DNA replication: enzymology and mechanisms

Zvi Kelman and Mike O'Donnell

Cornell University Medical College, New York, USA

Research into the enzymology of DNA replication has seen a multitude of highly significant advances during the past year, in both prokaryotic and eukaryotic systems. The scope of this article is limited to chromosomal replicases and origins of initiation. The multiprotein chromosomal replicases of prokaryotes and eukaryotes appear to be strikingly similar in structure and function, although future work may reveal their differences. Recent developments, elaborating the activation of origins in several systems, have begun to uncover mechanisms of regulation. The enzymology of eukaryotic origins has, until now, been limited to viral systems, but over the past few years, enzymology has caught a grip on the cellular origins of yeast.

Introduction

The structure of duplex DNA is so simple and elegant that one may have thought it would be simple to duplicate. However, since the isolation of DNA polymerase I first delivered replication mechanisms into the hands of enzymologists, the process has been shown to be far from simple. Over 20 proteins are utilized in Escherichia coli and probably more in higher organisms [1]. In fact, the individual functions of many of these proteins are still unknown and even more proteins remain to be identified. In overview, the process begins at a specific sequence called an origin upon which proteins bind and locally unwind the duplex, allowing invasion of a helicase. The helicase couples ATP hydrolysis to melt the duplex, and then the single-stranded DNA (ssDNA) is coated with ssDNA-binding protein (SSB). This SSB-ssDNA nucleofilament has no secondary structure and serves as the most efficient template for chromosomal replicases, multiprotein machineries characterized by their rapid and highly processive DNA synthesis. One strand of the chromosome is synthesized continuously, but because of the antiparallel structure of the duplex, the other strand is copied discontinuously as a series of fragments (Okazaki fragments). This discontinuous mode of replication necessitates frequent reinitiation of DNA chains, which are primed by short RNA primers synthesized by a primase. Most of the above proteins are thought to work together in one large replisome assembly for coordinated synthesis of both strands of the chromosome.

In this review, we will focus on recent developments in the study of origin activation and the mechanism of action of chromosomal replicases. These areas encompass only a subset of the recent advances in the enzymology of DNA replication.

Replication origins

The first (and still the only) cellular origin to have been activated in vitro was oriC, the origin of the E. coli chromosome [1]. Since that breakthrough, several phage and viral origins have been activated in vitro. Origin activation requires multiple copies of an origin-binding protein and usually additional origin accessory proteins, some of which have secondary functions as transcriptional activators. As proposed by Hatch Echols [2], the requirement for multiple proteins in site-specific processes, such as activation of origins and promoters, is imposed by the need for a high fidelity of action on specific sequences embedded in the bulk of chromosome DNA. These multiprotein-DNA complexes are termed specialized nucleoprotein structures (snups) [2]. Origin initiation events in E. coli, phage λ and simian virus 40 (SV40) are similar and appear to apply also to the more recently characterized systems of phage P1, bovine papilloma virus (BPV), herpes simplex virus 1 (HSV-1), and yeast (see Fig. 1). In most systems, the origin-binding protein is pre-
pared for origin activation and, for this review, we define this as stage I. Multiple copies of the origin-binding protein then assemble onto the origin with the accessory proteins to form the snup (stage II). This untwists a section of AT-rich DNA to form an 'open complex' (stage III), providing a point of entry for the helicase (stage IV). After helicase entry, the chromosome can be extensively unwound and replication forks can be assembled. Origin activation systems differ in the details of these steps and also in their regulation. In the following section, we will focus on these differences, with particular attention to recent developments.

Prokaryotic origins
Recent studies show that the activation of oriC is regulated at each stage (Fig. 1). In stage I, the DnaA origin-binding protein is inactive when bound to acidic phospholipids, but this repression can be relieved, and DnaA prepared for origin binding, by the DnaK heat shock protein or by phospholipase treatment [3]. In stage II, the origin snup is nucleated at four DnaA-binding sites in oriC, upon which 20 or more DnaA monomers assemble, along with several origin accessory proteins [1]. These accessory proteins include HU and IHF, which bend DNA, and they may act by helping to wrap DNA around the DnaA subunits [4°, 5]. Other accessory proteins include Fis and the recently identified Rob protein [6°]. They are not essential for the activation of the origin, and how they may modulate the process is not yet clear. In stage III, the origin snup unwinds DNA within three AT-rich 13-mer repeats to form the open complex and this reaction requires a supercoiled template [1]. Formation of the open complex is highly regulated [1]. Negative regulation is achieved by the inhibitor of cellular initiation, IciA, which competes with DnaA for interaction with the 13-mers and prevents the formation of the open complex [7∗, 8∗]. ATP also regulates the open complex formation. DnaA can bind either ATP or ADP tightly, but only the ATP form is active in unwinding the 13-mers [1]. The DnaA protein becomes inactive upon slowly hydrolyzing the ATP to ADP, and the ADP remains bound, thus preventing subsequent reinitiations. The DnaA protein can be readily reactivated by acidic phospholipids, which catalyze the exchange of ATP for the bound ADP [1, 9∗], but only when DnaA is bound to oriC; this may underlie the early observations of replicon attachment to the membrane [10]. RNA polymerase provides yet another level of regulation in determining the ability of R loops (RNA paired with one strand of the DNA duplex) to promote the open complex, and although the transcript need not enter oriC, it must be close [11]. In stage IV, the DnaB helicase enters the open complex, but its access absolutely requires the DnaC protein, a molecular matchmaker that delivers the hexameric DnaB into the origin in an ATP-dependent reaction, after which DnaC departs from the DNA [1]. Origin unwinding is bidirectional; therefore, two hexamers of DnaB must

---

**Fig. 1.** Regulation at different stages of origin activation in a number of systems. Origin activation has been subdivided into four stages: preparation of the origin-binding protein, assembly of the origin snup, local unwinding of origin DNA (the open complex), and entry of the helicase. Molecules that act on each stage are divided into positive (+) and negative (−) effectors and are shown for E. coli, phage λ, phage P1, yeast, SV40, BPV and HSV-1. A blank entry indicates that effectors of that step have not (yet) been identified. At the far right are listed the replication proteins needed to advance the replication fork.
be delivered to the open complex. Negative regulation is achieved by the presence of too much DnaC protein, which binds to and inhibits the DnaB helicase activity, possibly by preventing its translocation along DNA; therefore, a ‘fine balance’ of DnaC and DnaB is essential to the productivity of this stage [12].

Phage λ uses mainly host replication proteins for origin activation, with the exception of the phage-encoded O and P proteins, analogs of DnaA and DnaC [1]. The λ origin snup is composed of four dimers of λ O protein, which does not require pre-activation to assemble onto the origin. Formation of the open complex at three AT-rich 11-mer repeats requires the DNA to be supercoiled, but does not require ATP [13,14]. In stage IV, the λ P protein binds DnaB and delivers it to the origin, but the P protein remains tightly associated with DnaB and prevents its helicase activity [15].

In this regard, the P protein acts as a negative regulator, like excess DnaC, to prevent translocation of DnaB. Although ATP is not required for P protein to bring DnaB to the origin, it is required to break the interaction of P protein with DnaB, thus freeing the helicase for replication; this step is mediated by heat shock proteins DnaJ, DnaK and GrpE [16,17]. Transcription from the rightward promoter is needed for λ origin activation in vivo, and in vitro studies show that transcription disrupts HU-mediated negative repression (presumably of stage II or III) [18].

Recent studies on the activation of the lysogenic origin of phage P1 show that the phage-encoded RepA origin-binding protein is a native dimer, but the monomer strongly associates with the origin. In stage I, RepA is prepared for origin binding by the Dnaj, Dnak and GrpE heat shock proteins, which couple ATP hydrolysis to monomerize RepA [19**]. The origin snup comprises RepA bound at five sites and the host DnaA protein at five other sites. At least two of the DnaA-binding sites are essential for origin function, and in vitro studies show that the ADP form of DnaA protein is capable of activating the origin in combination with RepA (reviewed in [20]). Although stages III and IV have not been studied in detail, it seems likely that the open complex forms at five slightly AT-rich 7-mer repeats and then DnaC mediates delivery of DnaB to the open complex.

Eukaryotic viral origins

In the SV40 system (reviewed in [21]), the stage I preparation of the virally encoded T-antigen for origin binding is regulated in both positive and negative fashion. T-antigen must be phosphorylated at Thr124 for productive origin binding, and this modification can be performed by the cdc2 kinase [22]. Phosphorylation at other sites can inhibit T-antigen, and protein phosphatase 2A (PP2A) has been shown to activate T-antigen by removing phosphates from specific serine residues implicated in DNA-binding activity [23].

The p53 tumor suppressor protein is a negative regulator of stage I; it binds to T-antigen preventing its association with the origin [24]. Stage II, formation of the origin snup, is facilitated by ATP binding, which effects the assembly of two hexamers of T-antigen onto the origin (reviewed in [25]). The two hexamers are thought to encircle the origin DNA in SV40, because if the hexamers are pre-assembled in solution they do not bind the origin [26**]. Moreover, in electron micrographs the double hexamer of T-antigen on the origin does not appear to wrap DNA around it as with other origin snups, but rather the DNA appears to follow a straight path through the protein [27]. The transcriptional activator Sp1 acts as an origin accessory protein that perturbs the local histone distribution to make the origin accessible for initiation proteins [28]. However, the identity of the transcriptional activator is not important, as Sp1 can be exchanged for other activators [28], but the DNA-binding domain must be accompanied by an activation domain for origin function [29**].

Formation of the open complex in stage III is coincident with T-antigen binding to the origin, which results in the melting of 8 bp in the early palindrome and unwrapping of an AT tract [25]. These events require binding, but not hydrolysis, of ATP and do not require a supercoiled template [25]. In SV40, unlike the other systems discussed above, the stage IV helicase invasion is unique in that T-antigen itself is the helicase and thus encompasses the functions of the E. coli DnaA, B and C proteins. Stage IV is still regulated, however; for example, it has been shown recently that PP2A-treated T-antigen provides a cooperative interaction between the two hexamers, which facilitates stage IV [30**]. In addition, the human SSB, replication protein A (RP-A), consists of three non-identical subunits (p70, p32 and p14) and has been shown to be a positive effector of stage IV, as RP-A is needed for T-antigen to unwind the bulk of DNA [25]. A second point at which p53 may regulate replication has been identified recently in an interaction of p53 with the p70 subunit of RP-A (the ssDNA-binding subunit) that inactivates its ability to bind ssDNA [31**]. Also, the p32 subunit of RP-A is phosphorylated in a cell cycle dependent manner in the G1 to S transition by members of the cyclin cdc2 kinase family, and addition of this kinase to the SV40 replication system stimulates DNA synthesis [32**]. Phosphorylation of p32 is also performed by the DNA-activated protein kinase (GS Brush, CW Anderson, TJ Kelly, abstract 191, Eukaryotic DNA Replication Meeting, Cold Spring Harbor, September 1993).

The BPV snup is composed of two virally encoded proteins, E1 and E2. E1 is the origin-binding protein and a helicase (like T-antigen) [33**,34**] and E2 is an origin accessory protein (and transcriptional activator) that binds specific sequences and helps in the delivery of E1 to the origin via direct protein contacts [35*,36]. In this regard, E2 fulfills a stage IV function analogous to that of E. coli DnaC in delivery of the helicase into the origin. Regulatory aspects, exact functions of ATP and identification of an open complex remain for future studies, but the in vitro replication system and availability of pure E1 and E2 should yield this information in the near future.
In the HSV-1 system, the identification of all seven of the essential virally encoded replication proteins and the cloning and production of each of them in quantity has enabled a number of illuminating studies, although replication in vitro of a plasmid containing the origin has yet to be achieved. The origin snap consists of at least two dimers of the UL9 protein, which produces the open complex in an AT-rich region even in the absence of at least two dimers of the UL9 protein, which produces the open complex in an AT-rich region even in the absence of ATP [37]. UL9 has helicase activity, but it does not act catalytically like the other helicases discussed thus far. Instead, it must bind DNA stoichiometrically for unwinding and requires a stretch of ssDNA for duplex-unwinding activity [38, 39]. The UL9 helicase is stimulated by the virally encoded SSB (called ICP8) and it seems likely they function together at the origin to enlarge the open complex for the future replication fork.

Yeast chromosomal origin

In the past two years, rapid and exciting advances have been made in the study of the biochemistry of yeast cellular replication. Yeast origins are known as autonomous replication sequences (ARS) for their ability to confer autonomous replication on plasmids. Recent studies [40**] of the ARS1 origin show that it contains four sequence elements: the A element, which contains an 11 bp ARS core consensus sequence (ACS) common to all ARS elements, the B3 element, which is the binding site of the ABF1 transcription factor, and the B1 and B2 elements, which may interact with (as yet) unidentified origin accessory proteins. A large origin recognition complex (ORC), which binds to ARS1 at the A site, has now been purified from Saccharomyces cerevisiae [41**] and genomic footprinting experiments indicate the presence of the ORC on ARS1 in vivo [42**]. The ORC contains six subunits and requires ATP to bind the ARS. The exact function of the ORC is not certain, but a number of observations confirm that it plays a central role in replication. First, the ORC does not bind to single-site mutant forms of ACS that lack ARS activity [41**]. Second, mutations affecting the 70 kDa subunit of the ORC, identified because they produce a defect in function of the HMR silencer, show that the gene is essential for cell viability, and a conditional lethal allele is defective in chromosome replication [43*]. Third, each of the six genes encoding the ORC subunits are essential to yeast (SP Bell, R Kobayashi, B Stillman, abstract 15, Eukaryotic DNA Replication Meeting, Cold Spring Harbor, September 1993).

Recent developments in the study of the genetics of yeast replication have given researchers clues as to the roles of important replication proteins. Several genes of S. cerevisiae have been isolated, the products of which are needed for minichromosome maintenance (MCM) of ARS-containing plasmids. Three of these, MCM2, MCM3 and MCM5, are homologous in sequence and each is essential for cell viability [44*]. The intracellular location of MCM2, 3 and 5 is cell cycle regulated; they are moved into the nucleus upon completion of S phase and then moved into the cytoplasm at the start of S phase [44*, 45**]. It is hypothesized that in the nucleus the MCM proteins activate the ARS and are then transferred to the cytoplasm to prevent reinitiation [44*, 45**]; however, whether they act on origins directly and in a positive fashion must await further studies. In another exciting development, the cdcl8 gene of Schizosaccharomyces pombe has been shown to suppress mutations of cdcl10, a transcriptional activator that is needed for entry into S phase [46**]. The cdcl18 gene may be a major target of cdcl10 and, consistent with this, cdcl18 mutants fail to enter S phase [46**]. Further, cdcl18 may play a role in checkpoint control, as cdcl18 mutants are unable to prevent mitosis and rapid cell division even though the chromosomes are not fully duplicated. The cdcl18 gene sequence reveals a nucleotide-binding site, but whether the encoded protein plays a direct role (e.g. helicase) or an indirect role in replication must await future isolation and characterization of the gene product.

Events after helicase entry

Once the helicase has entered the chromosome, it presumably nucleates assembly of replication forks containing the primase and two replicative polymerases, one for each strand of DNA [1]. The leading and lagging strand polymerases in eukaryotes appear to be different polymerases (polymerases δ and ε), whereas prokaryotes use two copies of the same polymerase. The mechanism of replicase function is the second topic of this essay.

Chromosomal replicases

Replicases of E. coli, phage T4, yeast and humans

Organisms that span the evolutionary spectrum have been shown to possess replicases that appear to be similar in function and also in their actual structure. These are the E. coli DNA polymerase III holoenzyme (Pol III holoenzyme), the phage T4 replicase, and the polymerase δ (pol δ) of both yeast and humans. In each of these systems, the replicase encompasses several proteins that use ATP to initiate rapid and highly processive DNA synthesis. It can be thought of as having three components (see Table 1): the catalytic component, which contains the DNA polymerase and proofreading 3'→5' exonuclease activities, and the two categories of polymerase accessory proteins, a complex of accessory proteins and a single subunit processivity factor. The accessory proteins are needed to confer high processivity onto the catalytic polymerase. Over the past two years, a deeper understanding of the action of these accessory proteins has been developed and appears to apply generally to all four systems. Studies have shown that the single subunit accessory protein is a DNA-sliding clamp, which tethers the polymerase to DNA for high processivity.
The accessory protein complex functions as a 'clamp loader' that couples ATP to assemble the clamp protein on DNA in a two-step assembly process (see Fig. 2). The clamp loader recognizes the primed template and couples ATP to assemble the sliding clamp onto DNA to form a pre-initiation complex. This is followed by association of the polymerase to form the initiation complex (reviewed in [47]).

The best studied system, at least for this assembly reaction, is the Pol III holoenzyme of *E. coli*. The processivity factor of Pol III, the β subunit, is a dimer that freely slides along duplex DNA and is topologically linked to the DNA, in as much as it binds tightly to a nicked circular plasmid, but upon linearization freely slides off over the ends [48]. These results have been explained by the hypothesis that β encircles DNA like a doughnut [48]. The crystal structure of the β dimer shows that it is indeed in the shape of a ring capable of completely surrounding duplex DNA and reveals the unexpected feature that β, although only a dimer, has a sixfold appearance (Fig. 3) [49**]. This symmetry is the result of the three globular domains that comprise each monomer and the polypeptide chain backbone structures of these domains are nearly superimposable. The β dimer cannot assemble onto DNA independently, but in fact requires the five-protein γ complex clamp loader, a molecular matchmaker that hydrolyzes ATP to assemble the β dimer around DNA. The catalytic component of the Pol III holoenzyme, termed Pol III core, assembles with the β ring, which tethers the Pol III core to the DNA and continues to slide with it for rapid and highly processive synthesis. The γ complex acts catalytically to assemble multiple β clamp pre-initiation complexes on different primed templates. The

Table 1. Three part structure of chromosomal replicases of *E. coli*, phage T4, yeast and human.

<table>
<thead>
<tr>
<th>Component</th>
<th>E. coli</th>
<th>Phage T4</th>
<th>Eukaryotic</th>
</tr>
</thead>
<tbody>
<tr>
<td>I. Polymerase, 3′→5′ exonuclease</td>
<td>Pol III core</td>
<td>gp43</td>
<td>pol δ</td>
</tr>
<tr>
<td></td>
<td>α (pol)</td>
<td>p125</td>
<td></td>
</tr>
<tr>
<td></td>
<td>ε (exo)</td>
<td>p50</td>
<td></td>
</tr>
<tr>
<td></td>
<td>θ</td>
<td>pol ε</td>
<td></td>
</tr>
<tr>
<td>II. Accessory complex</td>
<td>γ complex</td>
<td>gp44–gp62 complex</td>
<td>RF-C (Activator-1)</td>
</tr>
<tr>
<td>(clamp loader)</td>
<td>DNA-dependent</td>
<td>γ</td>
<td>p128</td>
</tr>
<tr>
<td></td>
<td>ATPase</td>
<td>δ'</td>
<td>p37</td>
</tr>
<tr>
<td></td>
<td>Binds clamp</td>
<td>δ</td>
<td>p40</td>
</tr>
<tr>
<td></td>
<td>Binds SSB</td>
<td>Ψ</td>
<td>p38, p36</td>
</tr>
<tr>
<td></td>
<td>Unknown function</td>
<td>Ψ</td>
<td></td>
</tr>
<tr>
<td>III. Processivity factor</td>
<td>β</td>
<td>gp45</td>
<td>PCNA</td>
</tr>
<tr>
<td>(sliding clamp)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Replicase subunits and subassemblies are isolated as three pieces: the polymerase/exonuclease, the accessory complex (clamp loader), and the processivity factor (sliding clamp). For complexes, the individual subunits are listed along with the present knowledge of its associated function. The *E. coli* Pol III holoenzyme contains one further subunit called ε that binds one γ complex and two Pol III cores. The components of the yeast and human replicases have the same names and very similar structure; therefore, the yeast and human components are grouped under the heading 'Eukaryotic', but the subunit masses listed are particular to the human replicase.

Fig. 2. Assembly of a processive chromosomal replicase. In the *E. coli* system, the γ complex clamp loader (accessory complex) recognizes the ssDNA–dsDNA junction, binds the β subunit sliding clamp (processivity factor), and then couples hydrolysis of ATP to assemble the ring-shaped clamp around DNA forming the pre-initiation complex. The γ complex dissociates from the pre-initiation clamp and is capable of assembling other β clamps on other DNA molecules. The Pol III core associates with the β sliding clamp to form the initiation complex, which is capable of highly processive polymerization. This mechanism may apply generally to other replicases.
γ complex can be removed from the reaction prior to adding Pol III core without effect on the rate and processivity of DNA synthesis (for recent reviews, see [47,50,51]).

How similar are the T4 replicase and eukaryotic pol δ to E. coli Pol III holoenzyme? Both T4 and eukaryotes have an accessory complex that is analogous to the E. coli clamp loader γ complex. In yeast and human, the accessory complex is replication factor C (RF-C, also called Activator-1) (reviewed in [52]) and in T4, it is the complex of the gene 44 protein (gp44) and gene 62 protein (gp62) (gp44–gp62, reviewed in [53]). These accessory complexes are DNA-dependent ATPases that are stimulated by their respective clamp protein, and in all cases a primed template is the best DNA effector. Since 1992, the genes encoding the five subunits of human RF-C have been identified, as have the remaining four genes of the five-subunit E. coli γ complex [54−57,58−61]. Amino acid sequence comparison shows that several subunits of RF-C are homologous to the γ and δ subunits of E. coli γ complex and to the T4 gp44 [56]. The level of homology is sufficient to predict that they will have similar three-dimensional structures. Despite these structural similarities, it remains to be established whether the T4 and eukaryotic accessory complexes are truly clamp loaders and whether they need to remain associated with the polymerase during elongation. Recent experiments, however, are consistent with their being clamp loaders and indicate that they may not be needed during elongation.

In the T4 system, the processivity factor is the gene 45 protein (gp45), and in yeast and humans, it is the proliferating cell nuclear antigen (PCNA) (Table 1). But, do they really form rings like the E. coli β subunit? None of these proteins has a significant level of sequence homology with any other, and a major difference is that PCNA and gp45 are trimers, whereas β is a dimer, and they are only 2/3 the size of β. The sixfold symmetry of the β dimer held the explanation for these differences in size and aggregation state. A trimer of PCNA (and gp45) is of similar mass to a dimer of β so, if the PCNA and gp45 monomers were two-domain proteins, instead of three, then the trimer would contain a total of six domains like the β dimer. Indeed, a sequence alignment of gp45 and PCNA with the first two domains of β, using the structure of β as a guide, shows that the hydrophobic core residues are positionally conserved, lending strength to this hypothesis [49**]. As of September 1993, the answer was determined for the yeast PCNA, and the answer was a most definite yes, as the X-ray structure shows it to be a trimer in the shape of a ring with the same outside and inside diameters as the β dimer (J. Kuriyan, personal communication). Further, the affinity of human PCNA to DNA (placed there by RF-C) depends on the geometry of the DNA, being tightly retained on circular DNA, but freely sliding off upon linearization of the DNA; therefore, PCNA exhibits the same hallmarks of topological binding to DNA as the E. coli β subunit (N. Yao, Z. Kelman, Z. Dong, Z-Q Pan, J. Hurwitz, M. O'Donnell, unpublished data).

Now, back to the issue of whether the accessory proteins of T4 and eukaryotes are clamp loaders, and whether they must be present with the polymerase and clamp protein during elongation. Three recent reports in these systems indicate that the clamp–polymerase unit is all that is needed for processivity. Clever experiments in the yeast system show that on linear

---

Fig. 3. Crystal structure of the β DNA-sliding clamp of DNA polymerase III holoenzyme. The structure on the left is viewed 'face on' looking through the central cavity. The outside is a continuous layer of antiparallel sheet, which also forms the dimer interfaces (arrows). The central cavity is lined with 12 α helices, the only helices in the entire molecule. The structure on the right is the β dimer turned on its side, with duplex DNA modeled through the center. The two structurally distinct faces (A and B) are the result of the head-to-tail arrangement of the subunits. Reproduced from Molecular Biology of the Cell 1992, 3:955 by copyright permission of the American Society for Cell Biology.
DNA (but not circular) PCNA confers processivity onto pol δ in the complete absence of ATP and RF-C [62**]. This result was interpreted as showing that the PCNA ring threads itself onto the end of linear DNA and then couples with pol δ for processive synthesis, thus circumventing the need to open the ring, an action that requires the RF-C complex and ATP. In the T4 system, evidence that the gp44-gp62 complex is not needed during elongation has come from the observation that supply of a large excess of gp45 increases the processivity of the polymerase in the absence of the gp44-gp62 complex, a result that is similar to those from previous studies in the E. coli system and which enforces the idea of the clamp-polymerase as the processive unit [63*]. Furthermore, an elegant electron microscopy study has shown that the T4 sliding clamp on DNA appears as a 'hash-mark' (with similar dimensions to the β ring) through which DNA is threaded, and these 'hash-marks' appear in clusters indicating that they can slide [64**]. The size of the 'hash-mark' is insufficient to accommodate the mass of both the gp44-gp62 complex and the gp45 trimer and in light of similarity to other systems it is probably a gp45 clamp.

The actual mechanism by which the clamp loader assembles the clamp around DNA is a fascinating issue and is being addressed in all these systems, but detailed mechanisms are still unknown. Most of the work has been in the T4 and E. coli systems, in which the individual subunits of the clamp loaders are available in pure form. No individual subunit of either of these clamp loaders can assemble their respective clamp onto DNA; presumably this reaction is too complicated for just one protein. The five subunits of the E. coli γ complex are γ, δ, δ', χ, and ψ. Although only γ and δ are essential to place β onto DNA, the δ' subunit stimulates this reaction considerably [65,66]. DNA-dependent ATPase activity is produced by a mixture of γ and δ' [65] and γ is the presumed site of hydrolysis, as it is known to bind ATP [1]. The δ subunit forms a protein–protein complex with β [67]. Hence, it appears that γδ' recognizes the primed template, and the δ subunit functions to bring β into the structure for assembly around DNA and hydrolysis of ATP. At elevated ionic strength, such as exists in the cell, the χ and ψ subunits of the γ complex are also needed to initiate processive synthesis [68]. This is probably rooted in the fact that the γ complex associates with SSB-coated DNA [69*], an interaction mediated by χ that may give the γ complex the added grip that it needs to bind the template in elevated salt (Z Kelman, M O'Donnell, unpublished data).

The T4 clamp loader is composed of a tetramer of gp44 tightly associated with one protomer of gp62 [53]. By itself, the gp44 is a DNA-dependent ATPase (like E. coli γδ'), implying that it harbors the DNA-binding and ATPase sites [70]. The DNA-dependent ATPase of the gp44-gp62 complex is stimulated by gp45, whereas the gp44 ATPase is not, implying that gp62 interacts with gp45 (like E. coli δ) [70]. The gp44-gp62 complex interacts with gp32 (T4 SSB) bound to ssDNA (like E. coli χ), but it is not known whether the interaction is mediated by gp44 or gp62 [71]. Using footprinting methods, laser-induced UV-crosslinking of protein to DNA, and novel DNA–protein cross-linking agents, it has been found that the T4 gp44–gp62 complex needs only to bind ATP to bring the gp45 clamp to DNA, although it is not known whether the clamp surrounds DNA at this point [72,73**,74]. Further, these studies show that the gp44–gp62 complex works on the primed template junction to place the gp45 clamp on DNA, and then ATP hydrolysis induces a movement or the dissociation of gp44–gp62 from the DNA, presumably to make room for the polymerase [74,75].

Much less is known about the RF-C subunits, although with the recent identification of the genes encoding them we can expect much more information in the near future. The human p128 is known to bind DNA, as it was identified in 1993 in a south-western analysis as a DNA-binding protein called PO-GA [57*]. At the same time, a similar screen resulted in the isolation of the analogous RF-C subunit of the mouse [76]. The p37 and p40 proteins are the only RF-C subunits to be obtained in pure form thus far and studies have shown that p37 binds DNA, suggesting that p128 and p37 may be similar to γ and δ' (and T4 gp44) [77*]. The p40 binds ATP and PCNA, which suggests a functional analogy to δ (and T4 gp62) [77*]. A summary of these individual subunit functions is included in Table 1.

In all these systems, the three replicase components have only weak interactions with one another, at least in the absence of DNA. In the E. coli system, however, there is another subunit called τ that firmly binds one γ complex molecule and two molecules of Pol III core, thereby acting as a scaffold to hold all of these proteins tightly together in a 'holoenzyme' particle [67]. A dimeric polymerase complete with one clamp loader within a single molecular structure fits nicely with the need to synthesize two strands of DNA, one of which is synthesized discontinuously and requires multiple clamps to be loaded onto it (lagging strand). At present, the T4 and eukaryotic replicases have been purified as the three separable components and whether, in these systems, the clamp loaders maintain an association with the polymerase and clamp or have a functional equivalent of τ to hold them together remains to be determined.

Replicases of HSV-1 and phage T7

The other two replicases that are highly processive are those of HSV-1 and phage T7 [1]. These replicases require only one accessory protein for high processivity, UL42 for HSV-1 and E. coli thioredoxin for T7, and they do not need ATP. Thus, they lack a clamp loader. The crystal structure of thioredoxin bears no resemblance to a ring, so the molecular basis by which the accessory protein provides processivity to the polymerase is not clear. One simple possibility is that the polymerase has a cleft into which the DNA fits and the accessory protein seals off this cleft to trap DNA inside.
Interaction of sliding clamps with other proteins

The clamp proteins have recently been shown to share a common property, in that they interact with proteins other than the replicative polymerase. The best studied case is that of T4, in which the gp45 clamp interacts with RNA polymerase. Phage T4, as well as several viruses (e.g. SV40, adenovirus and HSV), has an early-to-late switch in gene expression, so that late genes (e.g. capsid proteins) are not expressed until the viral genome has been replicated. Elegant studies by Peter Geiduschek's research group have revealed that the underlying mechanism of the early-to-late switch is an interaction between the T4 replicase accessory proteins and the E. coli RNA polymerase (modified by T4 gp33 and gp55). In 1992, a continuation of these studies was published showing that this switch is the result of a tracking mechanism, whereby the accessory proteins assemble at a nick and slide along DNA, thus acting as a 'mobile enhancer' for late gene activation [78**]. Presumably the interaction is through the gp45 clamp and the modified RNA polymerase, although a role for the gp44-gp62 complex cannot be ruled out, as it was not removed from the system after formation of the sliding clamp. This mechanism may apply to eukaryotic viruses in general, as a baculovirus protein with 42% identity to PCNA is also needed for activation of late genes [79].

Use of the sliding clamp for other DNA metabolic processes appears to be the rule rather than the exception, as the clamps of E. coli and of human also interact with other proteins. In E. coli, the β clamp is utilized by Pol III and DNA polymerase II, and in humans and yeast, the PCNA clamp is utilized by both pol δ and pol ε [1]. In fact, the human PCNA has recently been shown to bind to members of the cyclin D and cyclin-dependent kinase families [80**]. The function of the latter interaction is unknown. Would the PCNA ring need to be on or off DNA to manifest the biological activity? It is tempting to speculate that the PCNA ring may tether cyclins and their associated kinases to DNA for some aspect of cell cycle control (e.g. mitotic checkpoint).

Conclusions

In summary, knowledge is expanding rapidly about replication mechanisms in several systems, too many, in fact, for them all to have been mentioned here. Chromosomal replicases appear to be quite similar in their mechanism for attaining high processivity. However, little is known about how they assemble at the origin or communicate with the helicase and primase, or how these processive enzymes release DNA rapidly and rebind new primers during discontinuous synthesis of the lagging strand. Likewise, the outline of events in activation of origins has been conserved in evolution from prokaryotes to eukaryotes, but important details in the manipulation of these sequences by proteins, the roles of ATP, and regulation of initiation in different systems is revealing fundamental diversification. The new information about yeast origin structure, and the proteins that act on it, lead us to anticipate future discoveries concerning the mechanisms by which eukaryotic cellular origins are activated, regulated and eventually integrated with signal transduction.

References and recommended reading

Papers of particular interest, published within the annual period of review, have been highlighted as:

- of special interest
- of outstanding interest


Urea treatment monomerizes RepA and completely bypasses the requirement for heat shock proteins for the repli- cation of an oriP plasmid, showing that heat shock proteins are required only at the nucleoprotein-stabilization step.


In the absence of SV40 DNA, ATP promotes hexamerization of T-antigen. This hexamer is stable on both DNA and a glyceraldehyde gradient, but cannot bind the origin. Inhibition of ATP monomerizes T-antigen and activates it for proper assembly onto the origin upon adding ATP and DNA. This suggests that the hexamers are required for the origin binding.


The multi-element nature of ARS1 was correlated to the known complexity of transcriptional promoters. This paper describes the purification of the origin recognition comments of ARS1. The multi-element nature of ARS1 was correlated to element A and ABF1 on B3. An additional footprint over element B2 and even then ATP was needed.

This report describes studies showing that defective MCM2 and MCM3, which are required for the Initiation of DNA Synthesis at Chromosomal Replication Origins in Yeast. This report describes studies showing that defective MCM2 and MCM3 proteins result in defects in origin-specific replication.

MCM5 is shown to be identical to CDC64, and mutants are shown to be defective in maintaining minichromosomes. The paper describes the cell cycle regulated subcellular localization of MCM5.

This paper reports limited sequence homology between the p40 subunit of RF-C and gp44 and E. coli. t and t.

PCNA was shown to be capable of loading onto the end of a long stretch of duplex DNA (0.55 kb), implying that it can slide long distances, but it could not load over ssDNA coated with SSB. These are similar mobility characteristics to those of the E. coli β sliding clamp.

An interesting observation that use of a large amount of gp45 results in clamp assembly on a circular DNA molecule without the gp44-gp62 complex or ATP.


This study identifies an association between y complex and ssDNA that requires SSB. The paper also shows that the y complex can place β onto DNA without SSB.


This paper shows that UV-induced cross-links between DNA and gp44–gp62 complex, indicative of a conformational change of gp44–gp62 relative to DNA upon placing gp45 on DNA. The authors present models by which gp45 is placed on DNA by gp44–gp62 complex.


This paper describes the cloning of the gene encoding RF-C. The authors note regions of homology between the carboxy-terminal portion of the large subunit of murine RF-C and the four smaller subunits of human RF-C.


The authors show that under some conditions p37 stimulates pol ε, but not pol δ. They also show that p40 inhibits pol δ, suggesting a p40-pol δ interaction.


The nicked site for assembly of the T4 sliding clamp was placed on one DNA ring of a two ring catenane; the other ring contained the T4 late promoter. The sliding clamp on one ring was unable to enhance transcription on the other ring.


80. XIONG Y, ZHANG H, BEACH D: D Type Cyclins Associated with Multiple Protein Kinases and the DNA Replication and Repair Factor PCNA. Cell 1992, 71:505-514.

It is proposed that PCNA is a member of multiprotein complexes containing p21 and combinatorial variations of cyclins D1 and 3 and cyclin-dependent kinases 2, 4 and 5.