DNA replication at high resolution

James L Keck and James M Berger

Several decades of research have delineated the roles of many proteins central to DNA replication. Here we present a structural perspective of this work spanning the past 15 years and highlight several recent advances in the field.

Address: Department of Molecular and Cell Biology, University of California, Berkeley, 229 Stanley Hall #3206, Berkeley, CA 94720, USA.

Correspondence: James M Berger
E-mail: jmberger@uclink4.berkeley.edu

Chemistry & Biology 2000, **7**:R63–R71

1074-5521/00/$ – see front matter

Introduction

Replication, the duplication of an organism's genetic material, is a fundamental process in biology. From viruses to multicellular organisms, life is inseparable from the genetic instructions that maintain and regulate its existence. Strong evolutionary pressures have, therefore, forced cells and viruses to ensure that the timing of replication is highly coordinated and that the process itself proceeds with as few errors as possible. Over the years, researchers have extensively examined these events, transforming replication into one of the most well-studied fields in biology.

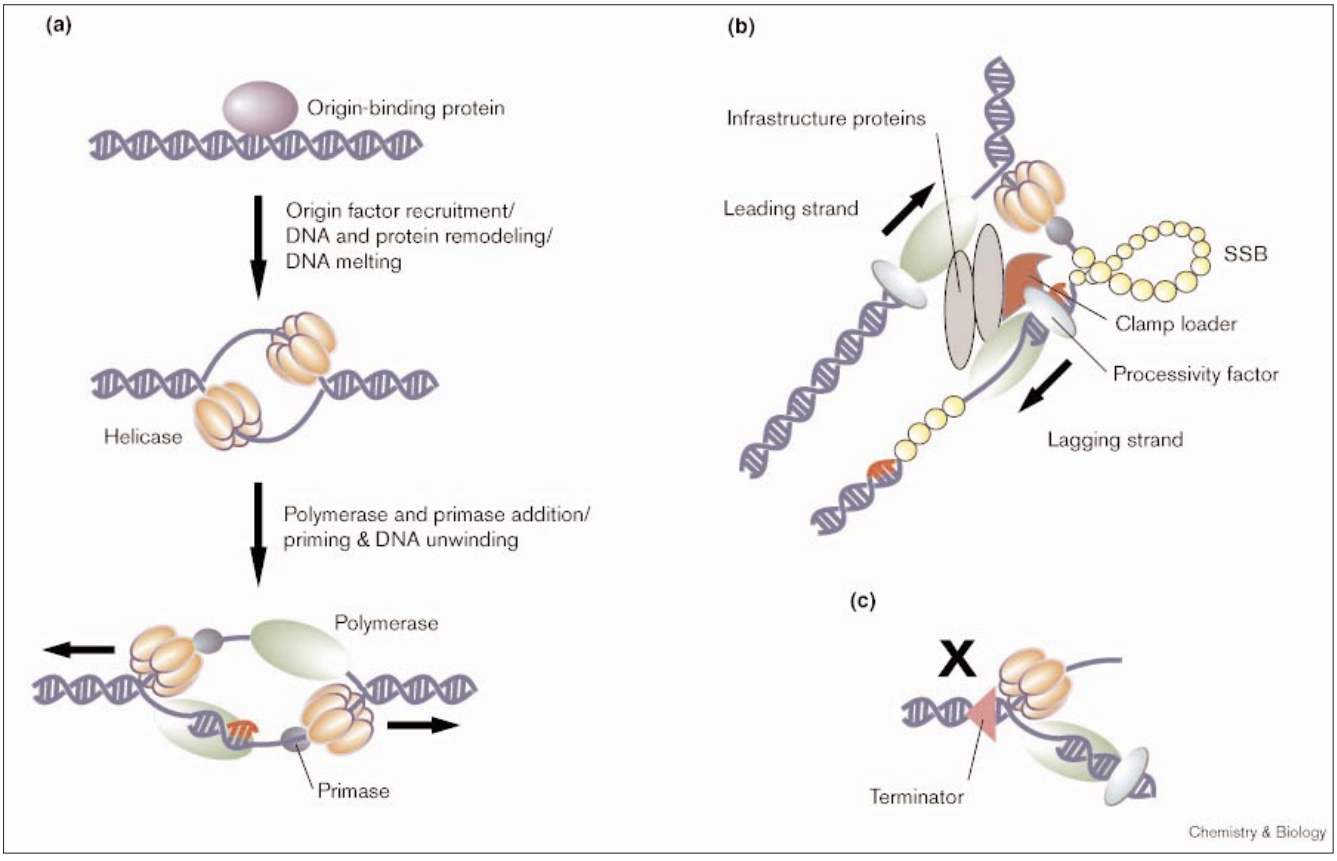
Initially, replication research focused on prokaryotes and bacteriophage, using genetic methods to broadly delineate and define the roles of numerous proteins central to the process. These experiments demonstrated that some replication proteins act only in initiation, whereas others are required throughout the reaction cycle [1]. As methods in molecular biology were invented and improved, researchers sharpened this view by studying individual replication proteins and protein complexes *in vitro*. The field has further expanded to include the complex reactions of eukaryotic replication. Together, these studies have helped develop working models for the catalytic properties of several replication proteins and have demonstrated that many of the components act cooperatively as multiprotein complexes.

Over the past 15 years, structural studies have begun to resolve this picture further, delivering high-resolution, three-dimensional images of many of the individual proteins central to replication. These models have allowed researchers to interpret the abundant genetic and biochemical data generated from decades of replication research, as well as to formulate novel hypotheses about the control and execution of the process. Breakthroughs continue today in all areas of replication study and have been the subjects of many recent reviews [2–7]. Here we will focus primarily on several advances in our understanding of prokaryotic and bacteriophage DNA replication that have been made using high-resolution structural biology. Most of the reactions of bacterial and phage replication have functional analogs in eukaryotic biology and are therefore meaningful across all kingdoms of life.

Bacterial DNA replication: a multiprotein event

DNA replication is a multi-staged reaction (Figure 1). First, a replication origin sequence within the genome is recognized by origin-binding protein components. Second, initiation of replication occurs through the recruitment of

Figure 1



Schematic diagram of bacterial DNA replication. Factors in grey are of unknown structure whereas factors in color have at least one known structure. Domains of primase were not discussed in this review but have been determined recently ([77]; J.M.B. *et al.*, unpublished observations). The color-coding scheme used here is employed throughout the other figures. Nucleic acids are shown either in blue (DNA) or red (RNA). Three stages of replication are shown based on current biochemical models in the literature: **(a)** initiation [3,7], **(b)** propagation [3,7,18] and **(c)** termination [73]. See the text for more details.

replisomal proteins and the assembly of replication forks at the origin. Third, the general replication reaction duplicates both strands of DNA. Fourth, replication terminates and the two daughter chromosomes are separated. We will describe these processes as they occur in the bacterium *Escherichia coli* before delving into the structures of known replication proteins.

Replication initiation in *E. coli* begins with the recognition of a short origin DNA sequence ('ori') by an initiator protein called DnaA (Figure 1a). DnaA binding alters the structure of the 'ori' DNA and provides an associative platform for targeting additional replication proteins [8,9]. The origin is then actively assembled into a bi-directional replication bubble with two oppositely directed replication forks. This process involves the recruitment of one helicase (DnaB) per replication fork by association with DnaA [10,11]. DnaB attracts a second important replication-fork enzyme called 'primase' (DnaG), which is a specialized, replication-priming RNA polymerase. Together, the helicase and primase form the core of an assembly referred to as the 'primosome' [12]. Because DNA polymerases are incapable of initiating DNA chains *de novo*, the primase portion of the primosome makes a short RNA chain, called a primer, from which DNA polymerase can extend its first few nucleotides. During these initial stages, and throughout the entire course of replication, the DnaB helicase catalyzes the unwinding of genomic DNA in an ATP-dependent reaction [13].

Once the DNA is melted and primed, the general DNA replication reaction begins (Figure 1b). The central enzyme in this stage of replication is the DNA polymerase, which synthesizes a single DNA strand by covalently linking together appropriate deoxynuclotide triphosphates (dNTPs) that are complementary to the template. Because all nucleic-acid polymerases add nucleoside triphosphates (NTPs) only to 3′-hydroxyl moieties on growing nucleic-acid chains, there is a strict directionality to DNA polymerase activity in the 5′→3′ direction.

This restraint forces replication of the antiparallel DNA strands to be asymmetric: one newly synthesised strand is made continuously (leading strand), while the other is synthesized in segments (lagging strand) with a new DNA polymerase initiating every ~1000 nucleotides [2]. In both cases, DNA polymerase III (pol III) is responsible for extending the nucleic-acid chains in *E. coli* [14,15]. Replication fork asymmetry also makes DNA polymerase initiation by primase obligatory throughout replication to ensure full duplication of the lagging strand. In order to maintain association between pol III and DNA, specialized proteins known as processivity factors act as 'sliding clamps' on duplex DNA. These proteins are loaded onto primed DNA by 'clamp-loader' proteins [16,17] and then associate with the polymerase, providing an additional tether between the polymerase and DNA. The activities of the leading- and lagging-strand polymerases in a replication fork are coordinated by a protein infrastructure forming a large complex referred to as the polymerase holoenzyme (pol-HE) [18]. As a replication fork progresses, single-stranded DNA (ssDNA) intermediates that are generated by the helicase are bound by ssDNA-binding proteins (SSBs) in a restricted conformation that ensures replication can proceed swiftly [2].

After copying the host chromosome, replication must terminate (Figure 1c). This process occurs in an organized manner, although certain details regarding this stage of replication are not as well understood as are initiation and elongation. In the *E. coli* chromosome, termination proteins bind specific DNA sequences and stop replication by inhibiting helicase translocation [19,20]. Following termination, the two completed copies of the genome are untangled by topoisomerases to allow independent segregation of the two daughter chromosomes. At this end point, the cell has catalyzed the replication of 4.6×10^6 basepairs of genomic DNA in 40 minutes, at the astonishing rate of 1000 nucleotides per second per replication fork [3].

Structures in DNA replication

Over the past 15 years, X-ray crystallographic and nuclear magnetic resonance (NMR) studies have resulted in the high-resolution structure determination of at least one example of nearly every protein involved in DNA replication. In Figure 1, replication proteins and complexes are colored to highlight known high-resolution structures; all proteins with color have at least one representative structure that defines the three-dimensional shape of the molecule. The overwhelming number of proteins for which structures have been solved indicates that DNA replication is fast becoming one of the best understood processes in biology at the molecular level. To provide an overview of the molecular details of replication, we describe examples of known replication protein structures, with a particular emphasis on recent examples from bacteria and phage.

Origin-binding proteins

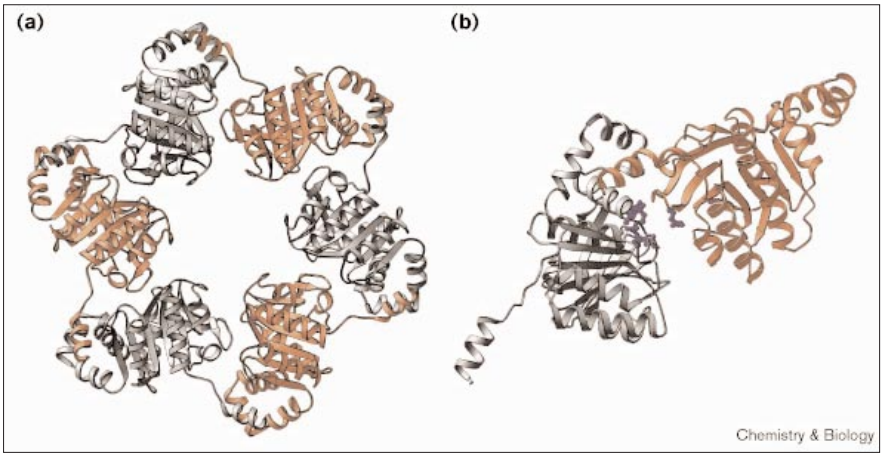
Although no structures have been solved for cellular origin-binding proteins, two such structures from viruses [21–23] and one from a plasmid replication system [24] have been described. The structures of these proteins are relatively distinct. All are seen to slightly bend DNA upon binding, but the functional significance of this feature is unknown. As these structures have recently been reviewed [25], we will not discuss them here.

Helicases

The double-stranded structure of genomic DNA is an impediment to replication because it buries the sequence information required to copy each strand. This feature makes unwinding double-stranded DNA (dsDNA) a requirement for replication, as well as for other types of reactions involving duplex nucleic acids, such as transcription. Because DNA unwinding is energetically unfavorable, cells have evolved enzymes known as helicases that couple the energy of NTP binding and hydrolysis to unwinding [26]. Families of helicases share several sequence and structural motifs, implying that there is a common unwinding mechanism used by these enzymes [27]. The differences that are observed between helicase groups are thought to be important for the specific roles that the enzymes serve *in vivo*. For example, one feature common to known replicative helicases is that they form hexameric rings that encircle DNA. Such a structure may be necessary for processive unwinding of DNA at replication forks.

Several high-resolution structures of helicases have been described [28–32], and recently the first high-resolution glimpse of a replicative hexameric helicase (T7 bacteriophage gp4 protein) was published [33]. The full-length T7 gp4 protein contains both primase and helicase activities as individual domains in a single protein [34]. These domains share sequence homology to the bacterial DnaB and DnaG families, respectively [35]. Similar to DnaB, gp4 oligomerizes to an active hexameric form and can unwind DNA processively and rapidly in the context of its replisome. Because all previously solved helicase structures are monomeric [28–32], the structural basis for this oligomerization has been unclear. To examine the structure of gp4, Ellenberger and coworkers [33] crystallized a truncated version of the protein (gp4-hel) that removed the primase domain and portions of the helicase domain. The effect of these truncations is the loss of both primase and helicase activities, but because an NTP hydrolysis activity was retained in the minimal construct, gp4-hel can be considered a good first representation of a hexameric helicase structure [33]. Solution of the gp4 structure represents a milestone in our understanding of replication at a structural level as it affords a visual explanation of many features unique to hexameric helicases.

Figure 2



Structure of the helicase domain of the T7 gp4 protein. **(a)** Ribbon diagram looking down the 6_1-symmetry axis of gp4-hel [33] (accession number: 1CR1). Although this aspect gives the appearance of a closed ring, the oligomeric structure is actually that of a right-handed filament. Monomers are alternatively colored grey or red to highlight the mode of oligomerization. **(b)** Detailed view of the shared thymidine triphosphate (TTP) binding site between two gp4-hel protomers. TTP is bound primarily by the protomer in grey, but a catalytically important arginine sidechain is contributed from the protomer in red. TTP and the arginine are shown in blue ball-and-stick models. Figures were generated using RIBBONS [78].

The structure of gp4-hel reveals that the truncated molecule forms a helical filament with 6_1 symmetry, rather than a closed ring with true sixfold symmetry (Figure 2a). This arrangement is somewhat surprising because electron micrographs of full-length gp4 had previously revealed a closed-circular hexameric structure for the molecule [36]. However, the six adjacent protomers of one filament turn in the crystal can be rearranged with simple translations and other minor modifications to assemble a structure with true sixfold symmetry that is consistent with the lower-resolution electron micrographs of gp4 [33]. Furthermore, Wigley and coworkers have observed that a slightly longer gp4 helicase construct forms a closed-ring structure and is active as a helicase; the structure of this gp4 domain is currently in progress (D. Wigley, personal communication). The 6_1-symmetric arrangement of proteins within the crystal is strikingly similar to the symmetry observed for *E. coli* RecA, a ssDNA-binding recombination protein [37]. Structural similarity between RecA and other helicase structures has been noted before, but gp4-hel is the first helicase for which a RecA-like quatenary arrangement is also observed [28,38].

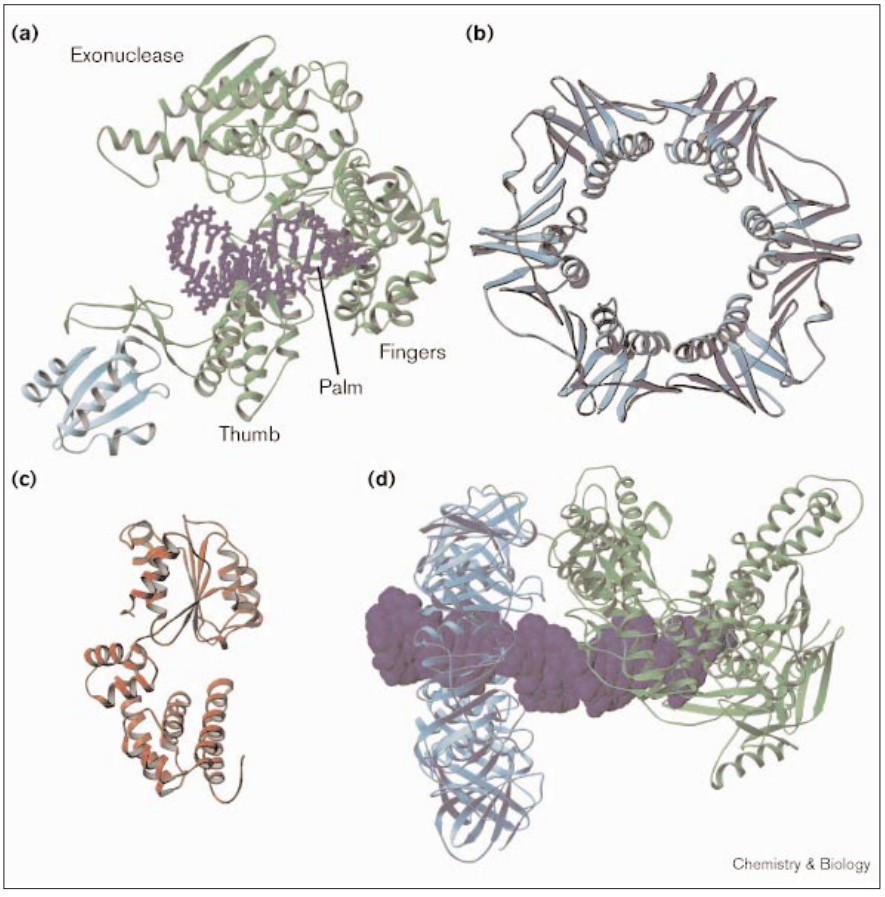
Oligomerization of the gp4-hel domains is accomplished primarily by domain swapping and, to a lesser extent, by shared active-site elements (Figure 2b). The amino-terminal 11 residues of each gp4 protomer are helical and form a subdomain with two helices that are present on the adjacent protomer. Domain swapping of this sort has been observed in a variety of oligomeric proteins and offers a means of coordinating the activities of one protomer with others in the complex [39]. Oligomeric helicases are well-noted for their cooperativity in nucleotide binding and hydrolysis [26]; the observation of a swapped helix linking helicase protomers offers an exciting new structural platform for understanding this enzymological feature.

All known helicases share sequence homology in two motifs important for NTP binding and hydrolysis, the Walker A and B boxes [27,40]. This homology also extends to three-dimensional similarity in the positions of active-site residues in known helicase structures, including gp4-hel [41]. However, the active-site of gp4-hel offers another clue explaining the cooperativity of NTP binding and hydrolysis in oligomeric helicases. Although the positions of the relevant conserved residues are consistent with monomeric helicase crystal structures, the gp4-hel active site is a composite region made up of conserved residues from two individual protomers. Specifically, a conserved active-site arginine residue is contributed from one protomer to another (Figure 2b). The result is that the NTP-binding and -hydrolysis site is shared between two protomers [33]; a shared active site is also observed in the structure of another hexameric ATPase, the F_1-ATP synthase [42]. Contrariwise, in monomeric helicase structures, the NTPase site is made up of two homologous domains contained within a single, contiguous polypeptide chain. It is therefore expected that cooperativity arises from the ability of residues in one protomer to 'sense' the presence or hydrolytic state of nucleotide in the active site of the adjacent protomer.

DNA polymerases

The first replication protein solved to high-resolution was the Klenow fragment of *E. coli* DNA polymerase I (pol I) [43]. Although this enzyme is not the replicative polymerase of *E. coli*, the pol I structure has been shown to define the general structure of all known replicative polymerases. Many excellent, up-to-date reviews of polymerase structure are available [44–48] so we will only briefly review their structural and functional highlights here. Over a dozen different DNA polymerase structures have been solved and each can be likened to a right hand with palm, thumb, and fingers domains (Figure 3a,d) [43]. The 'palm' portion of the polymerase acts as the catalytic center by binding two active-site divalent metals, as well as the 3′-hydroxyl nucleophile of the growing DNA chain

Figure 3



Structures of components in the polymerase holoenzyme. **(a)** T7 phage DNA polymerase (green) with bound *E. coli* thioredoxin processivity factor (light blue) and DNA (dark blue) [49]. Fingers, palm, thumb, and exonuclease regions are indicated. **(b)** Sliding clamp β subunit of *E. coli* pol-HE (light blue) [55]. **(c)** Clamp loader δ′ subunit of *E. coli* pol-HE (red) [60]. **(d)** Model of RB69 phage DNA polymerase–sliding clamp–DNA complex (green, light blue and dark blue, respectively) [51]. The accession numbers for (a–d) are 1T7P, 2POL, 1A5T, and1QE4, respectively. Figures were generated using RIBBONS [78].

and its α-phosphate target on the incoming dNTP [45]. This region of the polymerase forms a canonical fold that is shared across nearly all known polymerase structures, consisting of an antiparallel β sheet buttressed on one side with two α helices. The metal-liganding residues reside on the ends of two β strands of the sheet. The 'thumb' and 'fingers' regions form two canyon walls around the palm floor and contribute to template ssDNA, product dsDNA and dNTP binding. The folds of these domains vary significantly among polymerases but seem to carry out conserved mechanistic functions [45]. Many DNA polymerases also have exonuclease domains that either help proofread product dsDNA to ensure proper basepairing (3′→5′ exonuclease activity) or that remove the RNA primers at the 5′-initiation ends of DNA chains (5′→3′ exonuclease activity). The structures of several replicative DNA polymerases (from phage T7 [49] and RB69 [50,51], and from the archeabacterium *Thermococcus gorgonarius* [52]) have been determined to high-resolution, revealing an overall shape and metal-coordination geometry that are similar to pol I. These structural similarities demonstrate that the overall mechanism of polymerase activity is strongly conserved.

Processivity factors/clamp loaders

To ensure that DNA polymerases stay bound to their substrates throughout replication, cells and viruses have evolved proteins that enhance polymerase processivity [53,54]. These processivity factors fall into two structural categories: ring-shaped 'sliding clamps' that encircle dsDNA and are therefore topologically linked to the polymerase product, and globular accessory proteins that stabilize the polymerase–DNA complex (Figure 3a,b,d). Cells appear to use only the first mode of processivity enhancement. To date, a number of structures of sliding clamps have been determined: the dimeric β subunit of *E. coli* pol-HE [55], the trimeric gp45 protein from RB69 bacteriophage [51] and the trimeric proliferating cell nuclear antigen (PCNA) of *Saccharomyces cerevisiae* [56] and *Homo sapiens* [57].

Although the oligomerization state of clamps differs among species, their overall structures are strikingly similar (Figure 3b,d). The sliding clamps all form circular, pseudo-six-fold ring structures with a continuous β sheet framing the outside of the toroid and a display of α helices on the inside. Individual protomers of the

dimeric *E. coli* β protein have three nearly identical domains, whereas the trimeric clamps have two homologous domains per protomer. The result is that each clamp consists of six structurally similar domains, regardless of whether the functional form is dimeric or trimeric. The dimensions of the ring holes are similar and consistent with dsDNA fitting through the center (Figure 3d). It is interesting to note that the site for clamp–polymerase interactions has been identified in cocrystallization studies [51,57]. This identification has allowed one of the first reasonable attempts at building a high-resolution model of interacting replication proteins (see the section on future structural studies).

Besides the clamp mode of processivity enhancement, at least one phage DNA replication system uses the second strategy, namely that of binding a monomeric, globular protein that stabilizes the polymerase–DNA interaction. In the T7 phage replication system, the *E. coli* thioredoxin protein is coopted by the polymerase to act as an enhancer of replication processivity [58]. In this complex, thioredoxin binds to the DNA polymerase at the thumb domain and is thought to partially wrap around the dsDNA replication product [49] (Figure 3a). This mechanism therefore appears to be a simplified version of the sliding-clamp method.

Because sliding clamps form ring structures spontaneously, cell-replication systems require specialized clamp-loading proteins to open and position the clamp on DNA. Clamp loaders utilize cycles of ATP binding and hydrolysis to open sliding-clamp rings and load them onto primed DNA in a reaction that is coordinated with primase activity [18,59]. In *E. coli*, the clamp loader (γ complex) is composed of five subunits (γ, δ, δ′, χ and ψ). δ′ is the only component of known structure from this complex. δ′ forms a 'C-shaped' molecule with an NTP-binding site located in one of its domains [60] (Figure 3c). Although inactive as an ATPase, δ′ significantly resembles γ (the active clamp-loader ATPase subunit of the γ complex) in sequence and therefore helps in understanding the catalytic process of clamp loading [60]. It is thought that ATP binding and/or hydrolysis facilitates a conformational change in the γ subunit to cause the two domains of the 'C' to separate, thus opening the processivity ring and loading it onto primed DNA. The mechanistic details of this reaction are still unclear and await biochemical and structural insights into the workings of the γ–β supercomplex.

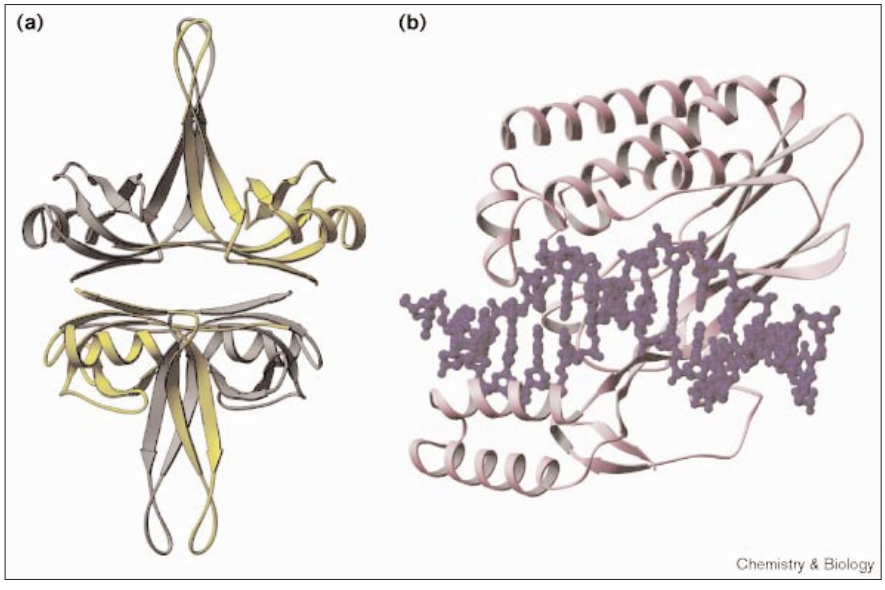
Single-stranded DNA-binding proteins

Unwinding of dsDNA by helicases produces ssDNA, the intermediate DNA structure used by polymerases in replication. This form of DNA is particularly susceptible to damage by nucleases and can also form intrastrand secondary structures that inhibit both DNA replication and other cellular DNA metabolic processes. To protect ssDNA and restrict its conformation, cells use SSBs to sequester the unpaired nucleic acid strands [61,62].

The structures of several SSBs have been determined, including phage T4 gp32 protein [63], *E. coli* SSB [64], the subunits of human replication protein A (RPA) [65,66], human mitochondiral SSB [67], several filamentous phage SSBs [68–70] and adenovirus SSB [71] (Figure 4a). Most of these structures share a common oligonucleotide/oligosaccharide-binding (OB) domain fold, which consists of a five-stranded, antiparallel β barrel that is puckered on one side. In known ssDNA–SSB complexes [63,65],

Figure 4

Structures of accessory replication proteins. **(a)** The *E. coli* SSB tetramer is shown in alternating yellow and grey protomers to highlight oligomerization [64] (accession number: 1KAW). **(b)** The *E. coli* tus–ter complex with tus shown in pink and the ter sequence in dark blue [76] (accession number: 1ECR). Figures were generated using RIBBONS [78].



ssDNA is bound in the puckered surface of the β barrel, similar to binding strategies seen in other types of OB folds [72]. As with the sliding clamps, differences in oligomeric state among different SSB species can be compensated for by repetition of homologous domains within the proteins. For example, *E. coli* SSB forms a homotetramer containing one OB fold per protomer [64], whereas human RPA is a heterotrimer composed of one large subunit with two OB folds and two smaller subunits, each bearing a single OB fold [65,66]. Both proteins, therefore, have four OB folds in their final active form, but arrive at that number in different ways. The importance of this fourfold arrangement for replicative SSBs is unclear.

Terminators

At the end of replication, ~180° away from the origin in a circular bacterial chromosome, replication forks are halted by replication termination proteins (called RTP in *Bacillus subtilis* or 'tus' in *E. coli*) [73]. These proteins bind to specific DNA sequences ('ter' sites) and inhibit replication-fork helicase activity in a polar fashion. For a given tus–ter or RTP–ter complex, therefore, only a single fork progressing from one direction will be terminated while a replication fork approaching from the other side is allowed to pass [74]. This polarity ensures that a clockwise-moving replication fork starting at ori (0°) passes through one tus–ter site at 179.9° but terminates at another, oppositely facing tus–ter at 180.1°. Meanwhile, the counterclockwise fork passes through the tus–ter at 180.1° but terminates at the 179.9° complex. In this way, the circular bacterial chromosome can be fully copied by bidirectional replication.

Two terminator proteins have been solved to high resolution, *B. subtillus* RPT [75] and a ter-bound form of *E. coli* tus [76] (Figure 4b). The structures are significantly different: RTP has a helix–turn–helix (HTH) domain that has been shown biochemically to be its DNA recognition surface, whereas tus binds ter primarily through β sheet–major groove interactions. The tus–ter complex is asymmetric, as would be expected given the polarity of its biochemical activity. Mechanisms to explain terminator activities invoke an asymmetric steric block that inhibits DNA helicases in one direction but allows translocation from the other direction. Because terminators also inhibit nonreplicative processes such as transcription in a polar fashion, it appears unlikely that a specific complex is formed between the replicative helicase and the termination apparatus [73].

Future structural studies in replication systems

Structural studies have helped reveal the molecular details of replication in several ways. First, determination of the structures of several examples of functionally analogous proteins reveals striking structural similarities between phage and bacteria, as well as with eukaryotes, often despite a lack of clear-cut sequence homology. This suggests that the basic replication reaction is ancient and conserved across the biosphere. Second, structures have helped answer several mechanistic questions about the reactions of replication, including the nature of processivity enhancement and the structural basis of dNTP addition to growing DNA chains. Finally, structural data have offered a visual basis for understanding replication that might not have arisen solely from biochemical or genetic studies.

What lies ahead? Figure 1 indicates that a majority of individual replication proteins have been determined to high-resolution. What is not known, however, is how these components fit together into the biologically relevant complexes that coordinate their activities. Structure determination of complexes is difficult, and is made even more difficult for replication proteins by the dynamic nature of these complexes. One approach to solving this problem is to crystallize minimal interactive regions from multiple proteins of known structure and to extrapolate these interactions to a model of the larger protein–protein complex. A second approach is to find stabilized homogeneous samples of replication complexes, either by covalently linking together proteins or perhaps by exploring replication complexes from thermophilic organisms. An example of the first approach is exemplified in a recent experiment that has led to a convincing polymerase–sliding clamp complex model from the bacteriophage RB69 [51]. In this work, Shamoo and Steitz [52] solved the structures of the sliding clamp and DNA-bound form of the replicative DNA polymerase from RB69. They then determined the structure of a complex between the clamp and the region of the DNA polymerase that links the two proteins, thus allowing them to build a model of the clamp–polymerase–DNA complex (Figure 3d). Studies employing both approaches in determining how multiprotein complexes assemble and coordinate their functions will dictate the future of structural research on replication.

Acknowledgements

We thank Dale Wigley for communicating results prior to publication and Peter Garber for careful reading of the manuscript. This review was supported by the G. Harold & Leila Y. Mathers Charitable Foundation (J.M.B.) and The Jane Coffin Childs Memorial Fund for Medical Research (J.L.K.).

References

1. Gross, J.D. (1971). DNA replication in bacteria. *Curr. Top. Microbiol. Immunol.* **74**, 39-74.
2. Kornberg, A. & Baker, T.A. (1992). *DNA Replication.* W.H. Freeman and Co., New York.
3. Marians, K.J. (1992). Prokaryotic DNA replication. *Annu. Rev. Biochem.* **61**, 673-719.
4. Hickey, R.J. & Malkas, L.H. (1997). Mammalian cell DNA replication. *Crit. Rev. Eukaryot. Gene Expr.* **7**, 125-157.
5. Boehmer, P.E. & Lehman, I.R. (1997). Herpes simplex virus DNA replication. *Annu. Rev. Biochem.* **66**, 347-84.
6. Waga S. & Stillman, B. (1998). The DNA replication fork in eukaryotic cells. *Annu. Rev. Biochem.* **67**, 721-751.
7. Baker, T.A. & Bell, S.P. (1998). Polymerases and the replisome: machines within machines. *Cell* **92**, 295-305.

8. Fuller, R.S., Funnell, B.E. & Kornberg, A. (1984). The dnaA protein complex with the *E. coli* chromosomal replication origin (oriC) and other DNA sites. *Cell* **38**, 889-900.
9. Bramhill, D. & Kornberg, A. (1988). Duplex opening by dnaA protein at novel sequences in initiation of replication at the origin of the *E. coli* chromosome. *Cell* **52**, 743-755.
10. Kaguni, J.M., Fuller, R.S. & Kornberg, A. (1982). Enzymatic replication of *E. coli* chromosomal origin is bidirectional. *Nature* **296**, 623-627.
11. Baker, T.A., Sekimizu, K., Funnell, B.E. & Kornberg, A. (1986). Extensive unwinding of the plasmid template during staged enzymatic initiation of DNA replication from the origin of the *Escherichia coli* chromosome. *Cell* **45**, 53-64.
12. Arai, K. & Kornberg, A. (1981). Unique primed start of phage phi X174 DNA replication and mobility of the primosome in a direction opposite chain synthesis. *Proc. Natl Acad. Sci. USA* **78**, 69-73.
13. LeBowitz, J.H. & McMacken, R. (1986). The *Escherichia coli* dnaB replication protein is a DNA helicase. *J. Biol. Chem.* **261**, 4738-4748.
14. Nusslein, V., Otto, B., Bonhoeffer, F. & Schaller, H. (1971). Function of DNA polymerase III in DNA replication. *Nat. New Biol.* **234**, 285-286.
15. Gefter, M.L., Hirota, Y., Kornberg, T., Wechsler, J.A. & Barnoux, C. (1971). Analysis of DNA polymerases II and III in mutants of *Escherichia coli* thermosensitive for DNA synthesis. *Proc. Natl Acad. Sci. USA* **68**, 3150-3153.
16. Wickner, S. (1976). Mechanism of DNA elongation catalyzed by *Escherichia coli* DNA polymerase III, dnaZ protein, and DNA elongation factors I and III. *Proc. Natl Acad. Sci. USA* **73**, 3511-3515.
17. O'Donnell, M.E. (1987). Accessory proteins bind a primed template and mediate rapid cycling of DNA polymerase III holoenzyme from *Escherichia coli*. *J. Biol. Chem.* **262**, 16558-16565.
18. Kelman, Z. & O'Donnell, M. (1995). DNA polymerase III holoenzyme: structure and function of a chromosomal replicating machine. *Annu. Rev. Biochem.* **64**, 171-200.
19. de Massy, B., Bejar, S., Louarn, J., Louarn, J.M. & Bouche, J.P. (1987). Inhibition of replication forks exiting the terminus region of the *Escherichia coli* chromosome occurs at two loci separated by 5 min. *Proc. Natl Acad. Sci. USA* **84**, 1759-1763.
20. Khatri, G.S., MacAllister, T., Sista, P.R. & Bastia, D. (1989). The replication terminator protein of *E. coli* is a DNA sequence-specific contra-helicase. *Cell* **59**, 667-674.
21. Hegde, R.S., Grossman, S.R., Laimins, L.A. & Sigler, P.B. (1992). Crystal structure at 1.7 Å of the bovine papillomavirus-1 E2 DNA-binding domain bound to its DNA target. *Nature* **359**, 505-512.
22. Bochkarev, A., Barwell, J.A., Pfuetzner, R.A., Bochkareva, E., Frappier, L. & Edwards, A.M. (1996). Crystal structure of the DNA-binding domain of the Epstein–Barr virus origin-binding protein, EBNA1, bound to DNA. *Cell* **84**, 791-800.
23. Luo, X., Sanford, D.G., Bullock, P.A. & Bachovchin, W.W. (1996). Solution structure of the origin DNA-binding domain of SV40 T-antigen. *Nat. Struct. Biol.* **3**, 1034-1039.
24. Komori, H., Matsunaga, F., Higuchi, Y., Ishiai, M., Wada, C. & Miki, K. (1999). Crystal structure of a prokaryotic replication initiator protein bound to DNA at 2.6 Å resolution. *EMBO J.* **18**, 4597-4607.
25. Edwards, A.M., Bochkarev, A. & Frappier, L. (1998). Origin DNA-binding proteins. *Curr. Opin. Struct. Biol.* **8**, 49-53.
26. Lohman, T.M. & Bjornson, K.P. (1996). Mechanisms of helicase-catalyzed DNA unwinding. *Annu. Rev. Biochem.* **65**, 169-214.
27. Gorbalenya, A.E. & Koonin, E.V. (1993). Helicases: amino acid sequence comparisons and structure-function relationships. *Curr. Opin. Struct. Biol.* **3**, 419-429.
28. Subramanya, H,S., Bird, L.E., Brannigan, J.A. & Wigley, D.B. (1996). Crystal structure of a DExx box DNA helicase. *Nature* **384**, 379-383.
29. Korolev, S., Hsieh, J., Gauss, G.H., Lohman, T.M. & Waksman, G. (1997). Major domain swiveling revealed by the crystal structures of complexes of *E. coli* Rep helicase bound to single-stranded DNA and ADP. *Cell* **90**, 635-647.
30. Kim, J.L., *et al.*, & Thomson, J.A. (1996). Crystal structure of the hepatitis C virus NS3 protease domain complexed with a synthetic NS4A cofactor peptide. *Cell* **87**, 343-355.
31. Yao, N., *et al.,* & Weber, P.C. (1997). Structure of the hepatitis C virus RNA helicase domain. *Nat. Struct. Biol.* **4**, 463-467.
32. Cho, H.S., *et al.,* & Oh, B.H. (1998). Crystal structure of RNA helicase from genotype 1b hepatitis C virus. A feasible mechanism of unwinding duplex RNA. *J. Biol. Chem.* **273**, 15045-15052.
33. Sawaya, M.R., Guo, S., Tabor, S., Richardson, C.C. & Ellenberger, T. (1999). Crystal structure of the helicase domain from the replicative helicase–primase of bacteriophage T7. *Cell* **99**, 167-177.
34. Bernstein, J.A. & Richardson, C.C. (1989). Characterization of the helicase and primase activities of the 63 kDa component of the bacteriophage T7 gene 4 protein. *J. Biol. Chem.* **264**, 13066-13073.
35. Ilyina, T.V., Gorbalenya, A.E. & Koonin, E.V. (1992). Organization and evolution of bacterial and bacteriophage primase-helicase systems. *J. Mol. Evol.* **34**, 351-357.
36. Egelman, E.H., Yu, X., Wild, R., Hingorani, M.M. & Patel, S.S. (1995). Bacteriophage T7 helicase/primase proteins form rings around single-stranded DNA that suggest a general structure for hexameric helicases. *Proc. Natl Acad. Sci. USA* **92**, 3869-3873.
37. Story, R.M., Weber, I.T. & Steitz, T.A. (1992). The structure of the *E. coli* recA protein monomer and polymer. *Nature* **355**, 318-325.
38. Yu, X. & Egelman, E.H. (1997). The RecA hexamer is a structural homologue of ring helicases. *Nat. Struct. Biol.* **4**, 101-104.
39. Bennett, M.J., Schlunegger, M.P. & Eisenberg, D. (1995). 3D domain swapping: a mechanism for oligomer assembly. *Protein Sci.* **4**, 2455-2468.
40. Walker, J.E., Saraste, M., Runswick, M.J. & Gay, N.J. (1982). Distantly related sequences in the alpha- and beta-subunits of ATP synthase, myosin, kinases and other ATP-requiring enzymes and a common nucleotide binding fold. *EMBO J.* **1**, 945-951.
41. Bird, L.E., Subramanya, H.S. & Wigley, D.B. (1998). Helicases: a unifying structural theme? *Curr. Opin. Struct. Biol.* **8**, 14-18.
42. Abrahams, J.P., Leslie, A.G., Lutter, R. & Walker, J.E. (1994). Structure at 2.8 Å resolution of F1-ATPase from bovine heart mitochondria. *Nature* **370**, 621-628.
43. Ollis, D.L., Brick, P., Hamlin, R., Xuong, N.G. & Steitz, T.A. (1985). Structure of large fragment of *Escherichia coli* DNA polymerase I complexed with dTMP. *Nature* **313**, 762-766.
44. Joyce, C.M. & Steitz, T.A. (1994). Function and structure relationships in DNA polymerases. *Annu. Rev. Biochem.* **63**, 777-822.
45. Brautigam, C.A. & Steitz, T.A. (1998). Structural and functional insights provided by crystal structures of DNA polymerases and their substrate complexes. *Curr. Opin. Struct. Biol.* **8**, 54-63.
46. Doublie, S. & Ellenberger, T. (1998). The mechanism of action of T7 DNA polymerase. *Curr. Opin. Struct. Biol.* **8**, 704-712.
47. Steitz, T.A. (1999). DNA polymerases: structural diversity and common mechanisms. *J. Biol. Chem.* **274**, 17395-17398.
48. Jager, J. & Pata, J.D. (1999). Getting a grip: polymerases and their substrate complexes. *Curr. Opin. Struct. Biol.* **9**, 21-28.
49. Doublie, S., Tabor, S., Long, A.M., Richardson, C.C. & Ellenberger, T. (1998). Crystal structure of a bacteriophage T7 DNA replication complex at 2.2 Å resolution. *Nature* **391**, 251-258.
50. Wang, J., Sattar, A.K., Wang, C.C., Karam, J.D., Konigsberg, W.H. & Steitz, T.A. (1997). Crystal structure of a pol alpha family replication DNA polymerase from bacteriophage RB69. *Cell* **89**, 1087-1099.
51. Shamoo, Y. & Steitz, T.A. (1999). Building a replisome from interacting pieces: sliding clamp complexed to a peptide from DNA polymerase and a polymerase editing complex. *Cell* **99**, 155-166.
52. Hopfner, K.P., *et al.*, & Angerer B. (1999). Crystal structure of a thermostable type B DNA polymerase from *Thermococcus gorgonarius*. *Proc. Natl Acad. Sci. USA* **96**, 3600-3605.
53. Kuriyan, J. & O'Donnell, M. (1993). Sliding clamps of DNA polymerases. *J. Mol. Biol.* **234**, 915-925.
54. Kelman, Z. (1997). PCNA: structure, functions and interactions. *Oncogene* **14**, 629-640.
55. Kong, X.P., Onrust, R., O'Donnell, M. & Kuriyan, J. (1992). Three-dimensional structure of the beta subunit of *E. coli* DNA polymerase III holoenzyme: a sliding DNA clamp. *Cell* **69**, 425-437.
56. Krishna, T.S., Kong, X.P., Gary, S., Burgers, P.M. & Kuriyan, J. (1994). Crystal structure of the eukaryotic DNA polymerase processivity factor PCNA. *Cell* **79**, 1233-1243.
57. Gulbis, J.M., Kelman, Z., Hurwitz, J., O'Donnell, M. & Kuriyan, J. (1996). Structure of the C-terminal region of p21(WAF1/CIP1) complexed with human PCNA. *Cell* **87**, 297-306.
58. Modrich, P. & Richardson, C.C. (1975). Bacteriophage T7 deoxyribonucleic acid replication *in vitro*. A protein of *Escherichia coli* required for bacteriophage T7 DNA polymerase activity. *J. Biol. Chem.* **250**, 5508-5514.
59. Yuzhakov, A., Kelman, Z. & O'Donnell, M. (1999). Trading places on DNA – a three-point switch underlies primer handoff from primase to the replicative DNA polymerase. *Cell* **96**, 153-163.
60. Guenther, B., Onrust, R., Sali, A., O'Donnell, M. & Kuriyan, J. (1997). Crystal structure of the δ′ subunit of the clamp-loader complex of *E. coli* DNA polymerase III. *Cell* **91**, 335-345.

61. Wold, M.S. (1997). Replication protein A: a heterotrimeric, single-stranded DNA-binding protein required for eukaryotic DNA metabolism. *Annu. Rev. Biochem.* **66**, 61-92.
62. Iftode, C., Daniely, Y. & Borowiec, J.A. (1999). Replication protein A (RPA): the eukaryotic SSB. *Crit. Rev. Biochem. Mol. Biol.* **34**, 141-180.
63. Shamoo, Y., Friedman, A.M., Parsons, M.R., Konigsberg, W.H. & Steitz, T.A. (1995). Crystal structure of a replication fork single-stranded DNA binding protein (T4 gp32) complexed to DNA. *Nature* **376**, 362-366.
64. Raghunathan, S., Ricard, C.S., Lohman, T.M. & Waksman, G. (1997). Crystal structure of the homo-tetrameric DNA binding domain of *Escherichia coli* single-stranded DNA-binding protein determined by multiwavelength X-ray diffraction on the selenomethionyl protein at 2.9 Å resolution. *Proc. Natl Acad. Sci. USA* **94**, 6652-6657.
65. Bochkarev, A., Pfuetzner, R.A., Edwards, A.M. & Frappier, L. (1997). Structure of the single-stranded-DNA-binding domain of replication protein A bound to DNA. *Nature* **385**, 176-181.
66. Bochkarev, A., Bochkareva, E., Frappier, L. & Edwards, A.M. (1999). The crystal structure of the complex of replication protein A subunits RPA32 and RPA14 reveals a mechanism for single-stranded DNA binding. *EMBO J.* **18**, 4498-4504.
67. Yang, C., Curth, U., Urbanke, C. & Kang, C. (1997). Crystal structure of human mitochondrial single-stranded DNA binding protein at 2.4 Å resolution. *Nat. Struct. Biol.* **4**, 153-157.
68. Folkers, P.J., Nilges, M., Folmer, R.H., Konings, R.N. & Hilbers, C.W. (1994). The solution structure of the Tyr41→His mutant of the single-stranded DNA binding protein encoded by gene V of the filamentous bacteriophage M13. *J. Mol. Biol.* **236**, 229-246.
69. Skinner, M.M., *et al.*, & Terwilliger, T.C. (1994). Structure of the gene V protein of bacteriophage f1 determined by multiwavelength X-ray diffraction on the selenomethionyl protein. *Proc. Natl Acad. Sci. USA* **91**, 2071-2075.
70. Folmer, R.H., Nilges, M., Konings, R.N. & Hilbers, C.W. (1995). Solution structure of the single-stranded DNA binding protein of the filamentous Pseudomonas phage Pf3: similarity to other proteins binding to single-stranded nucleic acids. *EMBO J.* **14**, 4132-4142.
71. Tucker, P.A., Tsernoglou, D., Tucker, A.D., Coenjaerts, F.E., Leenders, H. & van der Vliet, P.C. (1994). Crystal structure of the adenovirus DNA binding protein reveals a hook-on model for cooperative DNA binding. *EMBO J.* **13**, 2994-3002.
72. Murzin, A.G. (1993). OB(oligonucleotide/oligosaccharide binding)-fold: common structural and functional solution for non-homologous sequences. *EMBO J.* **12**, 861-867.
73. Bussiere, D.E. & Bastia, D. (1999). Termination of DNA replication of bacterial and plasmid chromosomes. *Mol. Microbiol.* **31**, 1611-1618.
74. Lee, E.H., Kornberg, A., Hidaka, M., Kobayashi, T. & Horiuchi, T. (1989). *Escherichia coli* replication termination protein impedes the action of helicases. *Proc. Natl Acad. Sci. USA* **86**, 9104-9108.
75. Bussiere, D.E., Bastia, D. & White, S.W. (1995). Crystal structure of the replication terminator protein from *B. subtilis* at 2.6 Å. *Cell* **80**, 651-660.
76. Kamada, K., Horiuchi, T., Ohsumi, K., Shimamoto, N., Morikawa, K. (1996). Structure of a replication-terminator protein complexed with DNA. *Nature* **383**, 598-603.
77. Pan, H. & Wigley, D.B (2000). Structure of the zinc-binding domain of *Bacillus stearothermophilus* DNA primase. *Structure* **8**, 231-239.
78. Carson, M. (1991). Ribbons 2.0. *J. Appl. Crystallogr.* **24**, 958-961.