contains an 11-bp consensus sequence that is highly conserved in ARSs. Region B3 may allow for an important DNA bend within ARS1.

21.2 Elongation

Once a primer is in place, real DNA synthesis (elongation) can begin. We have already identified the pol III holoenzyme as the enzyme that carries out elongation in *E. coli*, and DNA polymerases δ and ε as the enzymes that elongate the lagging and leading strands, respectively, in eukaryotes. The *E. coli* system is especially well characterized, and the data point to an elegant method of coordinating the synthesis of lagging and leading strands in a way that keeps the pol III holoenzyme engaged with the template so replication can be highly processive, and therefore very rapid. Let us focus on this *E. coli* elongation mechanism, beginning with a discussion of the speed of elongation.

Speed of Replication

Minsen Mok and Kenneth Marians performed one of the studies that measured the rate of fork movement in vitro with the pol III holoenzyme. They created a synthetic circular template for rolling circle replication, illustrated in Figure 21.6. This template contained a ^{32}P-labeled, tailed, full-length strand with a free 3′-hydroxyl group for priming. Mok and Marians incubated this template with either holoenzyme plus preprimosomal proteins and SSB, or plus DnaB helicase alone. At 10-sec intervals, they removed the labeled product DNAs and measured their lengths by electrophoresis. Panels (a) and (b) in Figure 21.7 depict the results with the two reactions, and Figure 21.7c shows plots of the rates of fork movement with the two reactions.

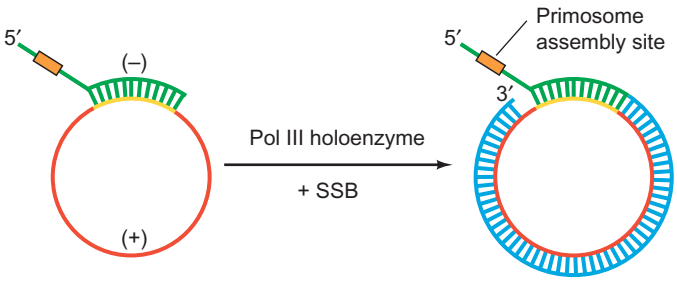


Figure 21.6 Synthesis of template used to measure fork velocity in vitro. Mok and Marians started with the 6702-nt positive strand (red) from the f1 phage and annealed it to a primer (green) that hybridized over a 282-nt region (yellow). This primer contained a primosome assembly site (orange). Mok and Marians elongated the primer with pol III holoenzyme and single-strand binding protein (SSB) to create the negative strand (blue). The product was a double-stranded template for multiple rounds of rolling circle replication, in which the free 3′-end could serve as the primer. (*Source:* Adapted from Mok, M. and K.J. Marians, The *Escherichia coli* preprimosome and DNA B helicase can form replication forks that move at the same rate. *Journal of Biological Chemistry* 262:16645, 1987.)

Both plots yielded rates of 730 nt/sec, close to the in vivo rate of almost 1000 nt/sec.

Furthermore, the elongation in these reactions with holoenzyme was highly processive. As we have mentioned, processivity is the ability of the enzyme to stick to its job a long time without falling off and having to reinitiate. This is essential because reinitiation is a time-consuming process, and little time can be wasted in DNA replication. To measure processivity, Mok and Marians performed the same elongation assay as described in Figure 21.7, but included either of two substances that would prevent reinitiation if the holoenzyme dissociated from the template. These substances were a competing DNA, poly(dA), and an antibody directed against the β-subunit of the holoenzyme. In the presence of either of these competitors, the elongation rate was just as fast as in their absence, indicating that the holoenzyme did not dissociate from the template throughout the process of elongation of the primer by at least 30 kb. Thus, the holoenzyme is highly processive in vitro, just as it is in vivo.

SUMMARY The pol III holoenzyme synthesizes DNA at the rate of about 730 nt/sec in vitro, just a little slower than the rate of almost 1000 nt/sec observed in vivo. This enzyme is also highly processive, both in vitro and in vivo.

The Pol III Holoenzyme and Processivity of Replication

The pol III core by itself is a very poor polymerase. It puts together about 10 nt and then falls off the template. Then it has to spend about a minute reassociating with the

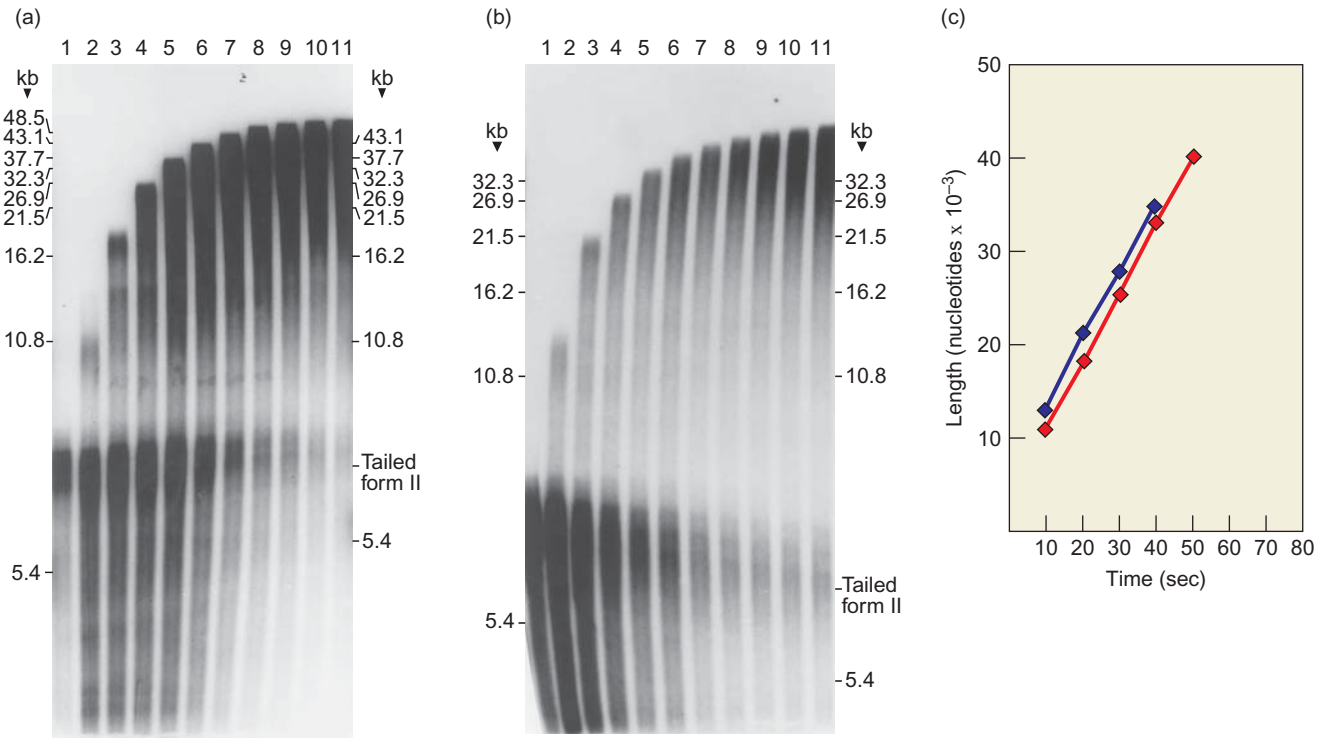


Figure 21.7 Measuring the rate of fork movement in vitro. Mok and Marians labeled the negative strand of the tailed template in Figure 21.6 and used it in in vitro reactions with pol III holoenzyme plus: **(a)** the preprimosomal proteins (the primosomal proteins minus DnaG); or **(b)** DnaB alone. They took samples from the reactions at 10-sec intervals, beginning with lanes 1 at zero time and lanes 2 at 10 sec, electrophoresed them, and then autoradiographed the gel. Recall that electrophoretic mobilities are a log function, not a linear function, of mass. The numbers on the left in each panel are marker sizes, not the sizes of DNA products. Panel **(c)** shows a plot of the results from the first five and four time points from panels (a) (red) and (b) (blue), respectively. (*Source:* Mok M. and K.J. Marians, The *Escherichia coli* preprimosome and DNA B helicase can form replication forks that move at the same rate. *Journal of Biological Chemistry* 262 no. 34 (5 Dec 1987) f. 6a–b, p. 16650. Copyright © American Society for Biochemistry and Molecular Biology.)

template and the nascent DNA strand. This contrasts sharply with the situation in the cell, where the replicating fork moves at the rate of almost 1000 nt/sec. Obviously, something important is missing from the core.

That "something" is an agent that confers processivity on the holoenzyme, allowing it to remain engaged with the template while polymerizing at least 50,000 nt before stopping—quite a contrast to the 10 nt polymerized by the core before it stops. Why such a drastic difference? The holoenzyme owes its processivity to a **"sliding clamp"** that holds the enzyme on the template for a long time. The β-subunit of the holoenzyme performs this sliding clamp function, but it cannot associate by itself with the preinitiation complex (core plus DNA template). It needs a **clamp loader** to help it join the complex, and a group of subunits called the **γ complex** provides this help. The γ complex includes the γ-, δ-, δ′-, χ-, and ψ-subunits. In this section, we will examine the activities of the β clamp and the clamp loader.

The β clamp One way we can imagine the β-subunit conferring processivity on the pol III core is by binding both the core complex and DNA. That way, it would tie the core to the DNA and keep it there—hence the term **β clamp.** In the course of probing this possibility, Mike O'Donnell and colleagues demonstrated direct interaction between the β- and α-subunits. They mixed various combinations of subunits, then separated subunit complexes from individual subunits by gel filtration. They detected subunits by gel electrophoresis, and activity by adding the missing subunits and measuring DNA synthesis. Figure 21.8 depicts the electrophoresis results. It is clear that α and ε bind to each other, as we would expect, because they are both part of the core. Furthermore, α, ε, and β form a complex, but which subunit does β bind to, α or ε? Panels (d) and (e) show the answer: β binds to α alone (both subunits peak in fractions 60–64), but not to ε alone (β peaks in fractions 68–70, whereas ε peaks in fractions 76–78). Thus, α is the core subunit to which β binds.

This scheme demands that β be able to slide along the DNA as α and ε together replicate it. This in turn suggests that the β clamp would remain bound to a circular DNA, but could slide right off the ends of a linear DNA. To test this possibility, O'Donnell and colleagues performed the experiment reported in Figure 21.9. The general strategy of this experiment was to load ^{3}H-labeled β dimers onto circular, double-stranded phage DNA with the help of the γ complex, then to treat the DNA in various ways to see if

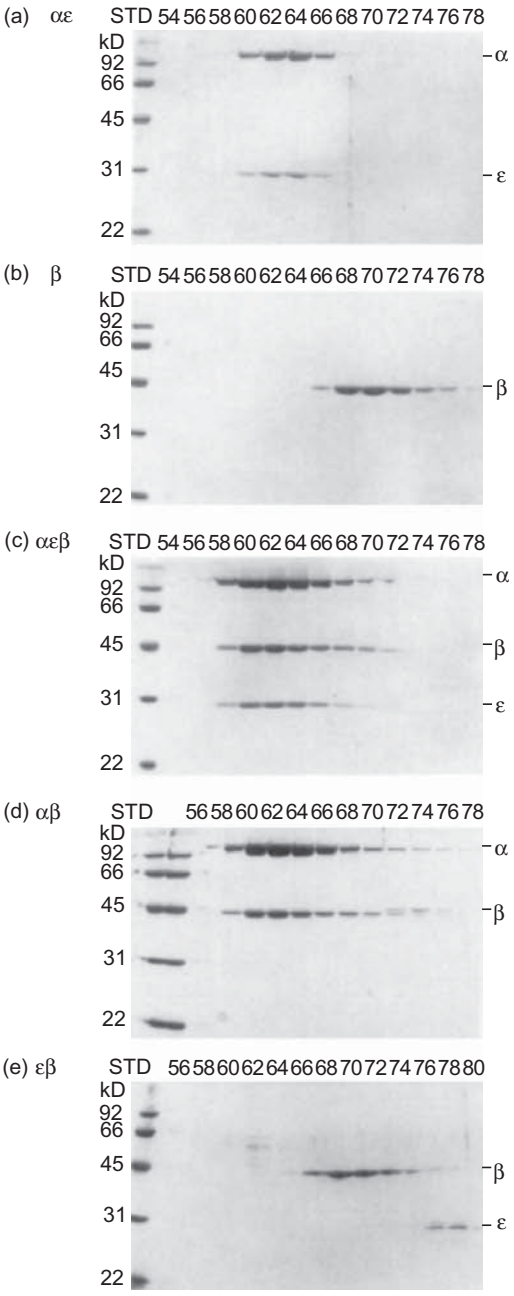


Figure 21.8 The Pol III subunits α and β bind to each other. O'Donnell and colleagues mixed various combinations of pol III subunits as follows: **(a)** α+ε; **(b)** β; **(c)** α+ε+β; **(d)** α+β; **(e)** ε+β. Then they subjected the mixtures to gel filtration to separate complexes from free subunits, then electrophoresed fractions from the gel filtration column to detect complexes. If a complex formed, the subunits in the complex should appear in the same fractions, as the α and ε fractions do in panel (a). (*Source:* Stukenberg, P.T., P.S. Studwell-Vaughn, and M. O'Donnell, Mechanism of the sliding β-clamp of DNA polymerase III holoenzyme. *Journal of Biological Chemistry* 266 no. 17(15 June 1991) figs. 2a–e, 3, pp. 11330–31. American Society for Biochemistry and Molecular Biology.)

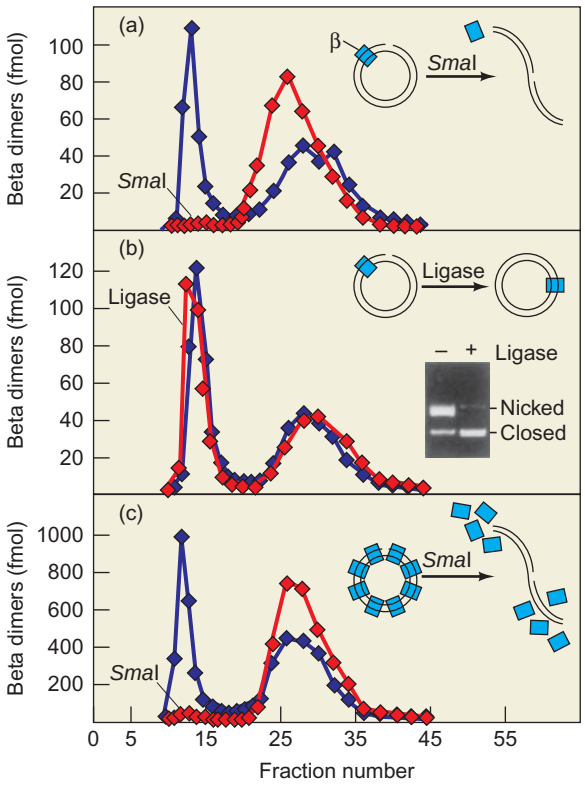


Figure 21.9 The β clamp can slide off the ends of a linear DNA. O'Donnell and colleagues loaded ^{3}H-labeled β dimers onto various DNAs, with the help of the γ complex, then treated the complexes in various ways as described. Finally, they subjected the mixtures to gel filtration to separate protein–DNA complexes (which were large and eluted quickly from the column, around fraction 15), from free protein (which was relatively small and eluted later, around fraction 28). **(a)** Effect of linearizing the DNA with *Sma*I. DNA was cut once with *Sma*I and then assayed (red). Uncut DNA was also assayed (blue). **(b)** Effect of removing a nick in the template. The nick in the template was removed with DNA ligase before assay (red), or left alone (blue). The inset shows the results of electrophoresis of DNAs before and after the ligase reaction. **(c)** Many β dimers can be loaded onto the DNA and then lost when it is linearized. The ratio of β dimers loaded onto DNA templates was increased by raising the concentration of β-subunits and lowering the concentration of DNA templates. Then the DNA was either cut with *Sma*I before assay (red) or not cut (blue). (*Source:* Stukenberg, P.T., P.S. Studwell-Vaughn, and M. O'Donnell, Mechanism of the sliding β-clamp of DNA polymerase III holoenzyme. *Journal of Biological Chemistry* 266 no. 17 (15 June 1991) fig. 3, p. 11331. American Society for Biochemistry and Molecular Biology.)

the β dimers could dissociate from the DNA. The assay for β-binding to DNA was gel filtration. Independent β dimers emerge from a gel filtration column much later than they do when they are bound to DNA.

In panel (a), the DNA was treated with *Sma*I to linearize the DNA, then examined to see whether the β clamp had slid off. It remained bound to circular DNA, but had dissociated from linearized DNA, apparently by sliding off the ends. Panel (b) demonstrates that the nick in the circular DNA is not what caused retention of the β dimer, because the nick can be removed with DNA ligase, and the

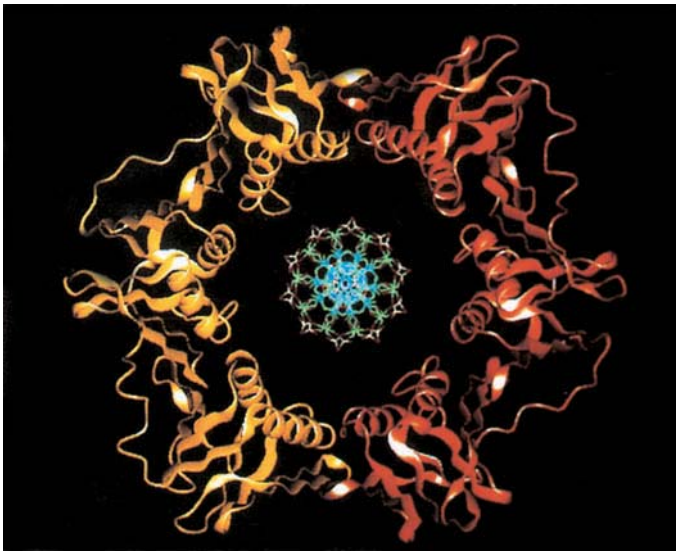


Figure 21.10 Model of the β dimer/DNA complex. The β dimer is depicted by a ribbon diagram in which the α-helices are coils and the β-sheets are flat ribbons. One β monomer is yellow and the other is red. A DNA model, seen in cross section, is placed in a hypothetical position in the middle of the ring formed by the β dimer. (*Source:* Kong, X.-P., R. Onrust, M. O'Donnell, and J. Kuriyan, Three-dimensional structure of the beta subunit of *E. coli* DNA polymerase III holoenzyme: A sliding DNA clamp. *Cell* 69 (1 May 1992) f. 1, p. 426. Reprinted by permission of Elsevier Science.)

β dimer remains bound to the DNA. The inset shows electrophoretic evidence that the ligase really did remove the nick because the nicked form disappeared and the closed circular form was enhanced. Panel (c) shows that adding more β-subunit to the loading reaction increased the number of β dimers bound to the circular DNA. In fact, more than 20 molecules of β-subunit could be bound per molecule of circular DNA. This is what we would expect if many holoenzymes can replicate the DNA in tandem.

If the β dimers are lost from linear DNA by sliding off the ends, one ought to be able to prevent their loss by binding other proteins to the ends of the DNA. O'Donnell's group did this in experiments, not shown here, by binding two different proteins to the ends of the DNA and demonstrating that the β dimers no longer fell off. Indeed, single-stranded tails at the ends of the DNA, even without protein attached, proved to be an impediment to the β dimers sliding off.

Mike O'Donnell and John Kuriyan used x-ray crystallography to study the structure of the β clamp. The pictures they produced provided a perfect rationale for the ability of the β clamp to remain bound to a circular DNA but not to a linear one: The β dimer forms a ring that can fit around the DNA. Thus, like a ring on a string, it can readily fall off if the string is linear, but not if the string is circular. Figure 21.10 is one of the models O'Donnell and Kuriyan constructed; it shows the ring structure of the β dimer, with a scale model of B-form DNA placed in the middle.

In 2008, O'Donnell and colleagues obtained the structure of a co-crystal of a β dimer bound to a primed DNA

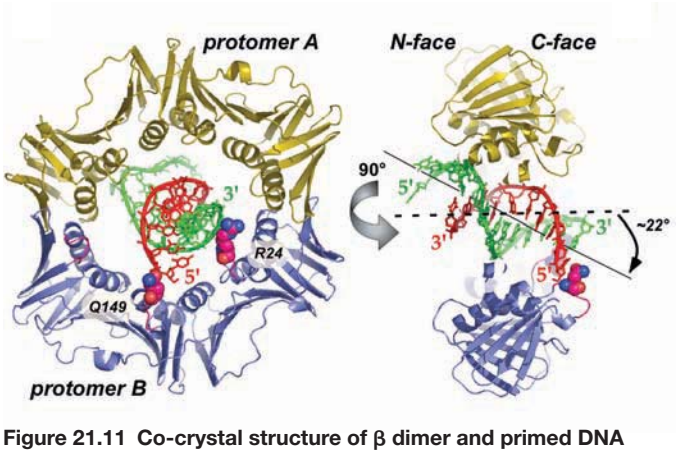


Figure 21.11 Co-crystal structure of β dimer and primed DNA template. The two β monomers (protomers A and B) are in gold and blue, with the primed DNA template in green and red. Magenta and blue space-filling models show the side chains of arginine 24 (R24) and glutamine 149 (Q149). The structure on the left is a front view; the structure on the right is a side view, which emphasizes the 22-degree tilt of the DNA. (*Source:* Georgescu et al., Structure of a sliding clamp on DNA. *Cell* 132 (11 January 2008) f. 3a, p. 48. Reprinted by permission of Elsevier Science.)

template. Figure 21.11 shows this crystal structure, which demonstrates that the β clamp really does encircle the DNA, as the model in Figure 21.10 predicted. However, this newer structure shows the actual geometry of DNA within the β clamp, and it contains a bit of a surprise: Instead of extending straight through the β clamp, like a finger through a ring, the DNA is tilted about 22 degrees with respect to a horizontal line through the clamp. Furthermore, the DNA contacts the side chains of two amino acids, arginine 24 and glutamine 149, both of which lie on the C-terminal face of the β clamp. This protein–DNA contact probably contributes to the tilt of the DNA with respect to the β dimer.

As mentioned in Chapter 20, eukaryotes also have a processivity factor called PCNA, which performs the same function as the bacterial β clamp. The primary structure of PCNA bears no apparent similarity to that of the β clamp, and the eukaryotic protein is only two-thirds the size of its prokaryotic counterpart. Nevertheless, x-ray crystallography performed by Kuriyan and his colleagues demonstrated that yeast PCNA forms a trimer with a structure arrestingly similar to that of the β clamp dimer: a ring that can encircle a DNA molecule, as shown in Figure 21.12.

SUMMARY The Pol III core (αε or αεθ) does not function processively by itself, so it can replicate only a short stretch of DNA before falling off the template. By contrast, the core plus the β-subunit can replicate DNA processively at a rate approaching 1000 nt/sec. The β-subunit forms a dimer that is ring-shaped. This ring fits around a DNA template and interacts with the α-subunit of the core to tether

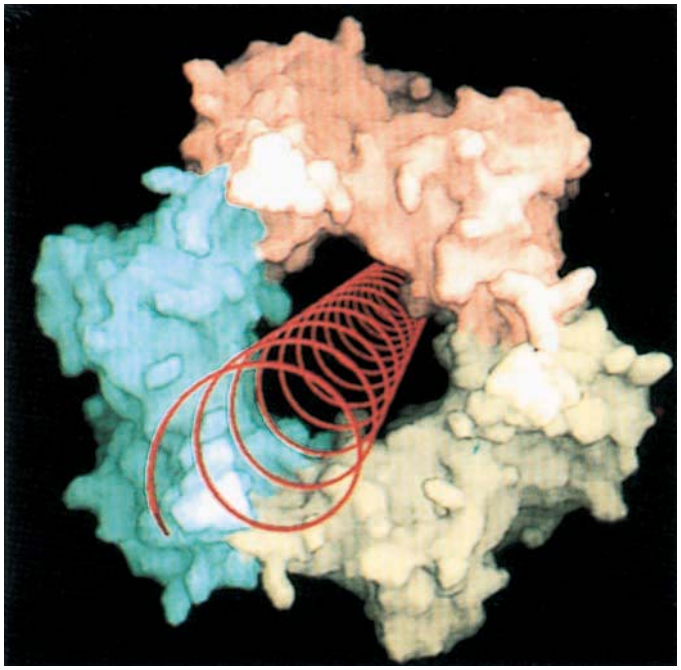


Figure 21.12 Model of PCNA–DNA complex. Each of the monomers of the PCNA trimer is represented by a different pastel color. The shape of the trimer is based on x-ray crystallography analysis. The red helix represents the probable location of the sugar–phosphate backbone of a DNA associated with the PCNA trimer. (*Source:* Krishna, T.S.R., X.-P. Kong, S. Gary, P.M. Burgers, and J. Kuriyan, Crystal structure of the eukaryotic DNA polymerase processivity factor PCNA. *Cell* 79 (30 Dec 1994) f. 3b, p. 1236. Reprinted by permission of Elsevier Science.)

the whole polymerase and template together. This is why the holoenzyme stays on its template so long and is therefore so processive. The eukaryotic processivity factor PCNA forms a trimer with a similar ring shape that can encircle DNA and hold DNA polymerase on the template.

The Clamp Loader O'Donnell and his colleagues demonstrated the function of the clamp loader in an experiment presented in Figure 21.13. These scientists used the α- and ε-subunits instead of the whole core, because the θ-subunit was not essential in their in vitro experiments. As template, they used a single-stranded M13 phage DNA annealed to a primer. They knew that highly processive holoenzyme could replicate this DNA in about 15 sec but that the αε core could not give a detectable amount of replication in that time. Thus, they reasoned that a 20-sec pulse of replication would allow all processive polymerase molecules the chance to complete one cycle of replication, and therefore the number of DNA circles replicated would equal the number of processive polymerases. Figure 21.13a shows that each femtomole (fmol, or 10^{-15} mol) of γ complex resulted in about 10 fmol of circles replicated in the presence of αε core and β-subunit. Thus, the γ complex acts

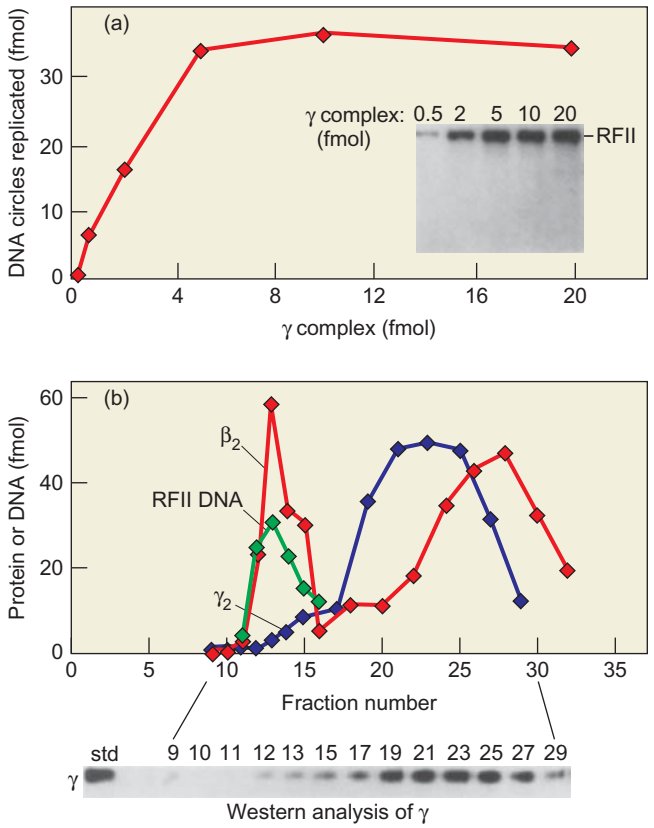


Figure 21.13 Involvement of β and γ complex in processivity. **(a)** The γ complex acts catalytically in forming a processive polymerase. O'Donnell and coworkers added increasing amounts of γ complex (indicated on the *x* axis) to a primed M13 phage DNA template coated with SSB, along with αε core, and the β-subunit of pol III holoenzyme. Then they allowed a 20-sec pulse of DNA synthesis in the presence of [α-^{32}P]ATP to label the DNA product. They determined the radioactivity of part of each reaction and converted this to fmol of DNA circles replicated. To check for full circle replication, they subjected another part of each reaction to gel electrophoresis. The inset shows the result: The great majority of each product is full-circle size (RFII). **(b)** The β-subunit, but not the γ complex associates with DNA in the preinitiation complex. O'Donnell and colleagues added ^{3}H-labeled β-subunit and unlabeled γ complex to primed DNA coated with SSB, along with ATP to form a preinitiation complex. Then they subjected the mixture to gel filtration to separate preinitiation complexes from free proteins. They detected the β-subunit in each fraction by radioactivity, and the γ complex by Western blotting, with an anti-γ antibody as probe (bottom). The plot shows that the β-subunit (as dimers) bound to the DNA in the preinitiation complex, but the γ complex did not. (*Source:* Stukenberg, P.T., P.S. Studwell-Vaughn, and M. O'Donnell, Mechanism of the sliding [beta]-clamp of DNA polymerase III holoenzyme. *Journal of Biological Chemistry* 266 (15 June 1991) f. 1a&c, p. 11329. American Society for Biochemistry and Molecular Biology.)

catalytically: One molecule of γ complex can sponsor the creation of many molecules of processive polymerase. The inset in this figure shows the results of gel electrophoresis of the replication products. As expected of processive replication, they are all full-length circles.

This experiment suggested that the γ complex itself is not the agent that provides processivity. Instead, the γ complex could act catalytically to add something else to the core

polymerase that makes it processive. Because β was the only other polymerase subunit in this experiment, it is the likely processivity-determining factor. To confirm this, O'Donnell and colleagues mixed the DNA template with ^{3}H-labeled β-subunit and unlabeled γ complex to form preinitiation complexes, then subjected these complexes to gel filtration to separate the complexes from free proteins. They detected the preinitiation complexes by adding αε to each fraction and assaying for labeled double-stranded circles formed (RFII, green). Figure 21.13b demonstrates that only a trace of γ complex (blue) remained associated with the DNA, but a significant fraction of the labeled β-subunit (red) remained with the DNA. (The unlabeled γ complex was detected with a Western blot using an anti-γ antibody, as shown at the bottom of the figure.) It is important to note that, even though the γ complex does not remain bound to the DNA, it plays a vital role in processivity by loading the β-subunit onto the DNA.

This experiment also allowed O'Donnell and colleagues to estimate the stoichiometry of the β-subunit in the preinitiation complex. They compared the fmol of β with the fmol of complex, as measured by the fmol of double-stranded circles produced. This analysis yielded a value of about 2.8 β-subunits/complex, which would be close to one β *dimer*/complex, in accord with other studies that suggested that β acts as a dimer.

Implicit in the discussion so far is the fact that ATP is required to load the β clamp onto the template. Peter Burgers and Kornberg demonstrated the necessity for ATP (or dATP) with an assay that did not require dATP for replication. The template in this case was poly(dA) primed with oligo(dT). The results showed that ATP or dATP is required for high-activity elongation of the oligo(dT) primer with dTMP.

How does the clamp loader pry apart the β dimer to allow it to clamp around DNA? O'Donnell, Kuriyan, and colleagues have determined the crystal structures of two complexes that give strong hints about how the clamp loader works. One of these was the structure of the active part of the clamp loader (a γδδ′ complex). The other was the structure of a modified β–δ complex composed of: a monomer of a mutant form of β (β_{mt}) that is unable to dimerize; and a fragment of δ that can interact with β.

The crystal structure of this modified β–δ complex showed that the interaction between δ and a β monomer would be expected to weaken the binding at one interface between the two β monomers in two ways. First, δ acts as a molecular wrench by inducing a conformational change in the β dimer interface such that it no longer dimerizes as readily. Second, δ changes the curvature of one β-subunit so that it no longer naturally forms a ring with the other subunit. Instead, it forms a structure that resembles a lock washer. Figure 21.14 illustrates these concepts. Notice that δ binds to only one β monomer in the β clamp (there is only one δ per β dimer in the pol III holoenzyme), so it weakens only one dimer interface, and therefore forces ring opening. If δ bound to both β monomers, it would presumably cause the two monomers to dissociate entirely.

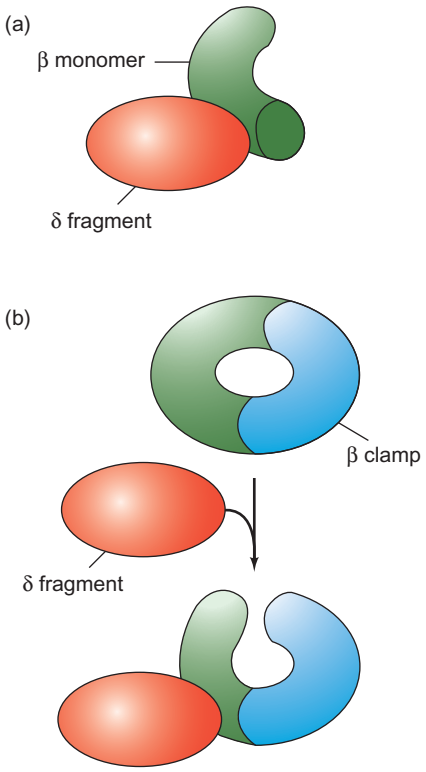


Figure 21.14 Model for the effect of δ binding on the β dimer. **(a)** Shape of the complex between the δ fragment and the β_{mt} monomer. **(b)** Effect of δ binding on the β clamp. The δ-subunit (or the δ fragment) causes the β dimer interface at the top to weaken and also changes the curvature of the β monomer on the left such that it can no longer form a complete circle with the other monomer. The result is an opening of the clamp. (*Source:* Adapted from Ellison, V. and B. Stillman, Opening of the clamp: An intimate view of an ATP-driven biological machine. *Cell* 106 [2001] p. 657, f. 3.)

These structural studies and earlier biochemical studies, some of which we will discuss later in this chapter, showed that δ on its own binds readily to a β monomer, but that δ in the context of the clamp loader complex cannot bind to the β clamp unless ATP is present. So the role of ATP appears to be to change the shape of the clamp loader to expose the δ-subunit so it can bind to one of the β-subunits and pry open the β clamp.

SUMMARY The β-subunit needs help from the γ complex (γ, δ, δ′, χ, and ψ) to load onto the DNA template. The γ complex acts catalytically in forming this processive αδβ complex, so it does not remain associated with the complex during processive replication. Clamp loading is an ATP-dependent process. The energy from ATP changes the conformation of the clamp loader such that the δ-subunit can bind to one of the β-subunits of the clamp. This binding opens the clamp and allows it to encircle DNA.

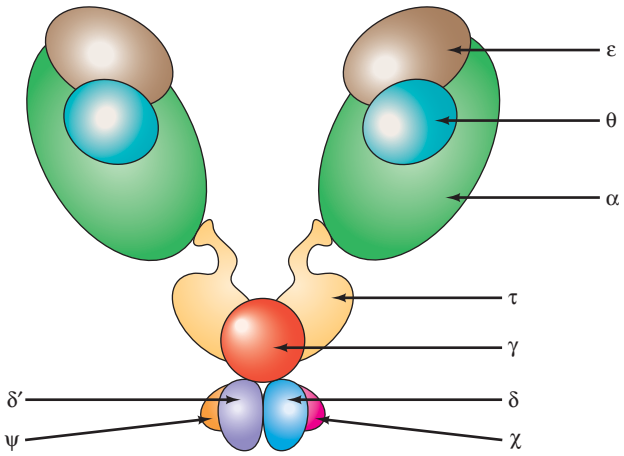


Figure 21.15 Model of the Pol III* subassembly. Note that two cores and two τ-subunits are present, but only one γ-complex (γ, δ, δ′, χ, and ψ). The τ-subunits are joined to the cores by their flexible C-terminal domains.

Lagging Strand Synthesis Structural studies on pol III* (holoenzyme minus the β clamp) have shown that the enzyme consists of two core polymerases, linked through a dimer of the τ-subunit to a clamp loader, as illustrated in Figure 21.15. The following reasoning suggests that the τ-subunit serves as a dimerizing agent for the core enzyme: The α-subunit is a monomer in its native state, but τ is a dimer. Furthermore, τ binds directly to α, so α is automatically dimerized by binding to the two τ-subunits. In turn, ε is dimerized by binding to the two α-subunits, and θ is dimerized by binding to the two ε-subunits. The two τ-subunits are products of the same gene that produces the γ-subunit. However, the γ-subunit lacks a 24-kDa domain (τ_c) at the C-terminus of the τ-subunits because of a programmed frameshift during translation. The two τ_c domains provide flexible linkers between the core polymerases and the γ complex.

The fact that the holoenzyme contains two core polymerases fits very nicely with the fact that two DNA strands need to be replicated. This leads directly to the suggestion that each of the core polymerases replicates one of the strands as the holoenzyme follows the moving fork. This is straightforward for the core polymerase replicating the leading strand, as that replication moves in the same direction as the fork. But it is more complicated for the core polymerase replicating the lagging strand, because that replication occurs in the direction opposite to that of the moving fork. This means that the lagging strand must form a loop, as pictured in Figure 21.16. Because this loop extends as an Okazaki fragment grows and then retracts to begin synthesis of a new Okazaki fragment, the loop resembles the slide of a trombone, and this model is sometimes called the "trombone model."

Because discontinuous synthesis of the lagging strand must involve repeated dissociation and reassociation of the

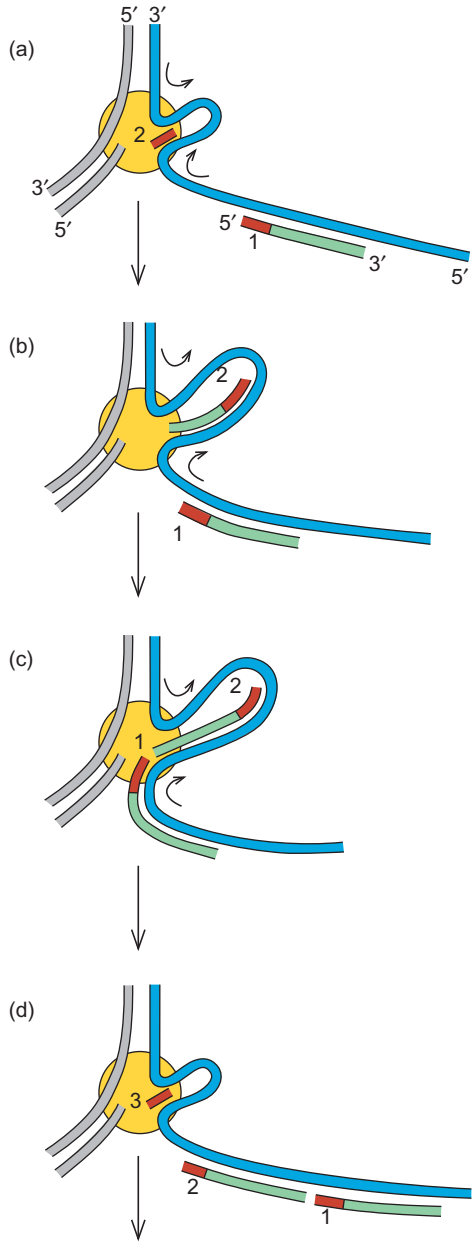


Figure 21.16 A model for simultaneous synthesis of both DNA strands. (a) The lagging template strand (blue) has formed a loop through the replisome (gold), and a new primer, labeled 2 (red), has been formed by the primase. A previously synthesized Okazaki fragment (green, with red primer labeled 1) is also visible. The leading strand template and its progeny strand are shown at left (gray), but the growth of the leading strand is not considered here. **(b)** The lagging strand template has formed a bigger loop by feeding through the replisome from the top and bottom, as shown by the arrows. The motion of the lower part of the loop (lower arrow) allows the second Okazaki fragment to be elongated. **(c)** Further elongation of the second Okazaki fragment brings its end to a position adjacent to the primer of the first Okazaki fragment. **(d)** The replisome releases the loop, which permits the primase to form a new primer (number 3). The process can now begin anew.

core polymerase from the template, this model raises two important questions: First, how can discontinuous synthesis of the lagging strand possibly keep up with continuous (or perhaps discontinuous) synthesis of the leading strand? If the pol III core really dissociated completely from the template after making each Okazaki fragment of the lagging strand, it would take a long time to reassociate and would fall hopelessly behind the leading strand. This would be true even if the leading strand replicated discontinuously, because no dissociation and reassociation of the pol III core is necessary in synthesizing the leading strand. A second, related question is this: How is repeated dissociation and reassociation of the pol III core from the template compatible with the highly processive nature of DNA replication? After all, the β clamp is essential for processive replication, but once it clamps onto the DNA, how can the core polymerase dissociate every 1–2 kb as it finishes one Okazaki fragment and jumps forward to begin elongating the next?

The answer to the first question seems to be that the pol III core making the lagging strand does not really dissociate completely from the template. It remains tethered to it by its association with the core that is making the leading strand. Thus, it can release its grip on its template strand without straying far from the DNA. This enables it to find the next primer and reassociate with its template within a fraction of a second, instead of the many seconds that would be required if it completely left the DNA.

The second question requires us to look more carefully at the way the β clamp interacts with the clamp loader and with the core polymerase. We will see that these two proteins compete for the same binding site on the β clamp, and that the relative affinities of the clamp for one or the other of them shifts back and forth to allow dissociation and reassociation of the core from the DNA. We will also see that the clamp loader can act as a clamp unloader to facilitate this cycling process.

Theory predicts that the pol III* synthesizing the lagging strand must dissociate from one β clamp as it finishes one Okazaki fragment and reassociate with another β clamp to begin making the next Okazaki fragment. But does dissociation of pol III* from its β clamp actually occur? To find out, O'Donnell and his colleagues prepared a primed M13 phage template (M13mp18) and loaded a β clamp and pol III* onto it. Then they added two more primed phage DNA templates, one (M13Gori) preloaded with a β clamp and the other (ϕX174) lacking a β clamp. Then they incubated the templates together under replication conditions long enough for the original template and secondary template to be replicated. They knew they would see replicated M13mp18 DNA, but the interesting question is this: Which secondary template will be replicated, the one with, or the one without, the β clamp? Figure 21.17 (lanes 1–4) demonstrates that replication occurred preferentially on the M13Gori template—the one with the β clamp. What if they put the β clamp on the other template instead? Lanes

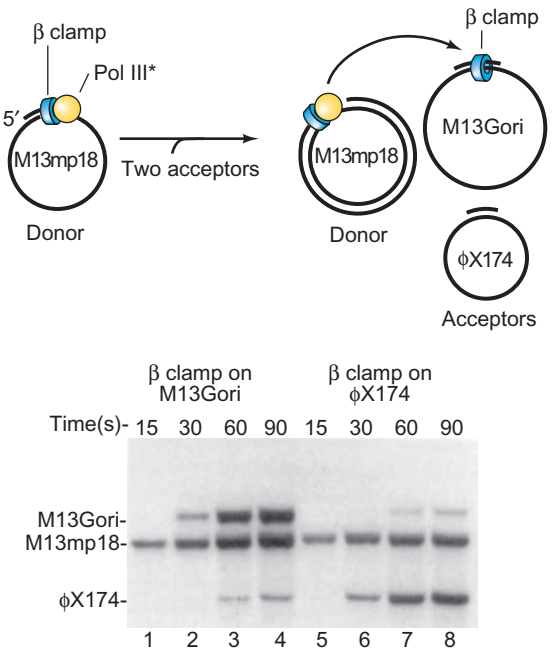


Figure 21.17 Test of the cycling model. If one assembles a pol III* complex with a β clamp on one primed template (M13mp18, top left) and presents it with two acceptor primed templates, one with a β clamp (M13Gori) and one without (ϕX174), the pol III* complex should choose the template with the clamp (M13Gori, in this case) to replicate when it has finished replicating the original template. O'Donnell and colleagues carried out this experiment, allowing enough time to replicate both the donor and acceptor templates. They also included labeled nucleotides so the replicated DNA would be labeled. Then they electrophoresed the DNAs and detected the labeled DNA products by gel electrophoresis. The electrophoresis of the replicated DNA products (bottom) show that the acceptor template with the β clamp was the one that was replicated. When the β clamp was on the M13Gori acceptor template, replication of this template predominated. On the other hand, when the β clamp was on the ϕX174 template, this was the one that was favored for replication. The positions of the replicated templates are indicated at left. (*Source:* Stukenberg, P.T., J. Turner, and M. O'Donnell, An explanation for lagging strand replication: Polymerase hopping among DNA sliding clamps. *Cell* 78 (9 Sept 1994) f. 2, p. 878. Reprinted by permission of Elsevier Science.)

5–8 show that in that case, the other template (ϕX174) was preferentially replicated. If the pol III* kept its original β clamp, it could have begun replicating either secondary template, regardless of which was preloaded with a β clamp. Thus, the results of this experiment imply that dissociation of pol III* from the template, and its β clamp, really does happen, and the enzyme can bind to another template (or another part of the same template), if another β clamp is present.

To check this conclusion, these workers labeled the β clamp with ^{32}P by phosphorylating it with [γ-^{32}P]ATP, then labeled pol III* with ^{3}H in either the θ- or τ-subunits, or in the γ complex. Then they allowed these labeled complexes to either idle on a gapped template in the presence of only dGTP and dCTP or to fill in the whole gap with all four dNTPs and thus terminate. Finally, they subjected the

reaction mixtures to gel filtration and determined whether the two labels had separated. When the polymerase merely idled, the labeled β clamp and pol III* stayed together on the DNA template. By contrast, when termination occurred, the pol III* separated from its β clamp, leaving it behind on the DNA. O'Donnell and coworkers observed the same behavior regardless of which subunit of pol III* was labeled, so this whole entity, not just the core enzyme, must separate from the β clamp and DNA template upon termination of replication.

The *E. coli* genome is 4.6 Mb long, and its lagging strand, at least, is replicated in Okazaki fragments only 1–2 kb long. This means that over 2000 priming events are required on each template, so at least 2000 β clamps are needed. Because an *E. coli* cell holds only about 300 β dimers, the supply of β clamps would be rapidly exhausted if they could not recycle somehow. This would require that they dissociate from the DNA template. Does this happen? To find out, O'Donnell and colleagues assembled several β clamps onto a gapped template, then removed all other protein by gel filtration. Then they added pol III* and reran the gel filtration step. Figure 21.18a shows that, sure enough, the β clamps dissociated in the presence of pol III*, but not without the enzyme. Figure 21.18b demonstrates that these liberated β clamps were also competent to be loaded onto an acceptor template.

It is clear from what we have learned so far that the β clamp can interact with both the core polymerase and the γ complex (the clamp loader). It must associate with the core during synthesis of DNA to keep the polymerase on the template. Then it must dissociate from the template so it can move to a new site on the DNA where it can interact with another core to make a new Okazaki fragment. This movement to a new DNA site, of course, requires the β clamp to interact with a clamp loader again. One crucial question remains: How does the cell orchestrate the shifting back and forth of the β clamp's association with core and with clamp loader?

To begin to answer this question, it would help to show how and when the core and the clamp loader interact with the β clamp. O'Donnell and associates first answered the "how" question, demonstrating that the α-subunit of the core contacts β, and the δ-subunit of the clamp loader also contacts β. One assay these workers used to reveal these interactions was **protein footprinting.** This method works on the same principle as DNase footprinting, except the starting material is a labeled protein instead of a DNA, and protein-cleaving reagents are used instead of DNase. In this case, O'Donnell and colleagues introduced a six-amino acid protein kinase recognition sequence into the C-terminus of the β-subunit by manipulating its gene. They named the altered product β^{PK}. Then they phosphorylated this protein in vitro using protein kinase and labeled ATP (an ATP derivative with an oxygen in the γ-phosphate replaced by ^{35}S); this procedure labeled the protein at its C-terminus.

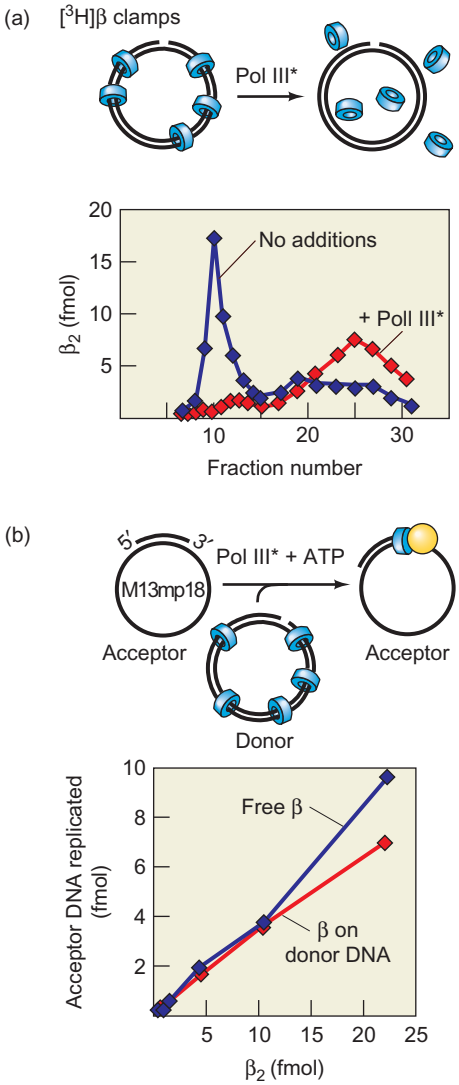


Figure 21.18 Pol III* has clamp unloading activity. (a) Clamp unloading. O'Donnell and colleagues used the γ complex to load β clamps (blue, top) onto a gapped circular template, then removed the γ complex by gel filtration. Then they added pol III* and performed gel filtration again. The graph of the results (bottom) shows β clamps that were treated with pol III* (red) were released from the template, whereas those that were not treated with pol III* (blue) remained associated with the template. **(b)** Recycling of β clamps. The β clamps from a donor β clamp–template complex treated with pol III* (red) were just as good at rebinding to an acceptor template as were β clamps that were free in solution (blue). (*Source:* Adapted from Stukenberg, P.T., J. Turner, and M. O'Donnell, An explanation for lagging strand replication: Polymerase hopping among DNA sliding clamps. *Cell* 78:883, 1994.)

(Note that this is similar to labeling a DNA at one of its ends for DNase footprinting.) First they showed that the δ-subunit of the clamp loader and the α-subunit of the core could each protect β^{PK} from phosphorylation, suggesting that both of these proteins contact β^{PK}.

Protein footprinting reinforced these conclusions. O'Donnell and colleagues mixed labeled β^{PK} with various proteins, then cleaved the protein mixture with two

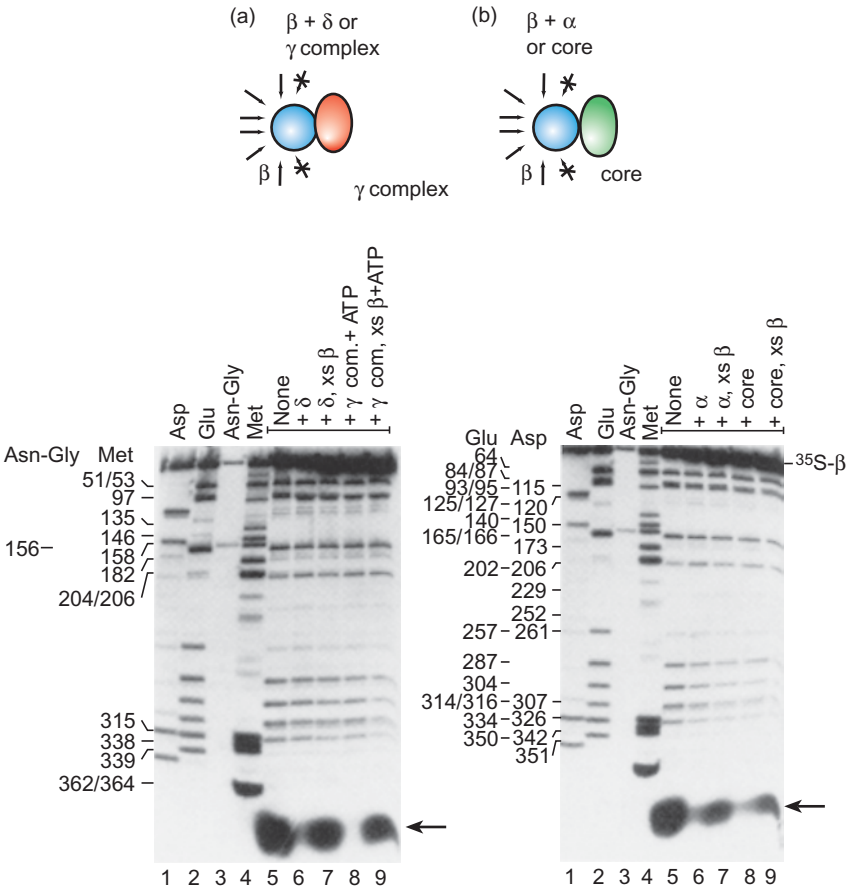


Figure 21.19 Protein footprinting of β with the γ complex and core polymerase. O'Donnell and colleagues labeled β^{PK} at its C-terminus by phosphorylation with protein kinase and [^{35}S]ATP. Then they mixed this end-labeled β with either δ or the whole γ complex (panel **a**) or with either α or the whole core (panel **b**). Then they subjected the protein complexes to mild cleavage with a mixture of pronase E and V8 protease to generate a series of end-labeled digestion products. Finally, they electrophoresed these products and autoradiographed the gel to detect them. The first four lanes in each panel are digestion products that serve as markers. The amino acid specificity of each treatment is given at top. Thus, in lane 1, the protein was treated with a protease that cleaves after aspartate (Asp) residues. Lane 5 in both panels represents β^{PK} cleaved in the absence of other proteins. Lanes 6–9 in both panels represent β^{PK} cleaved in the presence of the proteins listed at the top of each lane. The δ- and α-subunits and the γ and core complexes all protect the same site from digestion. Thus, they reduce the yield of the fragment indicated by the arrow at the bottom of the gel. The drawings at top illustrate the binding between the β clamp and either the γ complex (a) or the core (b), emphasizing that both contact the β clamp at the same places near the C-terminus of each β monomer and prevent cleavage there (arrows with Xs). (*Source:* Naktinis, V., J. Turner, and M. O'Donnell, A molecular switch in a replicating machine defined by an internal competition for protein rings. *Cell* 84 (12 June 1996) f. 3ab bottoms, p. 138. Reprinted by permission of Elsevier Science.)

proteolytic enzymes: pronase E and V8 protease. Figure 21.19 depicts the results. The first four lanes at the bottom of each panel are markers formed by cleaving the labeled β-subunit with four different reagents that cleave at known positions. Lane 5 in both panels shows the end-labeled peptides created by cleaving β in the absence of another protein. We observe a typical ladder of end-labeled products. Lane 6 in panel (a) shows what happens in the presence of δ. We see the same ladder as in lane 5, with the exception of the smallest fragment (arrow), which is either missing or greatly reduced in abundance. This suggests that the δ-subunit binds to β near its C-terminus and blocks a protease from cleaving there. If this δ–β interaction is specific, one should be able to restore cleavage of the labeled β^{PK} by adding an abundance of unlabeled β to bind to δ and prevent its binding to the labeled β^{PK}. Lane 7 shows that this is what happened. Lanes 8 and 9 in panel (a) are similar to 6 and 7, except that O'Donnell and coworkers used whole γ complex instead of purified δ. Again, the γ complex protected a site near the C-terminus of β^{PK} from cleavage, and unlabeled β prevented this protection.

Panel (b) of Figure 21.19 is just like panel (a), except that the investigators used the α-subunit and whole core instead of the δ-subunit and whole γ complex to footprint labeled β^{PK}. They observed exactly the same results: α or whole core protected the same site from cleavage as did δ or whole γ complex. This suggests that the core and the clamp loader both contact β at the same site, and that the α- and δ-subunits, respectively, mediate these contacts. In a further experiment, these workers used whole pol III* to footprint β^{PK}. Because pol III* contains both the core and the clamp loader, one might have expected it to yield a larger footprint than either subassembly separately. But it did not. This is consistent with the hypothesis that pol III* contacts β through either the core or the clamp loader, but not both at the same time.

If the β clamp can bind to the core or the clamp loader, but not both simultaneously, which does it prefer? O'Donnell and colleagues used gel filtration to show that when the proteins are free in solution, β prefers to bind to the clamp loader, rather than the core polymerase. This is satisfying because free β needs to be loaded onto DNA by the γ complex before it can interact with the core polymerase. However, that situation should change once the β clamp is loaded onto a primed DNA template; once that happens, β needs to associate with the core polymerase and begin making DNA. To test this prediction, O'Donnell and colleagues loaded ^{35}S-labeled β clamps onto primed M13 phage DNA and then added either ^{3}H-labeled clamp loader (γ complex) and unlabeled core, or ^{3}H-labeled core and unlabeled γ complex. Then they subjected these mixtures to gel filtration to separate DNA–protein complexes from free proteins.

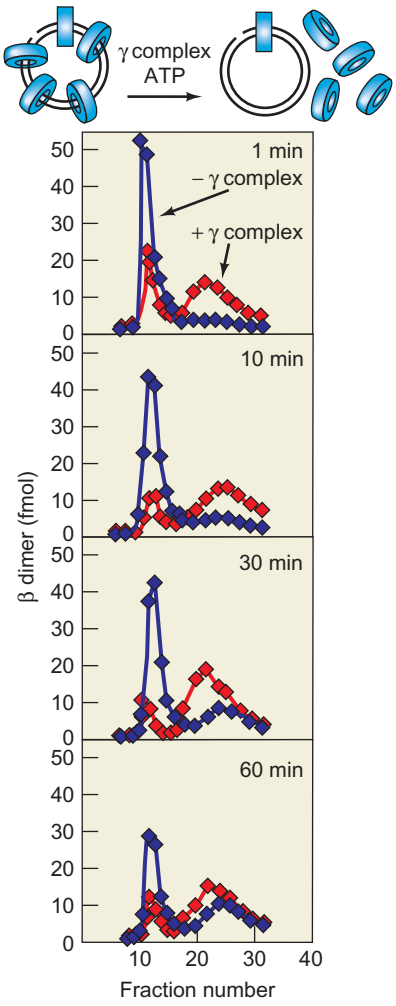


Figure 21.20 Clamp-unloading activity of the γ complex. O'Donnell and coworkers loaded β clamps onto a nicked circular DNA template, as shown at top, then incubated these complexes in the presence (red) or absence (blue) of the γ complex and ATP for the times indicated. Finally, they subjected the mixtures to gel filtration to determine how much β clamp remained associated with the DNA and how much had dissociated. The cartoon at top interprets the results: The γ complex and ATP served to accelerate the unloading of β clamps from the nicked DNA. (*Source:* Adapted from Naktinis, V., J. Turner, and M. O'Donnell, A molecular switch in a replication machine defined by an internal competition for protein rings. *Cell* 84:141, 1996.)

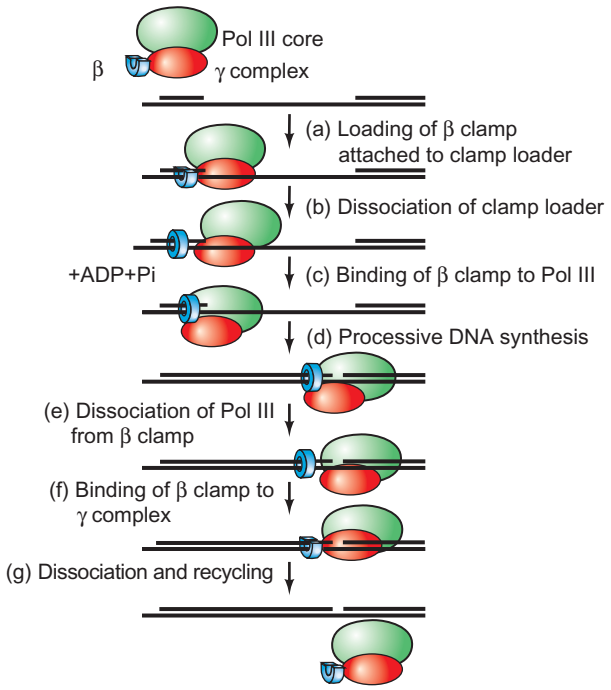


Figure 21.21 Summary of lagging strand replication. We begin with a β clamp associated with the γ complex part (red) of a pol III*. **(a)** The γ complex loads the β clamp (blue) onto a primed DNA template. **(b)** The γ complex, or clamp loader, dissociates from the β clamp. **(c)** The core (green) associates with the clamp. **(d)** The core and clamp cooperate to processively synthesize an Okazaki fragment, leaving just a nick between two Okazaki fragments. **(e)** The polymerase core dissociates from the clamp. **(f)** The γ complex reassociates with the β clamp. **(g)** The γ complex acts as a clamp unloader, removing the β clamp from the template. Now it is free to repeat the process, recycling to another primer on the template. (*Source:* Adapted from Herendeen, D.R. and T.J. Kelly, DNA polymerase III: Running rings around the fork. *Cell* 84:7, 1996.)

Under these conditions, it was clear that the β clamp on the DNA preferred to associate with the core polymerase. Almost no γ complex bound to the β clamp–DNA complex.

Once the holoenzyme has completed an Okazaki fragment, it must dissociate from the β clamp and move to a new one. Then the original β clamp must be removed from the template so it can participate in the synthesis of another Okazaki fragment. We have already seen that pol III* has clamp-unloading activity, but we have not seen what part of pol III* has this activity. O'Donnell and associates performed gel filtration assays that showed that the γ complex has clamp-unloading activity. Figure 21.20 illustrates this experiment. The investigators loaded β clamps onto a nicked DNA template, then removed all other proteins. Then they incubated these DNA–protein complexes in the presence and absence of the γ complex. We can see that the β clamps are unloaded from the nicked DNA much faster in the presence of the γ complex and ATP than in their absence.

Thus the γ complex is both a clamp loader and a clamp unloader. But what determines when it will load clamps and when it will unload them? The state of the DNA seems to throw this switch, as illustrated in Figure 21.21. Thus, when β clamps are free in solution and there is a primed template available, the clamps associate preferentially with the γ complex, which serves as a clamp loader to bind the β clamp to the DNA. Once associated with the DNA, the clamp binds preferentially to the core polymerase and sponsors processive synthesis of an Okazaki fragment. When the fragment has been synthesized, and only a nick remains, the core loses its affinity for the β clamp. The clamp reassociates with the γ complex, which now acts as a clamp unloader, removing the clamp from the template so it can recycle to the next primer and begin the cycle anew.

SUMMARY The pol III holoenzyme is double-headed, with two core polymerases attached through two τ-subunits to a γ complex. One core is responsible for (presumably) continuous synthesis of the leading strand, the other performs discontinuous synthesis of the lagging strand. The γ complex serves as a clamp loader to load the β clamp onto a primed DNA template. Once loaded, the β clamp loses affinity for the γ complex and associates with the core polymerase to help with processive synthesis of an Okazaki fragment. Once the fragment is completed, the β clamp loses affinity for the core polymerase and associates with the γ complex, which acts as a clamp unloader, removing the clamp from the DNA. Now it can recycle to the next primer and repeat the process.

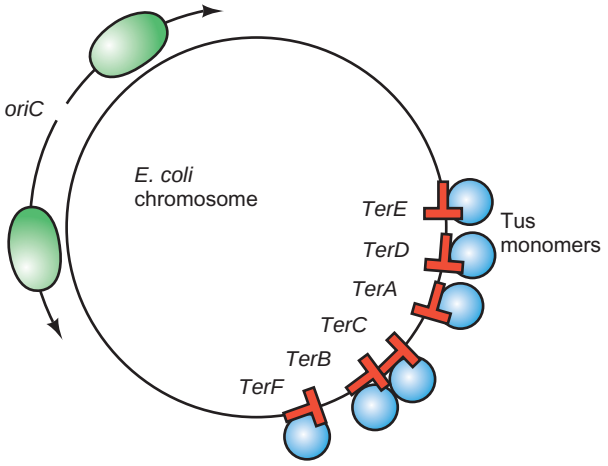


Figure 21.22 The termination region of the *E. coli* genome. Two replicating forks with their accompanying replisomes (green) are pictured moving away from *oriC* toward the terminator region on the opposite side of the circular *E. coli* chromosome. Three terminator sites operate for each fork: *TerE, TerD, and TerA* stop the counterclockwise fork; and *TerF, TerB,* and *TerC* stop the clockwise fork. The Tus protein binds to the terminator sites and helps arrest the moving forks. (*Source:* Adapted from Baker, T.A., Replication arrest. *Cell* 80:521, 1995.)

21.3 Termination

Termination of replication is relatively straightforward for λ and other phages that produce a long, linear concatemer. The concatemer simply continues to grow as genome-sized parts of it are snipped off and packaged into phage heads. But for bacteria and eukaryotes, where replication has a definite end as well as a beginning, the mechanisms of termination are more complex and more interesting. In bacterial DNA replication, the two replication forks approach each other in the terminus region, which contains 22-bp terminator sites that bind specific proteins. In *E. coli,* the terminator (*Ter*) sites are *TerA–TerF,* and they are arranged as pictured in Figure 21.22. The *Ter* sites bind proteins called **Tus** (for terminus utilization substance). Replicating forks enter the terminus region and pause before quite completing the replication process. This leaves the two daughter duplexes entangled. They must become disentangled before cell division occurs, or they cannot separate to the two daughter cells. Instead, they would remain caught in the middle of the cell, cell division would fail, and the cell would probably die. These considerations raise the question: How do the daughter duplexes become disentangled? For eukaryotes, we would like to know how cells fill in the gaps left by removing primers at the 5′-ends of the linear chromosomes. Let us examine each of these problems.

Decatenation: Disentangling Daughter DNAs

Bacteria face a problem near the end of DNA replication. Because of their circular nature, the two daughter duplexes remain entwined as two interlocking rings, a type of **catenane.** For these interlocked DNAs to move to the two daughter cells, they must be unlinked, or **decatenated.** If decatenation occurs before repair synthesis, a single nick will suffice to disentangle the DNAs, and a type I topoisomerase can perform the decatenation. However, if repair synthesis occurs first, a type II topoisomerase, which passes a DNA duplex through a double-stranded break, is required. *Salmonella typhimurium* and *E. coli* cells contain four topoisomerases: topoisomerases I–IV (topo I–IV). Topo I and III are type I enzymes, and topo II and IV are type II. The question is: Which topoisomerase is involved in decatenation?

Because DNA gyrase (topo II) acts as the swivel during DNA replication, many molecular biologists assumed that it also decatenates the daughter duplexes. But Nicholas Cozzarelli and his colleages demonstrated that **topo IV** is really the decatenating enzyme. They tested various temperature-sensitive mutant strains of *S. typhimurium,* a close relative of *E. coli,* for ability to decatenate dimers of the plasmid pBR322 in vivo at the permissive and nonpermissive temperatures. They showed that bacteria with mutations in the genes encoding the subunits of topo IV failed to decatenate the plasmid at the nonpermissive temperature (44°C) in the absence of norfloxacin. This suggests that topo IV is important in decatenation. Norfloxacin, by blocking DNA gyrase, halted DNA replication and presumably allowed subsequent decatenation by the small amount of residual topo IV, or by another topoisomerase. By contrast, the strain with the mutant DNA gyrase did not accumulate catenanes at the nonpermissive temperature, either in the presence or absence of norfloxacin, suggesting that this enzyme does not participate in decatenation. When they tested temperature-sensitive mutants of *E. coli,* Cozzarelli and colleagues observed similar behavior, indicating that topo IV also participates in decatenation in *E. coli.*

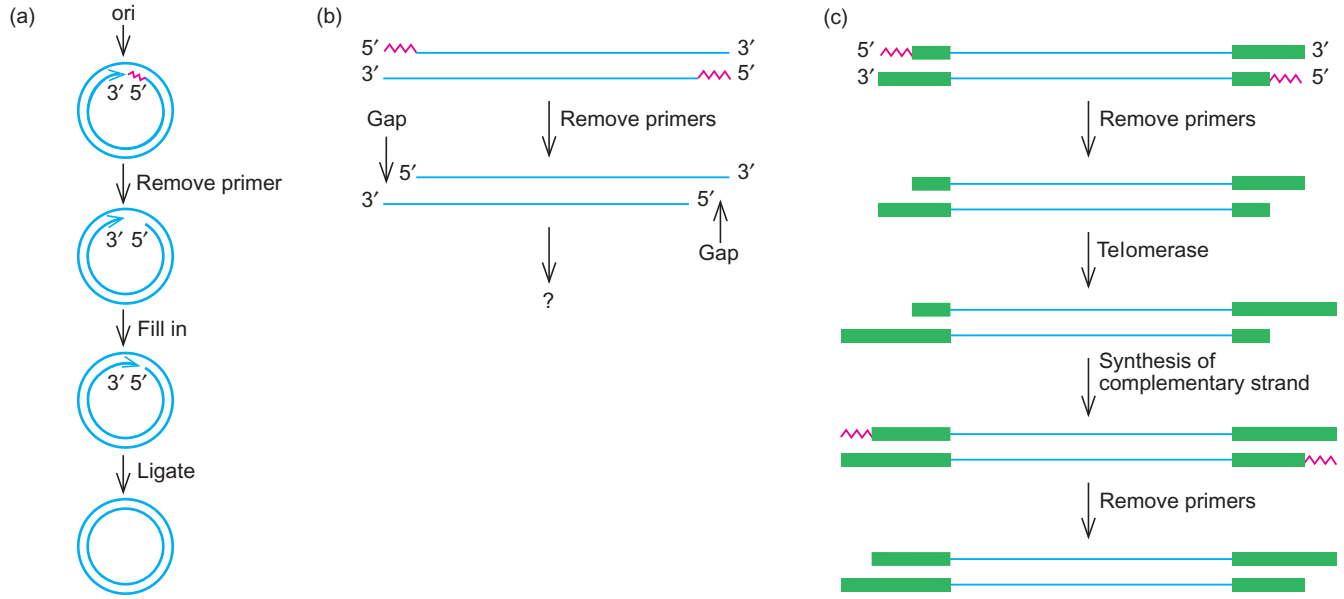


Figure 21.23 Coping with the gaps left by primer removal. (a) In bacteria, the 3′-end of a circular DNA strand can prime the synthesis of DNA to fill in the gap left by the first primer (pink). For simplicity, only one replicating strand is shown. **(b)** Hypothetical model to show what would happen if primers were simply removed from the 5′-end of linear DNA strands with no telomerase action. The gaps at the ends of chromosomes would grow longer each time the DNA replicated. **(c)** How telomerase can solve the problem. In the first step, the primers (pink) are removed from the 5′-ends of the daughter strands, leaving gaps. In the second step, telomerase adds extra telomeric DNA (green boxes) to the 3′-ends of the other daughter strands. In the third step, DNA synthesis occurs, using the newly made telomeric DNA as a template. In the fourth step, the primers used in step three are removed. This leaves gaps, but the telomerase action has ensured that no net loss of DNA has occurred. The telomeres represented here are not drawn to scale with the primers. In reality, human telomeres are thousands of nucleotides long. (*Source: (c)* Adapted from Greider, C.W. and E.H. Blackburn, Identification of a specific telomere terminal transferase activity in tetramere extracts. *Cell* 43 (Dec Pt1 1985) f. 1A, p. 406.)

Eukaryotic chromosomes are not circular, but they have multiple replicons, so replication forks from neighboring replicons approach one another just as the two replication forks of a bacterial chromosome approach each other near the termination point opposite the origin of replication. Apparently, this inhibits completion of DNA replication, so eukaryotic chromosomes also form catenanes that must be disentangled. Eukaryotic topo II resembles bacterial topo IV more than it does DNA gyrase, and it is a strong candidate for the decatenating enzyme.

SUMMARY At the end of replication, circular bacterial chromosomes form catenanes that must be decatenated for the two daughter duplexes to separate. In *E. coli* and related bacteria, topoisomerase IV performs this decatenation. Linear eukaryotic chromosomes also require decatenation during DNA replication.

Termination in Eukaryotes

Eukaryotes face a difficulty at the end of DNA replication that prokaryotes do not: filling in the gaps left when RNA primers are removed. With circular DNAs, such as those in bacteria, there is no problem filling all the gaps because another DNA 3′-end is always upstream to serve as primer (Figure 21.23a). But consider the problem faced by eukaryotes, with their linear chromosomes. Once the first primer on each strand is removed (Figure 21.23b), there is no way to fill in the gaps because DNA cannot be extended in the 3′→5′ direction, and no 3′-end is upstream, as there would be in a circle. If this were actually the situation, the DNA strands would get shorter every time they replicated. This is a termination problem in that it deals with the formation of the ends of the DNA strands, but how do cells solve this problem?

Telomere Maintenance Elizabeth Blackburn and her colleagues provided the answer, which is summarized in Figure 21.23c. The **telomeres,** or ends of eukaryotic chromosomes, are composed of repeats of short, GC-rich sequences. The G-rich strand of a telomere is added at the very 3′-ends of DNA strands, not by semiconservative replication, but by an enzyme called **telomerase.** The exact sequence of the repeat in a telomere is species-specific. In *Tetrahymena,* it is TTGGGG/AACCCC; in vertebrates, including humans, it is TTAGGG/AATCCC. Blackburn showed that this specificity resides in the telomerase itself and is due to a small RNA in the enzyme that serves as the template for telomere synthesis. This solves the problem:

The telomerase adds many repeated copies of its characteristic sequence to the 3′-ends of chromosomes. Priming can then occur within these telomeres to make the C-rich strand. There is no problem when terminal primers are removed and not replaced, because only telomere sequences are lost, and these can always be replaced by telomerase and another round of telomere synthesis.

Blackburn made a clever choice of organism in which to search for telomerase activity: *Tetrahymena,* a ciliated protozoan. *Tetrahymena* has two kinds of nuclei: (1) micronuclei, which contain the whole genome in five pairs of chromosomes that serve to pass genes from one generation to the next; and (2) macronuclei, in which the five pairs of chromosomes are broken into more than 200 smaller fragments used for gene expression. Because each of these minichromosomes has telomeres at its ends, *Tetrahymena* cells have many more telomeres than human cells, for example, and they are loaded with telomerase, especially during the phase of life when macronuclei are developing and the new minichromosomes must be supplied with telomeres. This made isolation of the telomerase enzyme from *Tetrahymena* relatively easy.

In 1985, Carol Greider and Blackburn succeeded in identifying a telomerase activity in extracts from synchronized *Tetrahymena* cells that were undergoing macronuclear development. They assayed for telomerase activity in vitro using a synthetic primer with four repeats of the TTGGGG telomere sequence and included a radioactive nucleotide to label the extended telomere-like DNA. Figure 21.24 shows the results. Lanes 1–4 each contained a different labeled nucleotide (dATP, dCTP, dGTP, and dTTP, respectively), plus all three of the other, unlabeled nucleotides. Lane 1, with labeled dATP showed only a smear, and lanes 2 and 4 showed no extension of the synthetic telomere. But lane 3, with labeled dGTP, exhibited an obvious periodic extension of the telomere. Each of the clusters of bands represents an addition of one more TTGGGG sequence (with some variation in the degree of completion), which accounts for the fact that we see *clusters* of bands, rather than *single* bands. Of course, we should observe telomere extension with labeled dTTP, as well as with dGTP. Further investigation showed that the concentration of dTTP was too low in this experiment, and that dTTP could be incorporated into telomeres at higher concentration. Lanes 5–8 show the results of an experiment with one labeled, and only one unlabeled nucleotide. This experiment verifed that dGTP could be incorporated into the telomere, but only if unlabeled dTTP was also present. This is what we expect because this strand of the telomere contains only G and T. Controls in lanes 9–12 showed that an ordinary DNA polymerase, Klenow fragment, cannot extend the telomere. Further controls in lanes 13–16 demonstrated that telomerase activity depends on the telomere-like primer.

How does telomerase add the correct sequence of bases to the ends of telomeres without a complementary DNA strand to read? It uses its own RNA constituent as a template. (Note that this is a template, not a primer.) Greider and Blackburn demonstrated in 1987 that telomerase is a ribonucleoprotein with essential RNA and protein subunits. Then in 1989 they cloned and sequenced the gene that encodes the 159-nt RNA subunit of the *Tetrahymena* telomerase and found that it contains the sequence CAACCCCAA. In principle, this sequence can serve as template for repeated additions of TTGGGG sequences to the ends of *Tetrahymena* telomeres as illustrated in Figure 21.25.

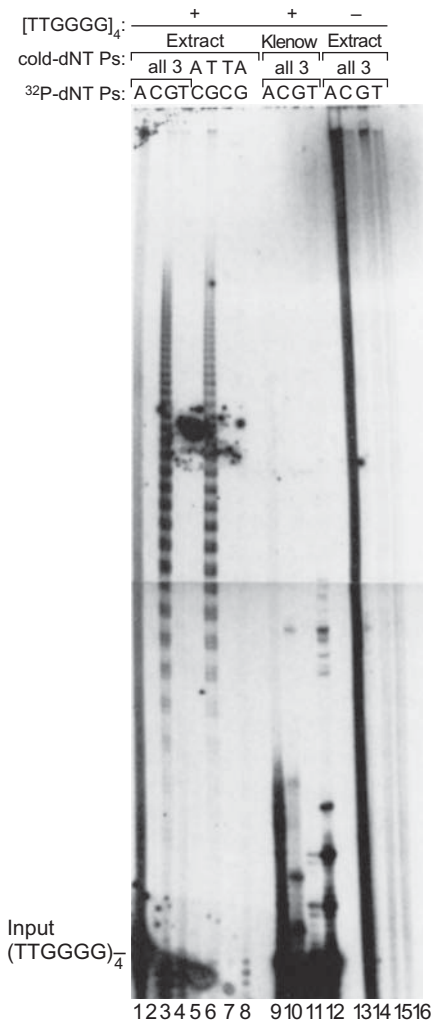


Figure 21.24 Identification of telomerase activity. Greider and Blackburn synchronized mating of *Tetrahymena* cells and let the offspring develop to the macronucleus development stage. They prepared cell-free extracts and incubated them for 90 min with a synthetic oligomer having four repeats of the TTGGGG telomere repeat sequence, plus the labeled and unlabeled nucleotides indicated at top. After incubation, they electrophoresed the products and detected them by autoradiography. Lanes 9–12 contained the Klenow fragment of *E. coli* DNA polymerase I instead of *Tetrahymena* extract. Lanes 13–16 contained extract, but no primer. Telomerase activity is apparent only when both dGTP and dTTP are present. (*Source:* Greider, C.W., and E.H. Blackburn, Identification of a specific telomere terminal transferase activity in tetramere extracts. *Cell* 43 (Dec Pt1 1985) f. 1A, p. 406. Reprinted by permission of Elsevier Science.)

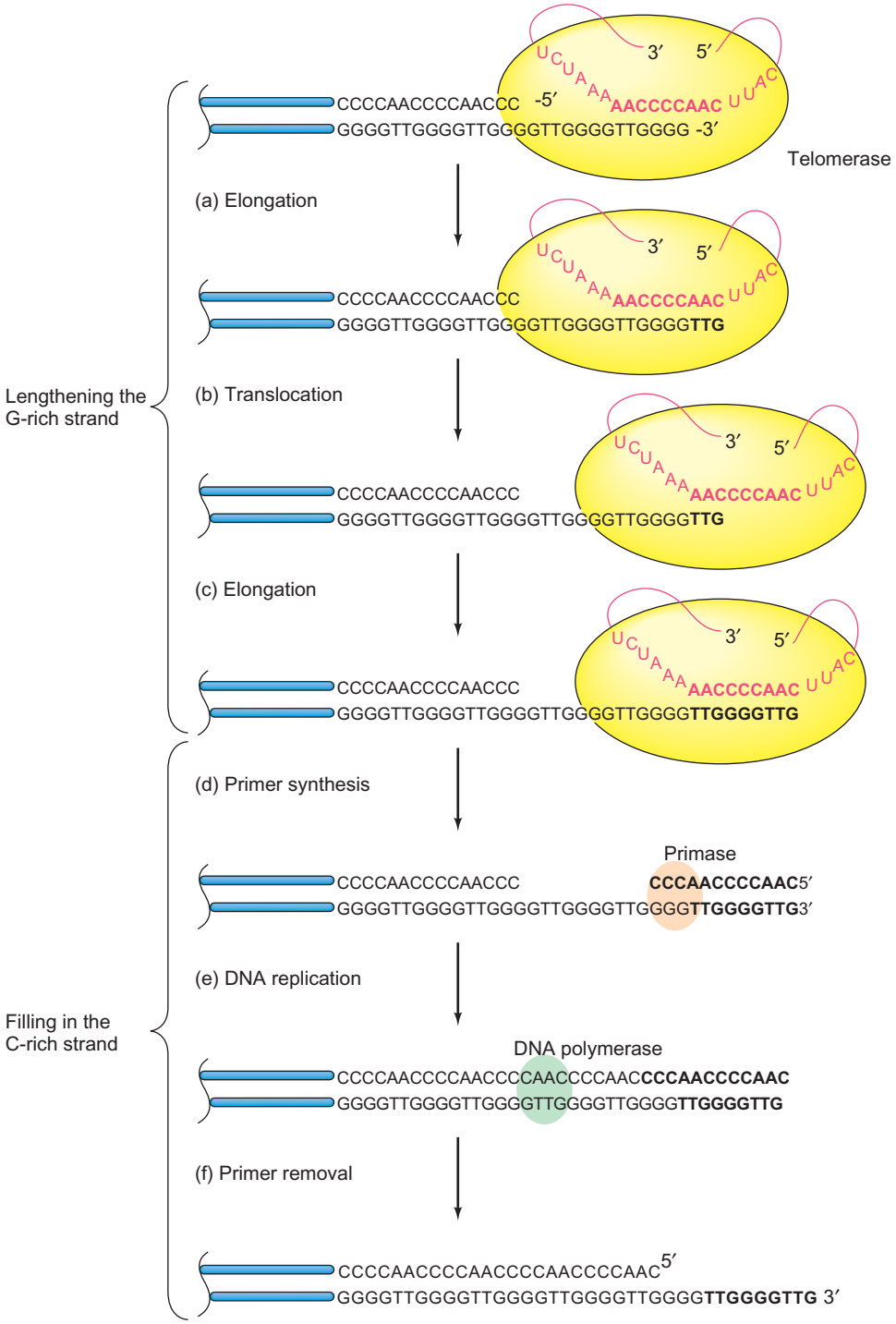


Figure 21.25 Forming telomeres in *Tetrahymena*. **(a)** Telomerase (yellow) promotes hybridization between the 3′-end of the G-rich telomere strand and the template RNA (red) of the telomerase. The telomerase uses three bases (AAC) of its RNA as a template for the addition of three bases (TTG, boldface) to the 3′-end of the telomere. **(b)** The telomerase translocates to the new 3′-end of the telomere, pairing the left-hand AAC sequence of its template RNA with the newly incorporated TTG in the telomere. **(c)** The telomerase uses the template RNA to add six more nucleotides (GGGTTG, boldface) to the 3′-end of the telomere. Steps (a) through (c) can repeat indefinitely to lengthen the G-rich strand of the telomere. **(d)** When the G-rich strand is sufficiently long (probably longer than shown here), primase (orange) can make an RNA primer (boldface), complementary to the 3′-end of the telomere's G-rich strand. **(e)** DNA polymerase (green) uses the newly made primer to prime synthesis of DNA to fill in the remaining gap on the C-rich telomere strand and DNA ligase seals the nick. **(f)** The primer is removed, leaving a 12–16-nt overhang on the G-rich strand.

Blackburn and her colleagues used a genetic approach to prove that the telomerase RNA really does serve as the template for telomere synthesis. They showed that mutant telomerase RNAs gave rise to telomeres with corresponding alterations in their sequence. In particular, they changed the sequence 5′-CAACCCCAA-3′ of a cloned gene encoding the *Tetrahymena* telomerase RNA as follows:

wt: 5′-CAACCCCAA-3′
1: 5′-CAACCCCCAA-3′
2: 5′-CAACCTCAA-3′
3: 5′-CGACCCCAA-3′

The underlined bases in each of the three mutants (1, 2, and 3) denote the base changed (or added, in 1). They introduced the wild-type or mutated gene into *Tetrahymena* cells in a plasmid that ensured the gene would be overexpressed. Even though the endogenous wild-type gene remained in each case, the overexpression of the transplanted gene swamped out the effect of the endogenous gene. Southern blotting of telomeric DNA from cells transformed with each construct showed that a probe for the telomere sequence expected to result from mutants 1 (TTGGGGG) and 3 (GGGGTC) actually did hybridize to telomeric DNA from cells transformed with these mutant genes. On the other hand, this did not work for mutant 2; no telomeric DNA that hybridized to a probe for GAGGTT was observed.

These results suggested that mutant telomerase RNAs 1 and 3, but not 2, served as templates for telomere elongation. To confirm this suggestion, Blackburn and colleagues sequenced a telomere fragment from cells transformed with mutant telomerase RNA 3. They found the following sequence:

5′-CTTTTACTCAATGTCAAAGAAATTATTAAATT$(GGGGTT)_{30}$
$(GGGGTC)_2$GGGGTT$(GGGGTC)_8$GGGGTTGGGGTC$(GGGGTT)_N$-3′

where the underlined bases must have been encoded by the mutant telomerase RNA. This nonuniform sequence differs stikingly from the normal, very uniform telomeric sequence in this species. The first 30 repeats appear to have been encoded by the wild-type telomerase RNA before transformation. These are followed by 11 mutant repeats interspersed with 2 wild-type repeats, then by all wild-type repeats. The terminal wild-type sequences may have resulted from recombination with a wild-type telomere, or from telomere synthesis after loss of the mutant telomerase RNA gene from the cell. Nevertheless, the fact remains that a significant number of repeats have exactly the sequence we would expect if they were encoded by the mutant telomerase RNA. Thus, we can conclude that the telomerase RNA does serve as the template for telomere synthesis, as Figure 21.25 suggests.

The fact that telomerase uses an RNA template to make a DNA strand implies that telomerase acts as a reverse transcriptase. Thus, Blackburn and others set about to purify the enzyme to prove that this is indeed how it works. Although the enzyme eluded purification for 10 years, Joachim Lingner and Thomas Cech finally succeeded in 1996 in purifying it from another ciliated protozoan, *Euplotes*. This telomerase contains two proteins, p43 and p123, in addition to the RNA subunit that serves as the template for extending telomeres. The p123 protein has the signature sequence of a reverse transcriptase, indicating that it provides the catalytic activity of the enzyme. We therefore call it **TERT,** for **telomerase reverse transcriptase.** Because this enzyme was discovered when the Human Genome Project was well along, it did not take long to find a complementary sequence in the human genome and use it to clone the human TERT gene, *hTERT,* in 1997.

Structural analysis has shown that the C-terminal part of the TERT protein contains the reverse transcriptase activity, and the N-terminal part binds to the RNA. In fact, the RNA appears to be tethered to the protein so as to give the RNA, which is hundreds of nucleotides long, considerable flexibility. This allows the RNA to fulfill its template role by moving with respect to the active site of the enzyme as each nucleotide is added to the growing telomere.

Until 2003, it appeared that the somatic cells of higher eukaryotes, including humans, lack telomerase activity, whereas germ cells retain this activity. Then, William Hahn and colleagues showed that cultured normal human cells do express telomerase at a low level, but only transiently, during S phase, when DNA is replicated. On the other hand, cancer cells have much higher telomerase activity, which is expressed constitutively—all the time. These findings have profound implications for the characteristics of cancer cells, and perhaps even for their control (see Box 21.1).

SUMMARY Eukaryotic chromosomes have special structures known as telomeres at their ends. One strand of these telomeres is composed of many tandem repeats of short, G-rich regions whose sequence varies from one species to another. The G-rich telomere strand is made by an enzyme called telomerase, which contains a short RNA that serves as the template for telomere synthesis. The C-rich telomere strand is synthesized by ordinary RNA-primed DNA synthesis, like the lagging strand in conventional DNA replication. This mechanism ensures that chromosome ends can be rebuilt and therefore do not suffer shortening with each round of replication.

Telomere Structure Besides protecting the ends of chromosomes from degradation, telomeres play another critical role: They prevent the DNA repair machinery from recognizing the ends of chromosomes as chromosome breaks and sticking chromosomes together. This inapproriate joining of chromosomes would be potentially lethal to the cell. Furthermore, cells have a DNA damage checkpoint that

BOX 21.1

Telomeres, the Hayflick Limit, and Cancer

Everyone knows that organisms, including humans, are mortal. But biologists used to assume that cells cultured from humans were immortal. Each individual cell would ultimately die, of course, but the cell *line* would go on dividing indefinitely. Then in the 1960s Leonard Hayflick discovered that ordinary human cells are not immortal. They can be grown in culture for a finite period—about 50 generations (or cycles of subculturing). Then they enter a period of senescence, when they slow down and then stop dividing, and finally they reach a crisis stage and die. This ceiling on the lifetime of normal cells is known as the Hayflick limit. But cancer cells do not obey any such limit. They *do* go on dividing generation after generation, indefinitely.

Investigators have discovered a significant difference between normal cells and cancer cells that may explain why cancer cells are immortal and normal cells are not: Human cancer cells contain abundant telomerase that is expressed constitutively, whereas normal somatic cells generally produce this enzyme only weakly and transiently. (Germ cells must retain telomerase, of course, to safeguard the ends of the chromosomes handed down to the next generation.) Thus, we see that cancer cells can repair their telomeres after every cell replication, but most normal cells cannot. Therefore, cancer cells can go on dividing without degrading their chromosomes, whereas normal cells' chromosomes grow shorter with each cell division. Sooner or later the telomeres are lost, and the ends of chromosomes that lack telomeres look like the ends of broken chromosomes. Most cells react to this apparent assault by halting their replication and ultimately by dying. But this does not happen to cancer cells; telomerase saves them from that fate.

One of the typical changes that occurs in a cell to make it cancerous is the reactivation of the telomerase gene. This leads to the immortality that is the hallmark of cancer cells. This discussion also suggests a potential treatment for cancer: Turn off the telomerase gene in cancer cells or, more simply, administer a drug that inhibits telomerase. Such a drug may not harm most normal cells because they have very little telomerase to begin with. Cancer researchers are hard at work on this strategy, but the discovery in 2003 that human fibroblasts in culture express low levels of *hTERT* and have a little telomerase activity casts some doubt on this idea. Further reservations come from the findings that expression of an inactive form of *hTERT,* or inhibiting the expression of normal *hTERT* by RNAi, causes premature senescence in human fibroblasts. The trick will be to kill cancer cells without dooming the patient's normal cells to an early death.

Some signs indicate that simply inhibiting the telomerase of cancer cells may not cause the cells to die. For one thing, knockout mice totally lacking telomerase activity survive and reproduce for at least six generations, though eventually the loss of telomeres leads to sterility. However, cells from these telomerase knockout mice can be immortalized, they can be transformed by tumor viruses, and these transformed cells can give rise to tumors when transplanted to immunodeficient mice. Thus, the presence of telomerase is not an absolute requirement for the development of a cancer cell. It may be that mouse cells have a way of preserving their telomeres without telomerase. We will have to see whether human cells behave differently.

Finally, immortalizing human cells in culture leads to the idea of immortalizing human beings themselves. Could it be that reactivating telomerase activity in human somatic cells would lengthen human lifetimes? Or would it just make us more susceptible to cancer? To begin answering this question, Serge Lichtsteiner, Woodring Wright, and their colleagues transplanted the *hTERT* gene into human somatic cells in culture, so these cells were forced to express telomerase activity. The results were striking: The telomeres in these cells grew longer and the cells went on dividing far past their normal lifetimes. They remained youthful in appearance and in their chromosome content. Furthermore, they did not show any signs of becoming cancerous. These findings were certainly encouraging, but they do not herald a fountain of youth. For now, that remains in the realm of science fiction.

detects damage and stops cell division until the damage can be repaired. Because chromosome ends without telomeres look like broken chromosomes, they invoke the checkpoint, so cells stop dividing and eventually die. If telomeres really looked the way they are pictured in Figures 21.23 and 21.25, little would distinguish them from real chromosome breaks. In fact, the critical telomere length in humans is 12.8 repeats of the core 6-bp sequence. Below that threshold, human chromosomes began to fuse. How do telomeres allow the cell to recognize the difference between a real chromosome end and a broken chromosome?

For years, molecular biologists pondered this question and, as telomere-binding proteins were discovered, they theorized that these proteins bind to the ends of chromosomes and in that way identify the ends. Indeed, eukaryotes from yeasts to mammals have a suite of telomere-binding

proteins that protect the telomeres from degradation, and also hide the telomere ends from the DNA damage factors that would otherwise recognize them as chromosome breaks. We will discuss the telomere-binding proteins from three groups of eukaryotes and see how they solve the telomere protection problem.

The Mammalian Telomere-Binding Proteins: Shelterin In mammals, the group of telomere-binding proteins is appropriately known as **shelterin,** because it "shelters" the telomere. There are six known mammalian shelterin proteins: **TRF1, TRF2, TIN2, POT1, TPP1,** and **RAP1.** TRF1 was the first of these proteins to be discovered. Because it bound to double-stranded telomere DNA, which includes repeats of the sequence TTAGGG, it was named **TTAGGG repeat-binding factor-1** (TRF1). TRF2 is a product of a paralog of the TRF1 gene (paralogs are homologous genes in the same organism), and it also binds to the double-stranded parts of telomeres. POT1 (**protection of telomeres -1**) binds to the single-stranded 3′-tails of telomeres, beginning at a position just 2 nt away from the 5′-end of the other strand. In this way it is positioned to protect the single-stranded telomeric DNA from endonucleases, and the 5′-end of the other strand within the double-stranded telomeric DNA from 5′-exonucleases. TPP1 is a POT1-binding protein. Indeed, it appears to be a partner of POT1 in a heterodimer. TIN2 (**TRF1-interacting factor-2**) plays an organizing role in shelterin. It connects TRF1 and TRF2 together, and connects the dimer TPP1/POT1 to TRF1 and TRF2. Finally, RAP1, with the uninformative name "repressor activator protein-1," binds to the telomere by interacting with TRF2.

Other proteins besides shelterin bind to telomeres, but shelterin proteins can be distinguished from the others in three ways: They are found only at telomeres; they associate with telomeres throughout the cell cycle; and they are known to function nowhere else in the cell. Other proteins may fulfill one of these criteria, but not two or all three.

Shelterin can affect the structure of telomeres in three ways. First, it can remodel the telomere into a loop called a **t-loop** (for "telomere-loop"). In 1999, Jack Griffith and Titia de Lange and their colleagues discovered that mammalian telomeres are not linear, as had been assumed, but form a DNA loop they called a t-loop. These loops are unique in the chromosome and therefore quite readily set the ends of chromosomes apart from breaks that occur in the middle and would yield linear ends to the chromosome fragments.

What is the evidence for t-loops? Griffith, de Lange and colleagues started by making a model mammalian telomeric DNA with about 2 kb of repeating TTAGGG sequences, and a 150–200-nt single-stranded 3′-overhang at the end. They added one of the telomere-binding proteins, TRF2, then subjected the complex to electron microscopy. Figure 21.26a shows that a loop really did form, with a ball of TRF2 protein right at the loop–tail junction. Such structures appeared about 20% of the time. By contrast, when these workers cut

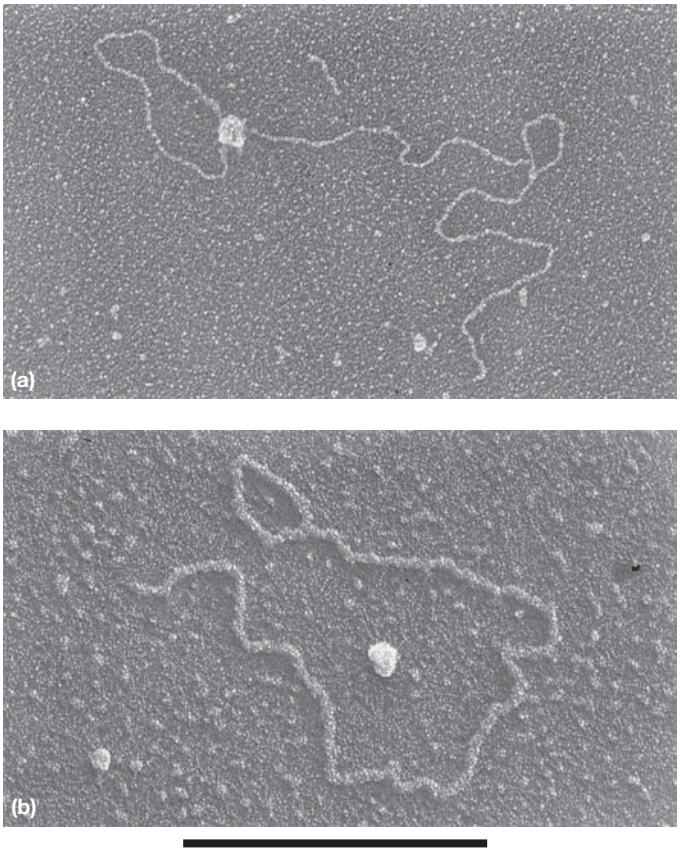


Figure 21.26 Formation of t-loops in vitro. (a) Direct detection of loops. Griffith and colleagues mixed a model DNA having a telomere-like structure with TRF2, then spread the mixture on an EM grid, shadowed the DNA and protein with tungsten, and observed the shadowed molecules with an electron microscope. An obvious loop appeared, with a blob of TRF2 at the junction between the loop and the tail. **(b)** Stabilization of the loop by cross-linking. Griffith and coworkers formed the t-loop as in panel (a), then cross-linked double-stranded DNA with psoralen and UV radiation, then removed the protein, spread the cross-linked DNA on an EM grid, shadowed with platinum and paladium, and visualized the shadowed DNA with an electron microscope. Again, an obvious loop appeared. The bar represents 1 kb. (*Source:* Griffith, J.D., L. Comeau, S. Rosenfield, R.M Stansel, A. Bianchi, H. Moss, and T. de Lange, Mammalian telomeres end in a large duplex loop. *Cell* 97 (14 May 1999) f. 1, p. 504. Reprinted by permission of Elsevier Science.)

off the single-stranded 3′-overhang, or left out TRF2, they found a drastic reduction in loop formation.

One way for a telomere to form such a loop would be for the single-stranded 3′-overhang to invade the double-stranded telomeric DNA upstream, as depicted in Figure 21.27. If this hypothesis is correct, one should be able to stabilize the loop with psoralen and UV radiation, which cross-link thymines on opposite strands of a double-stranded DNA. Because the invading strand base-pairs with one of the strands in the invaded DNA, this creates double-stranded DNA that is subject to cross-linking and therefore stabilization. Figure 21.26b shows the results of an experiment in which Griffith, de Lange, and coworkers cross-linked the model DNA with psoralen and UV, then

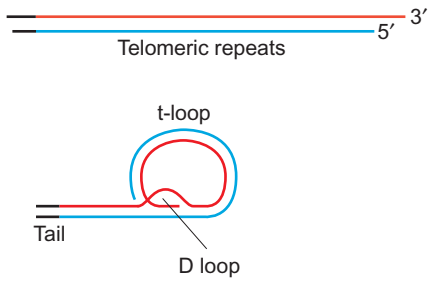


Figure 21.27 A model of a mammalian t-loop. The single-stranded 3′-end of the G-rich strand (red) invades the double-stranded telomeric DNA upstream, forming a long t-loop and a 75–200-nt displacement loop at the junction between the loop and the tail. A short subtelomeric region (black) is pictured adjoining the telomere (blue and red). (*Source:* Adapted from Griffith, D., L. Comeau, S. Rosenfield, R.M. Stansel, A. Bianchi, H. Moss, and T. de Lange, Mammalian telomeres end in a large duplex loop. *Cell,* 97:511, 1999).

deproteinized the complex, then subjected it to electron microscopy. The loop is still clearly visible, even in the absence of TRF2, showing that the DNA itself has been cross-linked, stabilizing the t-loop.

Next, these workers purified natural telomeres from several human cell lines and from mouse cells and subjected them to psoralen–UV treatment and electron microscopy. They obtained the same result as in Figure 21.26b, showing that t-loops appear to form in vivo. Furthermore, the sizes of these putative t-loops correlated well with the known lengths of the telomeres in the human or mouse cells, reinforcing the hypothesis that these loops really do represent telomeres.

To test further the notion that the loops they observed contain telomeric DNA, Griffith, de Lange and colleagues added TRF1, which is known to bind very specifically to double-stranded telomeric DNA, to their looped DNA. They observed loops coated with TRF1, as shown in Figure 21.28a.

If the strand invasion hypothesis in Figure 21.27 is valid, the single-stranded DNA displaced by the invading DNA (the displacement loop, or D-loop) should be able to bind *E. coli* single-strand-binding protein (SSB, recall Chapter 20) if the displaced DNA is long enough. Figure 21.28b demonstrates that SSB is indeed visible, right at the tail–loop junction. That is just where the hypothesis predicts we should find the displaced DNA.

Shelterin is essential for t-loop formation. In particular, TRF2 can form t-loops in a model DNA substrate. However, this remodeling reaction is weak in the absence of the other shelterin subunits. TRF1, the other telomere repeat-binding protein, is especially helpful, as it can bend, loop, and pair telomeric repeats. It is striking that this remodeling reaction can occur in vitro even in the absence of ATP. Based on all we know about shelterin proteins, de Lange proposed the model for t-loop formation depicted in Figure 21.29. Figure 21.29a shows the members of the shelterin complex bound to an unlooped telomere. Figure 21.29b is a model for the interaction of shelterin with a t-loop.

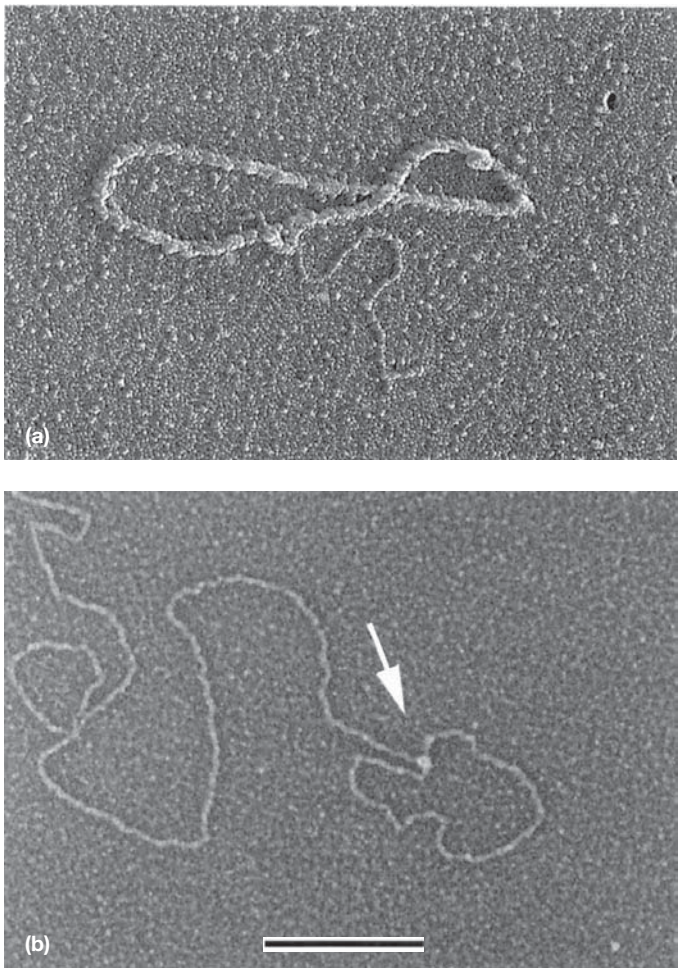


Figure 21.28 Binding of TRF1 and SSB to t-loops. (a) TRF1. Griffith, de Lange, and colleagues purified natural HeLa cell t-loops, cross-linked them with psoralen and UV radiation, and added TRF1, which binds specifically to double-stranded telomeric DNA. Then they shadowed the loop with platinum and paladium and performed electron microscopy. The t-loop, but not the tail, is coated uniformly with TRF1. **(b)** SSB. These workers followed the same procedure as in panel (a), but substituted *E. coli* SSB for TRF1. SSB should bind to single-stranded DNA, and it was observed at the loop–tail junction (arrow), where the single-stranded displacement loop was predicted to be. The bar represents 1 kb. (*Source:* Griffith, J.D., L. Comeau, S. Rosenfield, R.M. Stansel, A. Bianchi, H. Moss, and T. de Lange, Mammalian telomeres end in a large duplex loop. *Cell* 97 (14 May 1999) f. 5, p. 510. Reprinted by permission of Elsevier Science.)

Figure 21.29b also hints at an explanation for the paradox that POT1 is a single-stranded telomere-binding protein and yet the single-stranded telomeric DNA is hidden in the t-loop. But the figure shows that formation of the t-loop also creates a D-loop, and the displaced single-stranded region is a potential binding site for POT1. There is also the possibility that not all mammalian telomeres form t-loops. Any telomeres that remain linear would provide obvious binding sites for POT1.

The second way shelterin affects the structure of telomeres is by determining the structure of the end of the

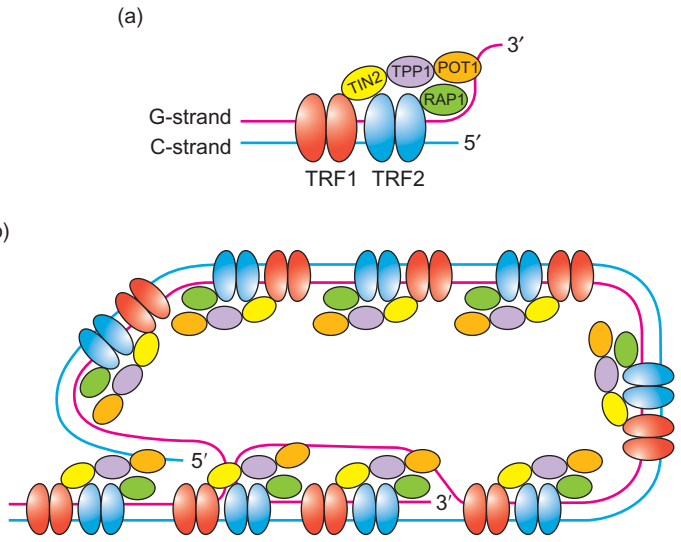


Figure 21.29 The shelterin-telomere complex. (**a**) Interaction with shelterin proteins and a linear telomere. TRF1 and TRF2 are shown interacting as dimers with the double-stranded part of the telomere, as POT1 interacts with the single-stranded part. The known interactions among shelterin proteins are also shown. (**b**) Model for the interaction of shelterin complexes with a t-loop. Colors are as in panel (a). Note the binding of POT1 (orange) to the single-stranded telomeric DNA in the D-loop, and the binding of TRF1 and TRF2 to the double-stranded telomeric DNA elsewhere in the t-loop.

telomere. It does this in two ways: by promoting 3′-end elongation, and protecting both the 5′- and 3′-ends from degradation. Finally, the third effect of shelterin on the structure of telomeres is to maintain telomere length within close tolerances. When the telomere gets too long, shelterin inhibits further telomerase action, limiting the growth of the telomere. POT1 plays a critical role in this process: When POT1 activity is eliminated, mammalian telomeres grow to abnormal lengths.

SUMMARY In mammals, telomeres are protected by a group of six proteins collectively known as shelterin. Two of the shelterin proteins, TRF1 and TRF2, bind to the double-stranded telomeric repeats. A third protein, POT1, binds to the single-stranded 3′-tail of the telomere. A fourth protein, TIN2, organizes shelterin by facilitating interaction between TRF1 and TRF2, and tethering POT1, via its partner, TPP1, to TRF2. Shelterin affects telomere structure in three ways: First, it remodels telomeres into t-loops, wherein the single-stranded 3′-tail invades the double-stranded telomeric DNA, creating a D-loop. In this way, the 3′-tail is protected. Second, it determines the structure of the telomeric end by promoting 3′-end elongation and protecting both 3′- and 5′-telomeric ends from degradation. Third, it maintains the telomere length within close tolerances.

Telomere Structure and Telomere-Binding Proteins in Lower Eukaryotes

Yeasts also have telomere-binding proteins, but they appear not to form t-loops. Thus, the proteins themselves must protect the telomere ends, without the benefit of hiding the single-stranded end within a D-loop. The fission yeast, *Schizosaccharomyces pombe,* has a group of telomere-binding proteins that resemble mammalian shelterin proteins. A protein called Taz1 plays the double-stranded telomere-binding role of mammalian TRF in fission yeast, and binds through Rap1 and Poz1 to a dimer of Tpz1 and Pot1. That resembles the TPP1-POT1 dimer in mammals, not only in structure, but in ability to bind to single-stranded telomeric DNA. These proteins can bind to a linear telomere, and they may also bend the telomere by 180 degrees by protein-protein interactions between proteins bound to the double-stranded telomere, and those bound to its single-stranded tail. This bending does not seem to form t-loops, however.

The budding yeast *Saccharomyces cerevisiae* also has telomere-binding proteins, but their evolutionary relationship to mammalian shelterin proteins is limited to one protein: Rap1. However, unlike mammalian RAP1, yeast Rap1 binds directly to double-stranded DNA, as the mammalian TRF proteins do. RAP1 has two partners, Rif1 and Rif2. In addition, a second protein complex, composed of Cdc13, Stn1, and Ten1, binds to the single-stranded telomeric tail.

Telomere-binding proteins were first discovered in the ciliated protozoan *Oxytricha*. This organism makes do with just two such proteins, TEBPα and TEBPβ, which are evolutionarily related to POT1 and TPP1 in mammals. These proteins bind to the single-stranded 3′-end of the organism's telomeres and protect them from degradation. By covering the ends of the telomeres, these proteins also prevent the telomeres from appearing like the ends of broken chromosomes—and all the negative consequences that would have.

SUMMARY Yeasts and ciliated protozoa do not form t-loops, but their telomeres are still associated with proteins that protect them. Fission yeasts have shelterin-like telomere-binding proteins, while budding yeasts have only one shelterin relative, Rap1, which binds to the double-stranded part of the telomere, plus two Rap1-binding proteins and three proteins that protect the single-stranded 3′-end of the telomere. The ciliated protozoan *Oxytricha* has only two telomere-binding proteins, which bind to the single-stranded 3′-ends of telomeres.

The Role of Pot1 in Protecting Telomeres In *S. pombe,* Pot1, instead of limiting the growth of telomeres, as mammalian POT1 does, plays a critical role in maintaining their integrity. Indeed loss of Pot1 can cause the loss of telomeres from this organism.

In 2001, Peter Baumann and Thomas Cech reported that they had found a protein in *S. pombe* that binds the single-stranded tails of telomeres. They named the *S. pombe* gene *pot1,* for protection of telomeres, and its product is now known as Pot1.

To test their hypothesis that *pot1* encodes a protein that protects telomeres, Baumann and Cech generated a $pot1^+/pot1^-$ diploid strain and germinated the spores from this strain. The $pot1^-$ spores gave rise to very small colonies compared with the colonies from $pot1^+$ spores. And the $pot1^-$ cells tended to be elongated, to show defects in chromosome segregation, and to stop dividing. All of these effects are consistent with loss of telomere function.

To test directly the effect of *pot1* on telomeres, Baumann and Cech looked for the presence of telomeres in $pot1^-$ strains by Southern blotting DNA from these strains and probing with a telomere-specific probe. Figure 21.30 shows the results. DNA from the $pot1^+$ strains, and from the diploid strains containing at least one $pot1^+$ allele, reacted strongly with the telomere probe, indicating the presence of telomeres. But DNA from the $pot1^-$ strains did not react with the probe, indicating that their telomeres had disappeared. Thus, the *pot1* gene product, Pot1p (or Pot1), really does seem to protect telomeres.

If Pot1 really protects telomeres, we would expect it to bind to telomeres. To check this prediction, Baumann and Cech cloned the *pot1* gene into an *E. coli* vector so it could be expressed as a fusion protein with a tag of six histidines (Chapter 4). They purified this fusion protein and used a gel mobility shift assay (Chapter 5) to detect its binding to either the C-rich or G-rich strand of the telomere, or a

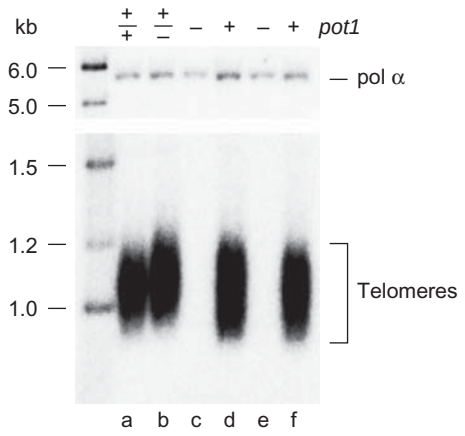


Figure 21.30 Fission yeast strains defective in *pot1* lose their telomeres. Baumann and Cech generated homozygous and heterozygous diploid, and $pot1^-$ and $pot1^+$ haploid strains of *S. pombe,* as indicated at top, then isolated DNA from these strains, digested the DNA with *Eco*RI, electrophoresed and Southern blotted the fragments, then probed the blot with a telomere-specific probe. As a control for uniform loading of the blot, the blot was also probed for DNA polymerase α, as indicated at top right. (*Source:* From Baumann and Cech, *Science* 292: p. 1172. © 2001 by the AAAS.)

double-stranded telomeric DNA. Figure 21.31a shows that Pot1 bound to the G-rich strand, but not to the C-rich or duplex DNA. Furthermore, an N-terminal fragment of Pot1 was even more effective in binding to the G-rich strand of the telomere (Figure 21.31b).

It is interesting that the phenotype of the $pot1^-$ strains, though it was originally quite aberrant, returned to normal after about 75 generations. The same effect had previously

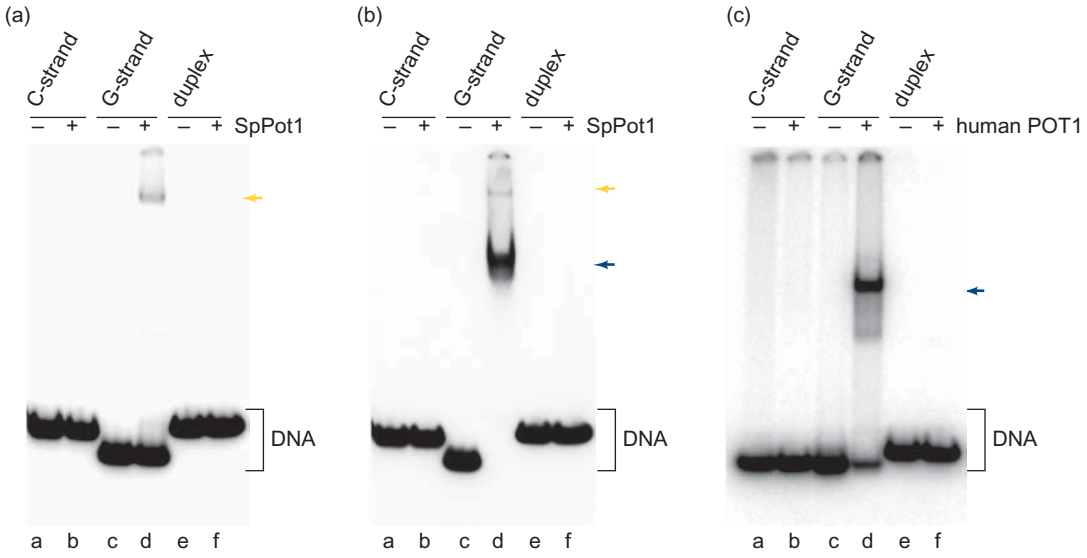


Figure 21.31 Pot1 binding to telomeric DNA. Baumann and Cech performed gel mobility shift experiments with *S. pombe* Pot1 and labeled *S. pombe* telomeric DNA **(a** and **b)** and human hPot1 and labeled human telomeric DNA **(c).** The telomeric DNA was either from the C-rich strand, the G-rich strand, or duplex DNA, as indicated at top. Panel **(a)** contained full-length Pot1. Panel **(b)** contained mostly an N-terminal fragment of Pot1, with slight contamination from full-length Pot1. Panel **(c)** contained an N-terminal fragment of human POT1. Arrows indicate the positions of shifted bands containing full-length Pot1 (yellow arrows) or N-terminal fragments of Pot1 or human POT1 (blue arrows). (*Source:* From Baumann and Cech, *Science* 292: p. 1172. © 2001 by the AAAS.)

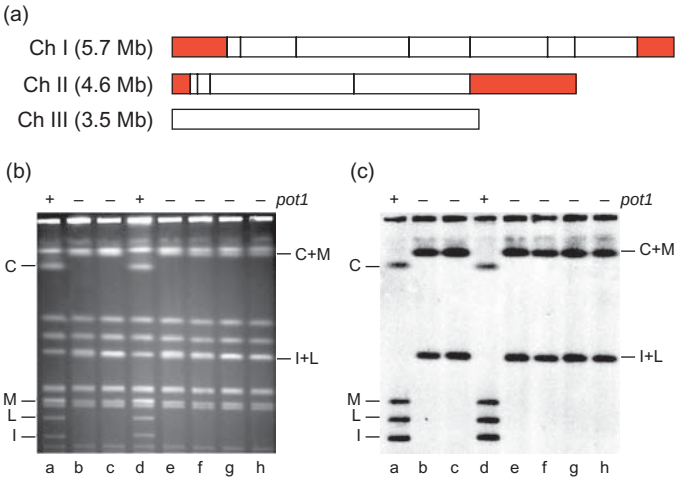


Figure 21.32 Surviving *Pot1*$^-$ stains have circularized chromosomes. (a) Maps of the three chromosomes of *S. pombe* showing the restriction sites for *Not*I as vertical lines. The terminal *Not*I fragments in chromosomes I and II are in red. Chromosome III is not cut by *Not*I. **(b)** Stained gel after pulsed-field gel electrophoresis of *Not*I DNA fragments from *pot1*$^+$ and *pot1*$^-$ cells, as indicated at top. The positions of terminal fragments (C, M, L, and I) of chromosomes I and II are indicated at left, and the positions of fused C+M and I+L fragments are indicated at right. **(c)** Baumann and Cech Southern blotted the gel from panel (b) and probed it with labeled DNA fragments C, M, L and I, representing the ends of chromosomes I and II. (*Source:* From Baumann and Cech. *Science* 292: p. 1172. © 2001 by the AAAS.)

been observed in strains lacking telomerase. This behavior can be explained if yeast chromosomes lacking telomeres can protect their ends by circularizing. To test this hypothesis, Baumann and Cech cleaved DNA from surviving *pot1*$^-$ strains with the rare cutter *Not*I (Chapter 4) and subjected the resulting DNA fragments to pulsed-field gel electrophoresis. If the chromosomes really had circularized, the *Not*I fragments at the ends of chromosomes should be missing and new fragments composed of the fused terminal fragments should appear. Figure 21.32 shows that this is exactly what happened for the two chromosomes tested, chromosomes I and II. The two fragments (I and L) normally at the ends of chromosome I were missing, and a new band (I+L), not present in *pot1*$^+$ strains, appeared. Similarly, the two fragments (C and M) normally at the ends of chromosome II were missing, and a new band (C+M) appeared. Thus, the chromosomes in *pot1*$^-$ strains really do circularize in response to loss of their telomeres.

The Role of Shelterin in Suppressing Inappropriate Repair and Cell Cycle Arrest in Mammals We have seen that telomeres prevent the cell from recognizing chromosome ends as chromosome breaks and invoking two processes that would threaten the life of the cell and even the organism. These processes are **homology-directed repair** (**HDR**) and nonhomologous end-joining (NHEJ, Chapter 20). HDR would promote homologous recombination between telomeres on separate chromosomes, or between telomeres and other chromosomal regions, resulting in potentially drastic shortening or lengthening of telomeres. The shortening would be especially dangerous because it could lead to loss of the whole telomere. NHEJ would lead to chromosome fusion, which is often lethal to the cell because the chromosomes do not separate properly during mitosis. If the cell doesn't die, the results could be even worse for the organism because they can lead to cancer.

In addition to HR and NHEJ, broken chromosomes also activate a checkpoint whereby the cell cycle can be arrested until the damage is repaired. If it is not repaired, the cells irreversibly enter a senescence phase and ultimately die, or they undergo a process called **apoptosis,** or programmed cell death, that results in rapid, controlled death of the cell. If normal chromosome ends invoked such a checkpoint, cells could not grow and life would cease. This is another reason that telomeres must prevent the cell from recognizing the normal ends of chromosomes as breaks.

Chromosome breaks do not by themselves activate cell cycle arrest. Instead, they are recognized by two protein kinases that autophosphorylate (phosphorylate themselves) and thereby initiate signal transduction pathways that lead to cell cycle arrest. One of these kinases is the ataxia telangiectasia mutated kinase (**ATM kinase**), which responds directly to unprotected DNA ends. Ataxia telangiectasia is an inherited disease caused by mutations in the ATM kinase gene. It is characterized by poor coordination (ataxia), prominent blood vessels in the whites of the eyes (telangiectasias), and susceptibility to cancer, among other symptoms.

The second kinase that senses chromosome breaks is the ataxia telangiectasia and Rad3 related kinase (**ATR kinase**), which responds to the single-stranded DNA end that appears when one DNA strand at a chromosome break is nibbled back by nucleases. As we have seen, mammalian telomeres have DNA ends that could activate the ATM kinase, and single-stranded DNA ends that could activate the ATR kinase, so both of these kinases need to be held in check at telomeres. How is this accomplished?

It is shelterin's job to repress both the ATM and ATR kinase at normal chromosome ends. One of shelterin's components, TRF2, represses the ATM kinase pathway. In fact, loss of TRF2 activity leads to the inappropriate activation of the ATM kinase at mammalian telomeres, which leads to cell cycle arrest. Another shelterin subunit, POT1, represses the ATR kinase pathway. When POT1 is inactivated, the ATM pathway remains repressed, but the ATR pathway is activated.

The simple formation of t-loops may explain the repression of the ATM pathway because the t-loops hide the DNA ends. However, t-loops cannot explain the repression of the ATR pathway, which is actually initiated by replication protein A (RPA), which binds directly to single-stranded DNA—and single-stranded DNA persists in the

D-loop part of a t-loop. Presumably, POT1 blocks binding of RPA to this single-stranded DNA simply by out-competing it for those binding sites. POT1 has an advantage over RPA in that it is automatically concentrated at telomeres by being part of the shelterin complex.

Shelterin also blocks the two DNA repair pathways that threaten telomeres: NHEJ and HDR. TRF2 represses NHEJ at telomeres during the G_1 phase of the cell cycle, before DNA replication, while POT1 and TRF2 team up to repress NHEJ at telomeres in the G_2 phase, after DNA replication. POT1 and TRF2 also collaborate to block HDR at telomeres. Ku (Chapter 20) can also block HDR at telomeres. This is interesting, because Ku's other role is to promote NHEJ when chromosomes are broken. Thus, telomeres must take advantage of Ku's ability to suppress HDR, while keeping in check its ability to promote NHEJ.

SUMMARY Unprotected chromosome ends would look like broken chromosomes and cause two potentially dangerous DNA repair activities, HDR and NHEJ. They would also stimulate two dangerous pathways (the ATM kinase and ATR kinase pathways) leading to cell cycle arrest. Two subunits of shelterin, TRF2 and POT1, block HDR and NHEJ. These two shelterin subunits also repress the two cell cycle arrest pathways. TRF2 represses the ATM kinase pathway, and POT1 represses the ATR kinase pathway.

SUMMARY

Primer synthesis in *E. coli* requires a primosome composed of the DNA helicase, DnaB, and the primase, DnaG. Primosome assembly at the origin of replication, *oriC,* occurs as follows: DnaA binds to *oriC* at sites called *dnaA* boxes and cooperates with RNA polymerase and HU protein in melting a DNA region adjacent to the leftmost *dnaA* box. DnaB then binds to the open complex and facilitates binding of the primase to complete the primosome. The primosome remains with the replisome, repeatedly priming Okazaki fragment synthesis, at least on the lagging strand. DnaB also has a helicase activity that unwinds the DNA as the replisome progresses.

The SV40 origin of replication is adjacent to the viral transcription control region. Initiation of replication depends on the viral large T antigen, which binds to a region within the 64-bp minimal *ori,* and at two adjacent sites, and exercises a helicase activity, which opens up a replication bubble within the minimal *ori.* Priming is carried out by a primase associated with the host DNA polymerase α.

The yeast origins of replication are contained within autonomously replicating sequences (ARSs) that are composed of four important regions (A, B1, B2, and B3). Region A is 15 bp long and contains an 11-bp consensus sequence that is highly conserved in ARSs. Region B3 may allow for an important DNA bend within ARS1.

The pol III holoenzyme synthesizes DNA at the rate of about 730 nt/sec in vitro, just a little slower than the rate of almost 1000 nt/sec observed in vivo. This enzyme is also highly processive, both in vitro and in vivo.

The pol III core (αε or αεθ) does not function processively by itself, so it can replicate only a short stretch of DNA before falling off the template. By contrast, the core plus the β-subunit can replicate DNA processively at a rate approaching 1000 nt/sec. The β-subunit forms a dimer that is ring-shaped. This ring fits around a DNA template and interacts with the α-subunit of the core to tether the whole polymerase and template together. This is why the holoenzyme stays on its template so long and is therefore so processive. The eukaryotic processivity factor PCNA forms a trimer with a similar ring shape that can encircle DNA and hold DNA polymerase on the template.

The β-subunit needs help from the γ complex (γ, δ, δ′, χ, and ψ) to load onto the complex. The γ complex acts catalytically in forming this processive αεβ complex, so it does not remain associated with the complex during processive replication. Clamp loading is an ATP-dependent process.

The pol III holoenzyme is double-headed, with two core polymerases attached through two τ-subunits to a γ complex. One core is responsible for (presumably) continuous synthesis of the leading strand, the other performs discontinuous synthesis of the lagging strand. The γ complex serves as a clamp loader to load the β clamp onto a primed DNA template. Once loaded, the β clamp loses affinity for the γ complex and associates with the core polymerase to help with processive synthesis of an Okazaki fragment. Once the fragment is completed, the β clamp loses affinity for the core polymerase and associates with the γ complex, which acts as a clamp unloader, removing the clamp from the DNA. Then it can recycle to the next primer and repeat the process.

At the end of replication, circular bacterial chromosomes form catenanes that must be decatenated for the two daughter duplexes to separate. In *E. coli* and related bacteria, topoisomerase IV performs this decatenation. Linear eukaryotic chromosomes also require decatenation during DNA replication.

Eukaryotic chromosomes have special structures known as telomeres at their ends. One strand of these telomeres is composed of many tandem repeats of short, G-rich regions whose sequence varies from one species to another. The G-rich telomere strand is made by an enzyme

called telomerase, which contains a short RNA that serves as the template for telomere synthesis. The C-rich telomere strand is synthesized by ordinary RNA-primed DNA synthesis, like the lagging strand in conventional DNA replication. This mechanism ensures that chromosome ends can be rebuilt and therefore do not suffer shortening with each round of replication.

In mammals, telomeres are protected by a group of six proteins collectively known as shelterin. Two of the shelterin proteins, TRF1 and TRF2, bind to the double-stranded telomeric repeats. A third protein, POT1, binds to the single-stranded 3′-tail of the telomere. A fourth protein, TIN2, organizes shelterin by facilitating interaction between TRF1 and TRF2, and tethering POT1, via its partner, TPP1, to TRF2. Shelterin affects telomere structure in three ways: First, it remodels telomeres into t-loops, wherein the single-stranded 3′-tail invades the double-stranded telomeric DNA, creating a D-loop. In this way, the 3′-tail is protected. Second, it determines the structure of the telomeric end by promoting 3′-end elongation and protecting both 3′- and 5′-telomeric ends from degradation. Third, it maintains the telomere length within close tolerances.

Yeasts and ciliated protozoa do not form t-loops, but their telomeres are still associated with proteins that protect them. Fission yeasts have shelterin-like telomere-binding proteins, while budding yeasts have only one shelterin relative, Rap1, which binds to the double-stranded part of the telomere, plus two Rap1-binding proteins and three proteins that protect the single-stranded 3′-end of the telomere. The ciliated protozoan *Oxytricha* has only two telomere-binding proteins, which bind to the single-stranded 3′-ends of telomeres.

Unprotected chromosome ends would look like broken chromosomes and cause two potentially dangerous DNA repair activities, HDR and NHEJ. They would also stimulate two dangerous pathways (the ATM kinase and ATR kinase pathways) leading to cell cycle arrest. Two subunits of shelterin, TRF2 and POT1, block HDR and NHEJ. These two shelterin subunits also repress the two cell cycle arrest pathways. TRF2 represses the ATM kinase pathway, and POT1 represses the ATR kinase pathway.

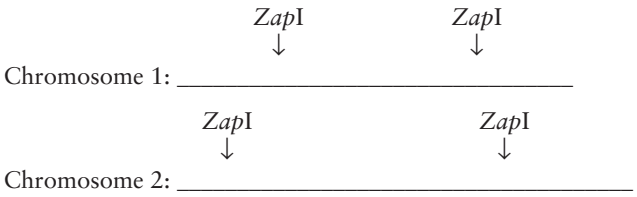
REVIEW QUESTIONS

1. Describe an assay to locate and determine the minimal length of an origin of replication.
2. List the components of the *E. coli* primosome and their roles in primer synthesis.
3. Outline a strategy for locating the SV40 origin of replication.
4. Outline a strategy for identifying an autonomously replicating sequence (ARS1) in yeast.
5. Outline a strategy to show that DNA replication begins in ARS1 in yeast.
6. Describe and give the results of an experiment that shows the rate of elongation of a DNA strand in vitro.
7. Describe a procedure to check the processivity of DNA synthesis in vitro.
8. Which subunit of the pol III holoenzyme provides processivity? What proteins load this subunit (the clamp) onto the DNA? To which core subunit does this clamp bind?
9. Describe and give the results of an experiment that shows the different behavior of the β clamp on circular and linear DNA. What does this behavior suggest about the mode of interaction between the clamp and the DNA?
10. What mode of interaction between the β clamp and DNA do x-ray crystallography studies suggest?
11. What mode of interaction between PCNA and DNA do x-ray crystallography studies suggest?
12. Describe and give the results of an experiment that shows that the clamp loader acts catalytically. What is the composition of the clamp loader?
13. Outline a hypothesis to explain how the clamp loader uses ATP energy to open the β clamp to allow entry to DNA.
14. How can discontinuous synthesis of the lagging strand keep up with synthesis of the leading strand?
15. Describe and give the results of an experiment that shows that pol III* can dissociate from its β clamp.
16. Describe a protein footprinting procedure. Show how such a procedure can be used to demonstrate that the pol III core and the clamp loader both interact with the same site on the β clamp.
17. Describe and give the results of an experiment that shows that the γ complex has clamp-unloading activity.
18. Describe how the β clamp cycles between binding to the core pol III and to the clamp unloader during discontinuous DNA replication.
19. Why is decatenation required after replication of circular DNAs?
20. Outline the evidence that topoisomerase IV is required for decatenation of plasmids in *Salmonella typhimurium* and *E. coli*.
21. Why do eukaryotes need telomeres, but prokaryotes do not?
22. Diagram the process of telomere synthesis.
23. Why was *Tetrahymena* a good choice of organism in which to study telomerase?
24. Describe an assay for telomerase activity and show sample results.
25. Describe and give the results of an experiment that shows that the telomerase RNA serves as the template for telomere synthesis.
26. Diagram the t-loop model of telomere structure.
27. What evidence supports the existence of t-loops?

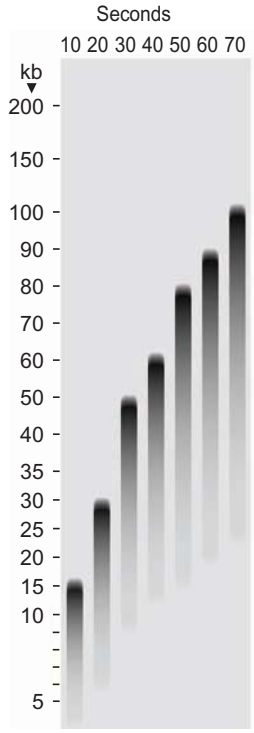
28. What evidence supports the strand-invasion hypothesis of t-loop formation?
29. Present a model for the structure of mammalian shelterin, showing each of the subunits, and how they participate in t-loop formation.
30. How does mammalian shelterin protect chromosome ends from HDR and NHEJ and block the two pathways leading to cell cycle arrest? What would be the consequences of failure to block each of these pathways?

ANALYTICAL QUESTIONS

1. Starting with the nucleotide sequence of the *hpot1* gene (or the amino acid sequence of hPot1) from humans, describe how you would search for a homologous gene (or protein) in another organism whose genome has been sequenced, such as the nematode *Caenorhabditis elegans*. Then describe how you would obtain the protein and test it for Pot1 activity.
2. You are investigating the *pot1* gene of a newly-discovered protozoan species. You find that cells with a defective *pot1* gene return to normal after 50 generations. Wild-type cells have only two chromosomes with the following restriction maps with respect to the restriction enzyme *Zap*I:



   Propose a hypothesis to explain how the mutant cells returned to normal, and describe an experiment you would perform to test it. Show the results you would obtain if your hypothesis is correct.
3. You are studying a eukaryotic virus with a 130-kb double-stranded DNA genome. You suspect that it has more than one origin of replication. Propose an experiment to test your hypothesis and find all of the origins.
4. You are investigating DNA replication in a new species of bacteria. You discover that this organism has a β clamp and pol III*, similar to their counterparts in *E. coli*. You want to know whether this β clamp and pol III* separate during idling and after termination on a model template. Describe the experiment you would use to answer this question. Include the assay for separation you would use, and present sample results.
5. You are investigating the elongation rate during replication of the DNA from a new extreme thermophile, *Rapidus royi*. Here are the results of electrophoresis on DNA elongated in vitro for various times. What is the elongation rate? Does it set a new world record?



6. Assuming they could be made in eukaryotes, what would be the advantages and disadvantages of primers made of DNA, rather than RNA? Would such primers eliminate the need for telomeres?

SUGGESTED READINGS

General References and Reviews

Baker, T.A. 1995. Replication arrest. *Cell* 80:521–24.
Blackburn, E.H. 1990. Telomeres: Structure and synthesis. *Journal of Biological Chemistry* 265:5919–21.
Blackburn, E.H. 1994. Telomeres: No end in sight. *Cell* 77:621–23.
Cech, T. R. 2004. Beginning to understand the end of the chromosome. *Cell* 116:273–79.
de Lange, T. 2001. Telomere capping—one strand fits all. *Science* 292:1075–76.
de Lange, T. 2005. Shelterin, the protein complex that shapes and safeguards human telomeres. *Genes and Development* 19:2100–10.
de Lange, T. 2009. How telomeres solve the end-protection problem. *Science* 326:948–52.
Ellison, V. and B. Stillman. 2001. Opening of the clamp: An intimate view of an ATP-driven biological machine. *Cell* 106:655–60.
Greider, C.W. 1999. Telomeres do D-loop-T-loop. *Cell* 97:419–22.
Herendeen, D.R. and T.J. Kelly. 1996. DNA polymerase III: Running rings around the fork. *Cell* 84:5–8.
Kornberg, A. and T.A. Baker. 1992. *DNA Replication,* 2nd ed. New York: W.H. Freeman.

Marx, J. 1994. DNA repair comes into its own. *Science* 266:728–30.
Marx, J. 1995. How DNA replication originates. *Science* 270:1585–86.
Marx, J. 2002. Chromosome end game draws a crowd. *Science* 295:2348–51.
Newlon, C.S. 1993. Two jobs for the origin replication complex. *Science* 262:1830–31.
Stillman, B. 1994. Smart machines at the DNA replication fork. *Cell* 78:725–28.
Wang, J.C. 1991. DNA topoisomerases: Why so many? *Journal of Biological Chemistry* 266:6659–62.
West, S.C. 1996. DNA helicases: New breeds of translocating motors and molecular pumps. *Cell* 86:177–80.
Zakian, V.A. 1995. Telomeres: Beginning to understand the end. *Science* 270:1601–6.

Research Articles

Arai, K. and A. Kornberg. 1979. A general priming system employing only *dnaB* protein and primase for DNA replication. *Proceedings of the National Academy of Sciences USA* 76:4309–13.
Arai, K., R. Low, J. Kobori, J. Shlomai, and A. Kornberg. 1981. Mechanism of *dnaB* protein action V. Association of *dnaB* protein, protein n′, and other prepriming proteins in the primosome of DNA replication. *Journal of Biological Chemistry* 256:5273–80.
Baumann, P. and T. Cech. 2001. Pot 1, the putative telomere end-binding protein in fission yeast and humans. *Science* 292:1171–75.
Blackburn, E.H. 1990. Functional evidence for an RNA template in telomerase. *Science* 247:546–52.
Blackburn, E.H. 2001. Switching and signaling at the telomere. *Cell* 106:661–73.
Bouché, J.-P., L. Rowen, and A. Kornberg. 1978. The RNA primer synthesized by primase to initiate phage G4 DNA replication. *Journal of Biological Chemistry* 253:765–69.
Brewer, B.J. and W.L. Fangman. 1987. The localization of replication origins on ARS plasmids in *S. cerevisiae*. *Cell* 51:463–71.
Georgescu, R.E., S.-S. Kim, O. Yuryieva, J. Kuriyan, X.-P. Kong, and M. O'Donnell. 2008. Structure of a sliding clamp on DNA. *Cell* 132:43–54.
Greider, C.W. and E.H. Blackburn. 1985. Identification of a specific telomere terminal transferase activity in *Tetrahymena* extracts. *Cell* 43:405–13.
Greider, C.W. and E.H. Blackburn. 1989. A telomeric sequence in the RNA of *Tetrahymena* telomerase required for telomere repeat synthesis. *Nature* 337:331–37.
Griffith, J.D., L. Comeau, S. Rosenfield, R.M. Stansel, A. Bianchi, H. Moss, and T. de Lange. 1999. Mammalian telomeres end in a large duplex loop. *Cell* 97:503–19.
Jeruzalmi, D., M. O'Donnell, and J. Kuriyan. 2001. Crystal structure of the processivity clamp loader gamma (γ) complex of *E. coli* DNA polymerase III. *Cell* 106:429–41.
Jeruzalmi, D., O. Yurieva, Y. Zhao, M. Young, J. Stewart, M. Hingorani, M. O'Donnell, and J. Kuriyan. 2001. Mechanism of processivity clamp opening by the delta subunit wrench of the clamp loader complex of *E. coli* DNA polymerase III. *Cell* 106:417–28.
Kong, X.-P., R. Onrust, M. O'Donnell, and J. Kuriyan. 1992. Three-dimensional structure of the β subunit of *E. coli* DNA polymerase III holoenzyme: A sliding DNA clamp. *Cell* 69:425–37.
Krishna, T.S.R., X.-P. Kong, S. Gary, P.M. Burgers, and J. Kuriyan. 1994. Crystal structure of the eukaryotic DNA polymerase processivity factor PCNA. *Cell* 79:1233–43.
Marahrens, Y. and B. Stillman. 1992. A yeast chromosomal origin of DNA replication defined by multiple functional elements. *Science* 255:817–23.
Mok, M. and K.J. Marians. 1987. The *Escherichia coli* preprimosome and DNA B helicase can form replication forks that move at the same rate. *Journal of Biological Chemistry* 262:16644–54.
Naktinis, V., J. Turner, and M. O'Donnell. 1996. A molecular switch in a replication machine defined by an internal competition for protein rings. *Cell* 84:137–45.
Stukenberg, P.T., P.S. Studwell-Vaughan, and M. O'Donnell. 1991. Mechanism of the sliding β-clamp of DNA polymerase III holoenzyme. *Journal of Biological Chemistry* 266:11328–34.