We are IntechOpen, the world’s leading publisher of Open Access books
Built by scientists, for scientists

3,900
Open access books available

116,000
International authors and editors

120M
Downloads

154
Countries delivered to

TOP 1%
Our authors are among the most cited scientists

12.2%
Contributors from top 500 universities

WEB OF SCIENCE™
Selection of our books indexed in the Book Citation Index in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com
Chapter 2

Replicative Helicases as the Central Organizing Motor Proteins in the Molecular Machines of the Elongating Eukaryotic Replication Fork

John C. Fisk, Michaille D. Chojnacki and Thomas Melendy

Additional information is available at the end of the chapter

http://dx.doi.org/10.5772/52335

1. Introduction

Major processes in the cell often involve the coordinated and efficient assembly of macromolecular complexes; such examples include: RNA transcription, DNA replication, translation, and cellular motion. These processes can be likened to miniature forms of machines, so-called “molecular machines” with multiple components and motors at their heart driving the systems. This term has been used by several researchers, which equate many of life’s inner workings as homologous to machines; albeit much more efficient than their macro-type counterparts [104]. In 1998, Bruce Alberts wrote an elegant article for Cell noting the inherent beauty of molecular biology’s machines, praising them and stating that as with all machines these macromolecular complexes must in turn contain an assortment of moving parts that act in a highly coordinated fashion with each other [1]. One such studied process is DNA replication, which has been extensively studied since the discovery of the DNA double helix. Due to the biological necessity for duplication of the genetic material, and the intricate link between the faithful replication of the genomic blueprint and its mismanagement leading to cancer, it is difficult to envision a process more important to human health than the study of DNA replication. The motor that drives the molecular machine that is DNA replication is the replicative DNA helicase. Replicative DNA helicases are well known as the motors that drive DNA replication forks along the DNA strands. But in more recent years it is becoming evident that replicative helicases also coordinate the necessary associations and dissociations of the various DNA replication complexes that need to act at the elongating replication fork. Here we
will review the current knowledge of how the molecular motors, replicative DNA helicases, coordinate the actions of the molecular machines that are elongating eukaryotic DNA replication forks.

2. Phases of DNA replication

The replication of DNA during the Synthesis (S) Phase of the cell is generally differentiated into distinct stages. The first is the binding and recognition of the origin of replication by origin binding proteins. For cellular replication in eukaryotes, these proteins are the Origin Recognition Complex (ORC) proteins, many of which belong to AAA+ family of cellular ATPases [20, 97]. To begin activation of the origin (i.e. licensing), two other proteins must act to make origins competent, Cdc6 and Cdt1 [5, 112]. These two proteins in turn are regulated by phosphorylation by Cdc7/Dbf4 as well as by geminin (in metazoans). The presence of ORC/Cdc6/Cdt1 are necessary for recruitment of the next set of vital DNA replication proteins, the minichromosome maintenance (MCM) proteins, which are components of the replicative DNA helicase [70, 115]. For many years, the MCM complex was proposed to be the replicative helicase; but both in vitro and in vivo studies could not verify that the MCM complex was in fact the DNA helicase necessary for eukaryotic replication [53, 68, 137]. However, it was well established that the six ‘core’ MCMs, MCM2-7, were essential for DNA replication and that their deletion was lethal in yeasts [125]. Additionally, MCMs appear to associate with chromatin just prior to S Phase, and dissociate from the chromatin as S Phase progresses, consistent with that of a DNA replication helicase [24, 117]. Only recently was it discovered that the MCM complex appears to be an incomplete DNA helicase, in that several additional proteins recruited during origin activation appear to be required to make up the DNA helicase holoenzyme. Cdc45 and the GINS (in Japanese Go-Ichi-Ni-San, which stands for the numbers 5-1-2-3 in the subunits Sld5, Psf1, Psf2, and Psf3) complex appear to make up the CMG (Cdc45-MCM-GINS), the complex multisubunit eukaryotic helicase [91], required for initiation of DNA replication. In spite of this elucidation of the CMG, the step-wise recruitment of these helicase components, and the complex nature of the post-translational modification steps required to reconstitute a functional CMG replicative DNA helicase, has severely constrained the ability to carry out detailed biochemical analyses of the eukaryotic DNA replication fork.

The formation of an active pre-replication complex at the origin, and the subsequent formation and activation of the CMG replicative DNA helicase allows for the recruitment of DNA polymerase α primase, which is necessary for the synthesis of RNA primers and a short DNA extension of those primers. Also recruited is RPA, the major ssDNA binding complex necessary to prevent the re-annealing of the DNA duplex [132], and topoisomerase I, which resolves the compression of the DNA helix caused by progression of the replication fork along the DNA duplex (Initiation of DNA replication). Following primer synthesis, the clamp loader, RFC, is loaded at the 5’ end of the primers, and RFC in turn loads the DNA polymerase processivity factor, PCNA. Due to the 5’->3’ nature of DNA
replication, synthesis occurs continuously on the leading strand through the recruitment and activity of DNA polymerase ε [106] and discontinuously on the lagging strand by DNA polymerase δ extension of the repeated primers laid down by DNA polymerase α primase [98]. The components of polymerase δ are also often found, not surprisingly, associated with the proteins involved in “processing” the lagging strand Okazaki fragments, namely those proteins involved in removing the RNA primers (see below) [65, 103]. During this elongation phase of DNA replication is when the majority of DNA synthesis occurs. However, while the ORC complex and other components of the origin recognition/licensing machinery are dispensable following origin firing [26], the heart of the DNA replication apparatus remains associated with the replicative helicase, the molecular motor that is actively unwinding the DNA duplex. How the replicative helicase interacts with components of the elongation machinery is probably the least understood remaining aspect of DNA replication and is the focus of this review. Many other components are implicated in the elongation phase of eukaryotic DNA replication, such as Mrcl (Claspin), which has been suggested to be involved in linking the helicase to the polymerases and has been found to be involved in the “uncoupling” of these two aspects of the fork during the DNA damage response [6, 56], and for regulating fork progression during uncompromised DNA synthesis [44, 78, 118, 122].

Following elongation, the RNA primers and the RNA-DNA linkages are removed through the actions of the flap endonuclease-1 (FEN1) nuclease and/or Pif1 helicase and Dna2 nuclease, assisted by RPA and DNA polymerase δ [74, 105, 108]. Following the removal of the primers, gaps are filled in, apparently by the action of the DNA polymerase δ and its cofactors, and the final DNA strands are ligated by DNA ligase I into long uninterrupted DNA chains. The removal of all the primers, filling of the subsequent gaps, and the final ligation of the products represent the completion of S-phase.

3. Model systems for elongation of DNA replication

As mentioned previously, eukaryotic cellular DNA replication is highly complicated, and only recently has the replicative DNA helicase finally been identified as MCM2-7 complexed with Cdc45 and the GINS complex (CMG) [91]; furthermore, the complex nature of assembly and regulation of this CMG replicative helicase has limited the ability to study the eukaryotic replication fork biochemically. However, early mechanistic studies of eukaryotic DNA replication were largely carried out using the small DNA tumor virus SV40 and to a lesser extent the papillomaviruses. What makes these viruses ideal models for the mechanistic study of eukaryotic DNA replication? One reason lies in their small genome size. To facilitate their duplication, these viruses make the most of their small number of ORFs by combining multiple replication functions into one or two proteins, and relying primarily on the host cell DNA replication machinery (see Table 1). In addition, the lack of these viruses utilizing the once-and-only-once per S Phase regulation of DNA replication means that their DNA replication systems were not subject to the complicated and constraining regulatory systems that control replication of cellular DNA.
SV40 DNA replication is driven by a single viral protein, SV40 large T-antigen (Tag), a protein that combines all the core DNA replication functions of the cellular initiation and origin activation proteins listed above for eukaryotic DNA replication. Tag recognizes and binds to the SV40 origin of replication, melts the DNA helix surrounding the origin, and establishes itself into a double hexameric structure. Tag then recruits the cellular DNA replication factors: RPA, topoisomerase I, and polymerase α primase. These four replication factors are all that is required for the initiation of SV40 DNA replication through the initial synthesis of RNA-DNA primers. Following these initiation events, the clamp loader, RFC, and the polymerase processivity factor, PCNA, are recruited and loaded, which leads to the binding and activity of DNA polymerase δ, which extends both lagging and leading strands in this viral DNA replication system. As in the mammalian system, Okazaki fragments are processed by FEN1, DNA helicase 2, and DNA ligase 1, completing synthesis of the viral DNA genomes. It was the early studies of this viral DNA synthesis system that elucidated these basic mechanisms of how eukaryotic DNA replication is carried out.

<table>
<thead>
<tr>
<th>Replication Step/Function</th>
<th>Mammalian</th>
<th>SV40</th>
<th>Papillomavirus (PV)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Origin Recognition/Initiator</td>
<td>Orc complex (2-6)</td>
<td>T-antigen (Tag)</td>
<td>E2/E1</td>
</tr>
<tr>
<td>pre-RC</td>
<td>Cdc6, Cdt1, Cdc45, Geminin, MCM10, Sld2(RecQL4), Sld3, Dpb11(TopBP1)</td>
<td>Tag</td>
<td>E2/E1</td>
</tr>
<tr>
<td>Helicase</td>
<td>MCM 2-7, GINS, Cdc45</td>
<td>Tag</td>
<td>E1</td>
</tr>
<tr>
<td>SSB</td>
<td>RPA</td>
<td>RPA</td>
<td>RPA</td>
</tr>
<tr>
<td>Torsional relaxation</td>
<td>Topoisomerase I</td>
<td>Topoisomerase I</td>
<td>Topoisomerase I</td>
</tr>
<tr>
<td>Clamp loader</td>
<td>RFC</td>
<td>RFC</td>
<td>RFC</td>
</tr>
<tr>
<td>Processivity factor</td>
<td>PCNA</td>
<td>PCNA</td>
<td>PCNA</td>
</tr>
<tr>
<td>DNA polymerases</td>
<td>DNA pol α primase, DNA pol δ</td>
<td>DNA pol α primase, DNA pol ε, DNA pol δ</td>
<td>DNA pol α primase, DNA pol δ</td>
</tr>
<tr>
<td>Accessory factors</td>
<td>Mrc1(Claspin)</td>
<td>None</td>
<td>None?</td>
</tr>
</tbody>
</table>

Table 1. Known and Proposed Components of the DNA replication complex

Similar findings were also found for another virus family, the papillomaviruses, which have also proven to be an apt model for cellular DNA replication mechanisms due to
their dependence on the host replication machinery. Initial studies were carried out in the bovine version BPV-1, and later corroborated with several human HPV isotypes. In general, papillomaviruses follow the same mode and progression of events found in SV40, except for the need for two viral proteins instead of the single Tag protein required for SV40. In addition, PV appears to require other cellular factors that SV40 does not [73, 80, 87], which to date remain unidentified. In papillomavirus DNA replication, the E2 protein assists and directs faithful viral origin recognition of E1 [79, 90, 110, 126], while E1 itself serves the role of the replicative DNA helicase, melting the DNA around the origin of replication and establishing itself as a double hexameric helicase. In a fashion similar to that of SV40 Tag, E1 also acts to recruit the cellular DNA replication proteins to the PV DNA replication fork [36, 113, 131]. E1 itself is a weak origin binding protein, but can bind to and unwind DNA even in the absence of E2 at high E1 concentrations, even on DNA without an apparent E1 binding sequence and is therefore relatively nonspecific without E2 [66]. Furthermore, following establishment of the double hexamer, the E2 protein is purportedly absent from subsequent steps of DNA replication, indicating E1 is the only viral protein implicated in the actual HPV elongating DNA fork [72]. Otherwise, these two small DNA viruses display very similar mechanisms of replication, especially during the elongation phase. So why rely on two very similar viruses as models and not just SV40? One reason is that by comparing and contrasting the DNA replication mechanisms in two subtly different systems, one gains further insight into the mechanisms of DNA replication. In specific aspects of DNA replication, one or the other virus might provide a more apt reflection of the mechanism of cellular DNA replication. Another reason lies in the diseases each virus causes and the implications for antiviral research. Although SV40 Tag is a potent transforming agent for cell culture due to its ability to inactivate p53, Rb protein, and many other components of the cell, SV40 itself does not appear to readily cause tumors in humans. Conversely, human papillomaviruses are the major cause of cervical, anogenital, and oral cancers and represent the major cause of infectious-agent-induced cancers in humans. These viruses represent historically important and still valuable models for DNA replication and can still be used to elucidate hitherto unknown mechanisms of mammalian DNA replication. Furthermore, the replicative DNA helicases of these viral DNA replication systems still provide the best biochemical system for investigating the role of DNA helicases in the elongation stage of eukaryotic DNA replication.

4. Replicative DNA helicases

When the structure of the DNA double helix was first proposed, one of the major questions concerning the replication of dsDNA was how the duplex would be opened to facilitate reading of the base sequence encoded by the DNA. The first such discovered protein that could carry out this function was the prokaryotic helicase of *E. coli*, discovered in 1976. All known helicases use the energy from NTPs to drive the remodeling of their substrate nucleic acids [75, 85]. Helicases are grouped into six superfamilies (SF1-SF6) and all
possess typical Walker A and B motifs involved in NTP binding and hydrolysis. The motor proteins of the macromolecular machines at DNA replication forks are all AAA+ module-containing helicases, which function to unwind the DNA helix and to drive the replication machinery along the DNA template. Another common characteristic of replicative helicases is that most form higher order oligomeric structures to facilitate their functions as DNA helicases at DNA replication forks. The MCM complex of the CMG cellular helicase, SV40 Tag, and PV E1 all form hexamers. Both Tag and E1 have been recently crystalized in their hexameric forms, which has contributed significantly to elucidating how these helicases function in splitting the DNA helix [32, 38, 71]. Further, Tag and E1, and later MCMs, were shown to form dimers of two hexamers [34, 36], which are presumed to act in bridging the two DNA replication forks, holding them together during replication fork progression, and creating a system whereby the template DNA is threaded through the DNA replication machinery in both directions simultaneously.

Various models have been proposed for how DNA helicases unwind the DNA helix. Some early proposals included the monomers binding to the DNA backbone and essentially rolling one DNA strand away from the other using the circular nature of the hexamer. Other models included a hexamer ‘embracing’ ssDNA, excluding it from its partner, or two hexamers acting at a distance pumping dsDNA through their central pore. Some studies indicate the double hexamers stay associated during elongation, and this led to a double hexameric DNA pumping mechanism that pumps dsDNA through the central pore somehow splitting the helix [42]. The more recent structural studies of the BPV1 E1 helicase bound to DNA, ATP, and ADP indicate an intricate hybrid model whereby the E1 hexamer pumps ssDNA through each central pore in a staircase type mechanism as ATP is bound and hydrolyzed by each subsequent E1 monomer [32, 33, 109]. In this model E1 uses the ATP binding/hydrolysis-induced conformational changes of the individual monomers to drive each nucleotide base of the enclosed ssDNA template through the central pore, displacing the hybridized (lagging-strand template) DNA strand freeing it to be available as a template for lagging strand DNA synthesis [32]. Although the model for helicase action based on the SV40 Tag structure was not the same, the Tag structure was done in the absence of ssDNA, and the structural information on the Tag hexamer would be consistent with a helicase model similar to that of E1.

5. Helicase interactions with replication proteins that initiate elongation

As stated previously, DNA replication proteins commonly recruited by both of these viral replicative helicases are: RPA, topoisomerase I, and DNA polymerase α primase. In this section, we will look closer at the individual and combinatorial interactions between the helicase and these necessary DNA replication factors that are intimately involved in both the initiation and elongation stages of DNA replication. In many cases, studies have focused on specific interactions, often detailed down to specific amino acid residues required for recruitment of these factors. Various groups have used the powerful ability to investigate the interactions of these factors with the viral helicases both in vitro and in vivo, to elegantly
demonstrate the importance of these molecular contacts. For each of these three DNA replication factors, we will look into the extensive work that has been performed in the SV40 system with Tag, then in the PV system with E1. Following this, we will briefly touch on the mammalian system, highlighting some of the similarities between the viral and the mammalian host systems.

6. Helicase interactions with replication proteins that initiate elongation: Topoisomerase I

The unwinding action of the DNA replication fork driving along the DNA helix creates torsional stress and overwound DNA that must be relieved to allow replication to proceed. Topoisomerases are enzymes that help relieve this stress and aid in maintaining chromosome structure and integrity by modifying DNA topology, and resolving specific DNA structures that arise from cellular processes such as DNA repair, replication, transcription, recombination and chromosome compaction [13]. These processes result in compression (positive supercoiling) of the DNA helix and the entanglement of DNA segments and chromosomal regions that can lead to cytotoxic or mutagenic breaks in the DNA if left unmanaged [127]. Hence, topoisomerases play a vital role in living cells, particularly during DNA replication.

Enzymatically, topoisomerases act through the action of a nucleophilic tyrosine; the enzyme cleaves one or more DNA strands and generates an enzyme-DNA complex that serves to prevent the release of nicked or broken DNA that could possibly result in chromosome damage [127]. After passage of one or more DNA strands through this transient break(s), the topoisomerase re-ligates the strands leaving the original DNA sequence intact. Though all topoisomerases have this feature in common, topoisomerases are separated into two classes, type I and type II, depending on whether they cleave one or two strands of DNA, respectively [127]. Type I topoisomerases act on one strand, and generally pass a single DNA strand through the transient break, while type II topoisomerases break both DNA strands and generally pass dsDNA through the transient break. Type I topoisomerases generally work in front of replication or transcription forks, to relax positive supercoils in a highly processive manner; while type II topoisomerases are involved in untangling intertwined duplex DNA such as that found in newly replicated molecules or during chromosome resolution during cell division [30].

Topoisomerases have roles in each of the major replicative phases: initiation, fork progression and termination. During DNA replication in eukaryotes, topoisomerases have been observed to bind directly to the replication origin to aid in activation in the initiation phase [45, 127]. During strand synthesis, topoisomerases are required to alleviate compression of the DNA helix caused by positive supercoiling that results from DNA unwinding, which is mediated by replicative helicases [127]. Topoisomerases are also required for daughter strand resolution. Eukaryotes rely on topoisomerase I (topo I) to fulfill the initiation and elongation functions during DNA replication [127].
Human topo I is an ATP-independent, 100-kDa monomeric protein capable of relaxing positive or negative superhelical twists by making a transient single-strand break, thus relieving the tension generated by the replicative helicases during the DNA-unwinding process [61, 127, 135]. Topo I can be divided into four domains: the highly charged NH₂-terminal domain; the conserved core domain; a short, positively charged linker domain, which links the N-terminal domain to the core domain; and the highly conserved COOH-terminal domain, which contains the active-site tyrosine [116]. Due to the topologically constrained nature of a circular dsDNA molecule, it is no surprise that topo I is required for the replication of the genomes of small circular double-stranded DNA viruses. The role of topo I in DNA replication of the small DNA circular DNA viruses was first noted when it was observed that the extent of DNA replication in SV40 DNA replication in vitro was limited by the level of topoisomerase activity; addition of topo I to crude extracts stimulated SV40 DNA replication in vitro [51]. This effect could have been due to an enhanced rate of chain elongation resulting from an increased efficiency of unlinking of the parental DNA strands [135], or merely due to the presence of limited levels of topo I in the extracts used. Ultimately it was shown that the DNA replication of SV40 and PV both require topo I [134, 136].

While the role of topoisomerases in DNA replication had always been presumed to be due to their need to resolve topological constraint, more recent studies have indicated that topo I plays additional, highly specific, roles in DNA replication of the small DNA viruses, SV40 and PV. Topo I appears to be involved in the very earliest stages of DNA replication, namely origin recognition. It is evident that topo I is stably associated with the initiation complex and is one of the first cellular proteins to be recruited to the initiation machinery [11,45]. Topo I was shown to preferentially associate with the fully formed Tag double hexamer initiation complexes and to be recruited to the initiation complex prior to the beginning of unwinding [11]. This stable association of topo I with Tag results in an increased specificity of Tag for duplex unwinding at the origin by inhibiting unwinding at non-origin sites [39]. Perhaps for this reason, topo I was observed to be required at initiation to stimulate DNA replication in vitro, and was shown to have no effect on replication if introduced during the elongation phase, indicating it enhanced the synthesis of fully replicated DNA molecules by forming essential interactions with Tag and enabling initiation [45,11]. In contrast, topo I specifically enhances origin binding of PV E1 several-fold, but has no effect on non-origin binding [14]. After origin binding, E1 recruits topo I to the replication fork through direct protein interactions and the relaxation activity is strongly enhanced [14,4]. This enhancement of topo I is critical to relax the supercoiling created by the progressing replication fork during the elongation phase of DNA replication. Notably, although topo I plays a significant role in where Tag unwinds the DNA, topo I does not activate origin binding or unwinding and does not structurally distort the DNA [39]. Nonetheless, the similarities in these findings indicate that topo I plays an active role in origin recognition/specificity for the replication of both of these small DNA viral systems. Moreover, following initiation, the topo I-helicase complex remains stably associated and moves with the replication fork during DNA replication [45].
Topoisomerases have been proposed to act together with DNA helicases as “swivelases”, tightly coordinating DNA duplex unwinding with the topoisomerase relaxing activity during DNA replication [15, 30, 61]. With the progression of the replication fork and unwinding of duplex DNA, topo I is needed to release the torsion created by the progressing replication fork [37]. Optimally topo I should be present and its activity regulated to suit the pace of the helicase [37]. This suggested that there might be direct interactions between the helicases and topo I, and that might be modulation of function due to these interactions. The early finding that topo I was localized at SV40 DNA replication forks supported this concept [4], as did evidence that topo I played an important role in the elongation phase of SV40 DNA replication. Reports of the interactions between SV40 TAg and E1 with topo I were also consistent with the swivelase model [15, 133]. The demonstration that E1 stimulates the enzymatic activity of topo I up to seven-fold and that SV40 TAg also stimulates topo I activity (R. Clower and T. Melendy, unpublished results) provided the first evidence of the cooperative nature of this interaction predicted by the swivelase model [15]. Based on these studies it is clear that the viral helicases interact productively with topo I at DNA replication forks forming active coordinated swivelase molecular machines.

The physical interactions between the viral helicases and topo I have been investigated. In 1996, it was found that two independent regions of Tag, one N-terminal and one C-terminal, bind to the cap region of topo I (see Fig. 1), and binding can take place while DNA-bound. Similarly, for PV E1 it was also observed that topo I binds two distinct regions within E1, within E1’s DNA binding domain (DBD) and at the C-terminus [15, 45]. The E1 C-terminal region was shown to enhance topo I relaxation activity, and to a lesser extent, so did a truncation that included the DBD with additional sequence, flanking either side of the DBD [15]. More detailed studies identified mutants in the DNA binding domain of Tag that were unable to unwind the DNA and were partially defective in their association with topo I, suggesting that this interaction maybe important for proper unwinding of viral DNA at replication forks [114]. More recently, four specific amino acid residues within the C-terminal domain of Tag when mutated were shown to exhibit decreased topo I binding and to abolish SV40 DNA replication in vitro and to have dramatic effects on virus production in vivo [61]. These were the critical results that conclusively demonstrated the vital nature of the helicase-topo I interaction for DNA replication. Though first only believed to be involved in the relaxation of overwound DNA during replication fork progression, topo I has proven to be an integral part of the entire replication process in SV40 DNA replication, including critical roles in initiation and even in RNA-DNA primer synthesis in the elongation phase [37, 60, 61, 123]. Though less well-studied, topo I has been observed to be similarly important in these stages of PV DNA replication. These viral systems are vital models for eukaryotic DNA replication, and as of yet these biochemical studies cannot be recapitulated for cellular DNA replication. The only evidence to date of corroboration of these findings for chromosomal DNA replication is the co-purification of topo I with the GINS-MCM complex [39].
Figure 1. General replication domains of the SV40 T-antigen and papillomavirus E1 helicases. The known domains for SV40 Tag (upper) and HPV-11 E1 (lower) are indicated by horizontal lines. Four domains in E1 have limited homology with SV40 T-antigen (regions A-D). A nuclear localization signal has only been elucidated for BPV-1 E1, therefore the HPV-11 NLS is currently only speculated to be in the analogous sequence area.

7. Helicase interactions with replication proteins that initiate elongation: Replication protein A

One of the first proteins identified as necessary for eukaryotic DNA replication is arguably also one of the most important DNA binding proteins in the cell, the ssDNA binding complex, Replication Protein A (RPA). RPA is a heterotrimeric complex conserved in all eukaryotes, and also shows strong homology to the ssDNA binding proteins of archaeabacteria [57, 59]. The human RPA complex is comprised of three subunits, RPA70, RPA32, and RPA14, and the complex binds to ssDNA with extremely high affinity (approximately 10^{-9} to 10^{-10} M [62]), showing much lower affinity for dsDNA. RPA binds ssDNA with a distinct polarity, in a 5’->3’ orientation [22, 51]. Like SSB [132], RPA is required for DNA replication in vivo; knockdown of the largest RPA subunit, RPA70, using siRNA results in inhibition of DNA synthesis [25]. The presence of RPA
as the ssDNA binding protein is critical in keeping the DNA double helix from reannealing during DNA replication, as well as protecting the exposed ssDNA from nuclease attack. And while other non-related ssDNA binding proteins (such as E. coli or T4 SSB) can support some of these functions (such as ssDNA stabilization and stimulation of the processive DNA polymerases) RPA is specifically required for the early initiation steps of replication, including primer synthesis and stimulation of the DNA polymerase activity of DNA polymerase α primase [10, 81]. RPA is also involved in many DNA recombination and DNA repair pathways, acting as a central coordinator of DNA metabolism [52, 132].

RPA exhibits several DNA binding states. RPA70 has three ssDNA binding sites or oligonucleotide binding (OB) domains and RPA32 has one OB domain [8, 121]. When only RPA70 interacts, this is a lower affinity compacted state, binding to only 8-10 nts. When all four OB domains bind, this represents a higher affinity extended mode that spans ~30 nts [7]. The ability of other proteins to facilitate these binding modes in turn impact the binding of RPA to ssDNA, either covering or exposing various stretches of ssDNA. Since several other proteins bind to RPA through its OB domains, this facilitates a model in which RPA cooperatively hands off and orients the binding of each DNA replication protein through increasing affinity with the subsequent factor [64, 89, 138].

7.1. RPA loading onto ssDNA by replicative DNA helicases

RPA plays many roles in the initial steps of elongation as well as throughout DNA replication. Due to its role in ssDNA stabilization, RPA is one of first proteins required following the unwinding of dsDNA. The critical question here is how this process is coordinated in relation to the double hexameric helicase. The RPA heterotrimer itself makes direct contact with the helicase, be it MCM, SV40 Tag, or PV E1 [3, 43, 77, 95, 101, 130]. The first such studied interaction was through Tag, which interacts with RPA through the helicase’s origin binding domain (OBD) (Figure 1). The importance of this interaction is implied by the absolute necessity for RPA for SV40 replication, RPA cannot be replaced by ssDNA binding proteins from E. coli or even RPA from S. cerevisiae [11, 58, 88]. In turn, RPA interacts with Tag through both its RPA70 and possibly to a lesser degree its RPA32 subunits. In PV DNA replication, the E1 helicase interaction with RPA is also critical for viral DNA replication. E1 directly binds to RPA through its largest subunit, RPA70, but does not appear to bind to RPA32 or RPA14 (unlike Tag which binds RPA70 and RPA32) [43, 52, 69, 77]. Similar to the SV40 system, RPA binds to the PV E1 helicase through its major dsDNA binding domain (Figure 1, Fisk JC and T. Melendy, unpublished data).

Evaluation of the multiple interactions between RPA, E1 and ssDNA in various combinations led to development of a novel model for how DNA helicases may ‘load’ ssDNA binding proteins onto ssDNA being displaced through helicase action [77]. RPA binds well to the E1 protein, but only in the absence of free ssDNA. When RPA was prebound to short (~10 nt) stretches of ssDNA, thereby adopting the short compacted form of RPA, it still bound to E1 as well as RPA not bound to DNA. However, when RPA was bound to longer ssDNA templates (~30 nt or longer), consistent with RPA being in its fully-engaged extended form, RPA would no longer bind to E1. This implied a ‘releasing mechanism’ by which the E1-
RPA interaction would be released upon RPA binding to ssDNA in RPA’s extended form. Based on this data, a model was developed in which free, non-ssDNA-bound RPA is bound by E1. As the E1 helicase unwinds the dsDNA, producing ssDNA, it positions the RPA to bind to the newly exposed ssDNA, releasing RPA from the helicase complex (see Figure 2). As the helicase progresses, subsequent helicase monomers bring subsequent RPA molecules to the ssDNA continuously displaced by helicase action [77]. Very similar results were later shown for SV40 T-antigen, leading to a nearly identical model for RPA placement onto ssDNA during SV40 T-antigen helicase progression [9, 54]. Of course, this simplified model does not take into account topo I or polymerase α primase interactions, but it does suggest how the newly produced ssDNA can be rapidly coated with RPA to prevent reannealing or hairpin formation, and to protect from nuclease attack.

![Figure 2. Generalized model for RPA deposition by replicative helicases.](image)

For simplicity, only a singular replication fork is shown. Free, unbound RPA interacts with a monomer of the helicase. As the helicase rotates relative to the DNA helix, the RPA bound monomer comes in juxtaposition to the freshly unwound ssDNA. A ‘hand off’ occurs, whereby RPA binds to the ssDNA released by the helicase. Upon RPA binding ssDNA in the ‘elongated’ (~30 bp-bound) form, it can no longer be bound by the helicase monomer. This ‘release’ action leaves RPA bound to the newly exposed ssDNA, and allows the helicase to progress along the DNA template. As additional ssDNA is exposed this process is repeated, creating an array of RPA coating the ssDNA lagging strand template. (Adapted from [78]).
8. Helicase interactions with proteins that initiate elongation: DNA polymerase alpha-primase

In bacteria and the T4 bacteriophage, the importance of the primase is clear as they are linked physically to the helicase, which is necessary for efficient lagging strand synthesis [19, 92]. In the T7 bacteriophage, this is even more evident as the primase is actually fused to the functional hexameric helicase [31, 99, 102]. In a more complex fashion, in the mammalian system, GINS/ctf4 are required to link the helicase to the catalytic core of DNA polymerase (pol) alpha [40, 120, 140]. Clearly the interaction between primase and the helicase machinery is conserved throughout evolution.

Pol α primase was the first eukaryotic polymerase discovered in 1957 and was thought to be the only replicative DNA polymerase. The later discovery of the proofreading and highly processive polymerases δ and ε indicated that this was not the case [49, 50]. Pol α primase is a heterotetrameric complex comprised of a large p180 catalytic subunit, the regulatory p68 “B” subunit, and the two primase subunits of p55 and p49. Pol α primase is critical for first synthesizing an approximately 10 nt RNA primer, followed by a short ~20-30 nt DNA extension [23, 41, 119]. Polymerase switching then occurs on this RNA/DNA primer through the action of the eukaryotic clamp-loading complex, RFC, which loads the eukaryotic sliding clamp, PCNA, and then a processive DNA polymerase (DNA pol δ or ε) for synthesis of both leading and lagging DNA strands [124, 128]. RFC is integral here, as it competes with RPA for the end of the primer, disrupting the RPA-pol α interaction and allowing polymerase switching [138]. As with many of the core aspects of eukaryotic DNA replication, the functions of pol α primase were largely elucidated using the SV40 system. Pol α primase is absolutely essential for SV40 DNA replication in vitro [94]. Tag interacts with three subunits of pol α primase [17, 27-29, 48, 100, 129]. Recent work has demonstrated the importance of the Tag-p68 interaction for facilitating priming in both cell-free systems and in monkey cell culture [46, 139]. Mutations in Tag that abrogate the Tag-p68 interaction, but do not affect the interactions with p180 or primase, severely decrease priming (in the SV40 monoplymerase assay, which uses a plasmid with the SV40 ori, and purified Tag, pol α primase, RPA, and topo I, all of which are subjects of this review [86]). The amino residues in SV40 Tag shown to be critical for interaction with pol α p68 are H395, R548, K550, and K616, all of which are highly conserved between Tags from other polyomaviruses [139]. Interestingly, the helicase activity of Tag is dispensable for primosome activity (stimulation of priming by pol α primase), indicating that this effect on priming is likely due to the protein-protein interaction between the helicase and the pol, and not some indirect role of DNA helicase action [47]. In general, the Tag-pol α primase interaction mediates a process that allows the helicase to hand off the pol-primase to the ssDNA to enable primer synthesis [16, 35, 88]. Obviously there would also need to be interplay between the helicase and the two proteins competing for the exposed ssDNA, RPA and pol α primase.

The interactions between E1 and pol α primase show some differences between those found with SV40 Tag. Early studies indicated that the p180 catalytic subunit interacted with the N-terminal half of E1, while the p68 subunit interacted with the C-terminus of the helicase [18,
A later study then looked closely at the role of the E1 interaction with pol α primase in regards to supporting HPV-11 DNA replication in vitro [2]. This study confirmed the earlier findings by indicating that E1 interacts with the pol p68 regulatory subunit through its C-terminal half (Figure 1). The presence of E2, whose trans-activation domain binds a similar region of E1, stimulates the E1-p68 interaction; but E2 and p68 nonetheless compete for [2, 83]. This is consistent with a step-wise mechanism whereby E2 helps E1 assemble into a functional helicase, which is then recognized by p68 of the pol α primase complex. No interaction with the pol p180 subunit was detected in these latter studies. Whether this was due to subtle differences in the BPV-1 E1 used in the earlier study and HPV-11 E1 in the later studies has not been further investigated. Regardless, PV E1 appears to interact with pol α primase within the same E1 domain as the ATPase/helicase function. Further studies are necessary to determine if similar residues in E1 as those in SV40 Tag facilitate the binding to pol α primase; these studies may be beneficial as E1 may only use this subunit to bind and recruit pol α primase to the viral replication machinery.

9. Interactions between replication proteins that initiate elongation: coordination

While the earlier sections have alluded to interplay between the multiple cellular replication factors that interact with the viral helicases during DNA replication, the complexity of the interplay between these interactions is what truly epitomizes the term Molecular Machines.

9.1. RPA’s involvement in de-repression of priming

While the interaction of the viral helicases with RPA has been shown to have a direct effect, apparently through the placement of RPA on the ssDNA being displaced by helicase action, this interaction has also been shown to play another vital role in DNA replication: de-repression of priming. RPA binds directly to pol α primase [10, 28, 96], and can stimulate the fidelity and processivity of pol α primase activity [10, 81]. However, when RPA is present in excess, which it is in human cell nuclei [76], RPA strongly represses synthesis of primers by pol α-primase, likely due to the high affinity of RPA out-competing pol α primase for the ssDNA template [16, 88]. While Tag and pol α primase are required for correct initiation of SV40 DNA replication [27, 130], and the interaction between Tag and pol α-primase is sufficient for stimulation of RNA/DNA primer synthesis by pol α-primase on ssDNA [16], these are insufficient for efficient primer synthesis when there is competition with ssDNA binding proteins. Tag can de-repress primer synthesis by pol α-primase, but only when the ssDNA template is coated by RPA, and not by other ssDNA binding proteins or evolutionarily divergent RPA [88]. The interaction between Tag and RPA is vital for de-repression of priming [88, 111]. E1 has similarly been shown to interact with RPA, and RPA is required for PV DNA replication (and RPA cannot be replaced by other ssDNA binding proteins in PV DNA replication). So while the E1-RPA interaction has not been shown to be essential for priming de-repression during PV DNA replication, this is nonetheless likely to be the case.
9.2. Topo I’s involvement in priming

Similarly, in addition to its roles in origin recognition/specificity and release of DNA helix compression during elongation, another role for topo I was elucidated when it was observed that topo I induces pol α-primase to synthesize larger amounts of primers with higher molecular weight [60]. In this study, Tag mutants that failed to bind topo I normally did not participate in the synthesis of expected amounts of primers or large molecular weight DNA molecules, indicating that the association of topo I with the C-terminal Tag binding site is required for these processes. Whether this is due to a direct effect on Tag function at the replication fork, or due to an indirect effect on pol α-primase through Tag (analogous to the effect of the RPA-Tag effect on priming by pol α-primase described above) is unclear. Additionally, topo I was shown to bind directly to RPA, and RPA binds directly to pol α-primase, and can stimulate its DNA polymerase activity. It is unclear whether or not RPA may be influencing the interaction of Topo I with pol α-primase, or vice versa [60]. However these interactions are integrated, the binding of topo I to the helicase domain of Tag significantly enhances the synthesis of DNA-RNA primers and their extension by pol α-primase.

9.3. Helicase interactions with other proteins involved in elongation

What of helicase interaction with the other proteins involved in DNA replication elongation? In the model systems of SV40 and PV little has been elucidated about any direct interactions. Of the proteins involved in elongation, very little is known about the role of helicase interaction with pol δ, RFC, PCNA, or the proteins involved in primer removal: RNaseH, DNA2, Fen I, or DNA ligase I. In the accepted model of SV40 DNA replication, the first primers synthesized by pol α primase on the two strands at the origin become the primers for the leading strand of the opposite fork [124]. After recruitment of RFC, PCNA and pol δ, the leading strand polymerase continuously tracks along behind the helicase action. Since the helicase, in this case Tag, unwinds dsDNA at the relatively slow rate of approximately 200 bp/min [93] while pol δ/RFC/PCNA polymerizes at about 80 nts/sec [12], it is reasonable to speculate that the slower speed of the helicase limits the polymerase in such a way to coordinate the entire machinery mechanism. However, the speeds of polymerases are often assayed on artificial templates, and this rate for pol δ/RFC/PCNA is faster than the measured rate of eukaryotic replication forks (~ 2 kb/min). Conversely the measured speed of Tag is far slower than the measured rate for eukaryotic replication forks. It is likely that coordination between the various factors and complexes involved in the replication fork lead to the final replication fork rate that is not dependent on any one factor, but is a characteristic of the coordinated complex. Indeed, it is critical that these machines are tightly regulated; without a tight molecular machine at the fork, there would be wild exposure of ssDNA via the helicase leading to DNA damage signaling. It should also be noted here that DNA pol ε is not needed in SV40 DNA replication [141]. This finding may be due to the lack of a need for two replicative helicases to duplicate small virus genomes. Alternately, DNA pol ε and TopBP1 (Dbp11) play roles in initiation in mammalian replication; this role may be dispensable or even inter-
fere with the Tag/E1 initiator functions [82, 84]. In *E. coli* DNA replication it has been shown that the tau subunit actually links the leading strand DNA polymerase to the replicative helicase, dnaB [63] (which tracks along the lagging strand template, unlike the case for the SV40 and PV replicative helicases, that track along the leading strand template). It remains possible that these viral replicative helicases may have heretofore unobserved interactions with additional cellular factors involved in the elongation stages of replication that play important roles in DNA replication. This is a potential area for future study.

9.4. Extrapolation to the cellular chromosomal replication fork

The cellular ‘replicative helicase’ is still poorly defined. Some have designated the human CMG helicase (a large 11 subunit complex comprising Cdc45 and the MCMs and GINS sub-complexes [91]) to be the replicative helicase, while others have designated the RPC, the “replisome progression complex”, comprised of the CMG in complex with Mrc1 (Claspin), Tof1 (Tim or Timeless), Csm3 (Swi3/Tipin), Ctf4 (And-1), and the FACT heterodimer (Spt16, and Pob3 (SSRP1) as the ‘true replicative helicase’ [39]. This study found that Mcm10 and topo I associate weakly with this RPC complex, although it is unclear with which specific subunit. It is unknown if the MCM helicase itself interacts with topo I; however, considering the elaborate number of regulatory subunits now known in the eukaryotic helicase supercomplex, this may not be necessary, and may be unlikely. The GINS complex of CMG can bind to and directly stimulate the activity of pol α-primase [21]. A later study showed that the Ctf4 subunit couples the MCMs to pol α-primase and the Mrc1 subunit interacts with polymerase ε [40]. Other studies have found that both Mcm10 and Cdc45 interact with pol α-primase and also found that loss of Mcm10 in yeast led to uncoupling of the MCMs from pol α-primase and resulted in large stretches of ssDNA, a potent DNA damage signal [67, 107]. In human cells, Mcm10 has been suggested to interact with and regulate pol α-primase levels and prevent inappropriate induction of DNA damage [14]. RPA interacts with many components of the RPC, including Mcm3-7, Cdc45, and Claspin (Mrc1) and requires Mcm for chromatin localization [95]. It is intriguing that only RPA appears to directly interact with the MCMs in eukaryotes; this may be due to the intimate linkage with ssDNA and the helicase machine and the highest priority of multicellular organisms to prevent the aberrant signaling of DNA damage through ssDNA coating by RPA. Additionally, in the absence of the RPC interacting protein Mcm10 or in the presence of a mutant zinc finger bearing Mcm10, RPA is also prevented from loading [55]. In general, the major components of the elongation machinery interact with the replicative helicase in eukaryotes through multiple layers of regulation as the RPC complex, a feature that is nonexistent in the simplified machinery presented by these small DNA viral systems. These viral factories simplify the entire complex by using their own central multifunctional helicases. But this simplification has led to the ability to use these viral systems as models where the biochemical nature and functions of these important interactions that occur at the interface of initiation and elongation can be studied.
10. Conclusion

Replicative DNA helicases, modeled by the SV40 and PV DNA replication systems, play complex roles coordinating the multiple actions of multiple DNA replication factors at eukaryotic replication forks. Their interactions with topo I are involved in origin recognition/specificity, DNA helix decompression function, and primer synthesis. Their interactions with pol α-primase are vital for primer synthesis. Their interactions with RPA are involved in loading of RPA onto ssDNA, and de-repression of priming on RPA-coated ssDNA. And the complex interplay between all these factors is intricate, highly-regulated, and appears to be coordinated at least in large part, through the action of the replicative helicases.

Using this wealth of knowledge about the viral replication forks, we have assembled a likely model of replication elongation using the viral helicases as the central molecular machine at the fork. For ease of the various steps of elongation, only a single helicase is pictured in this model (Figure 3). Following assembly of the replication machinery at the viral origin, there is a very intricate four-way interaction comprised of the helicase, topo I, RPA and pol α primase. Topo I has two interactions with helicase; one within the N-terminal half of the helicase and one within the C-terminus. Through these interactions the topo I-helicase interaction assists in helicase origin recognition and creates the swivelase; a machine that couples the unwinding of the DNA duplex with relaxation of torsional stress. During elongation, topo I is likely in front of the helicase to facilitate the easing of positive supercoiling, likely through interactions with the helicase N-terminus. The helicase encircles the leading strand of the newly unwound DNA, actively pumping the leading strand template through the central channel of the helicase and away from the lagging strand replication machinery. While the leading strand template is bound to the central channel and the helicase domain, the lagging strand template is therefore left relatively unprotected. To facilitate a protective role at this point, the OBD of the helicase binds to free RPA, which swings into place as the helicase turns, actively loading RPA onto the lagging strand template. This serves in the role of nuclease protection, as well as preventing aberrant ssDNA structures. However, this coating of the lagging strand template is counterproductive to the process of priming. Therefore, at regular intervals roughly equivalent to the length an Okazaki fragment, the helicase interacts with pol α primase and RPA to facilitate the placement of the pol α primase onto the template, possibly while simultaneously removing RPA in a localized fashion, so that pol α primase can synthesize the RNA-DNA primer. It is intriguing to speculate that it is through this regular placement that Okazaki fragments are placed and spaced; primarily through helicase action and its protein-protein interactions with the primase. Although given the size of eukaryotic Okazaki fragments, it is likely that interactions with histones may play a role as well. The coordinated and highly regulated roles of the multi-subunit DNA helicase in modulating the proteins and their protein-protein interactions involved in the late initiation and elongation stages of DNA replication clearly play a central organizing role in the molecular machine that is the eukaryotic DNA replication fork.
Using the proposed helicase model presented in [33], the replicative helicase is shown oriented with the N-terminal OBD facing towards the unwound dsDNA. For simplicity, only one of the two hexamers is shown. The interaction of topo I with the OBD of the helicase both assists it in origin binding/specificity, and targets it to the incoming dsDNA, where topoisomerase action is vital for replication fork progression. The interaction of RPA within the OBD is involved in the process of directing loading RPA onto the ssDNA newly exposed by helicase action. The interaction of the helicase domain with pol α primase stimulates primer synthesis; and the interaction of the helicase with RPA allows for pol α primase to synthesize primers even in the presence of RPA, through localized RPA removal or ‘priming de-repression’. As each primer is synthesized, RFC, in coordination with RPA, loads PCNA and DNA pol δ onto the 3’ DNA end to allow for processive DNA synthesis. The various interactions of the helicase with topo I, RPA, and DNA pol α primase, as well as other interactions between the cellular factors themselves, coordinately the complex interplay necessary for replication fork function.

Author details

John C. Fisk, Michaeille D. Chojnacki and Thomas Melendy

Department of Microbiology & Immunology, University at Buffalo School of Medicine & Biomedical Sciences, Buffalo, NY, USA
References


MCM2-7 helicase to DNA polymerase alpha within the eukaryotic replisome. Embo J 28:2992-3004.


Proceedings of the National Academy of Sciences of the United States of America 103:10236-10241.


ing protein required for eukaryotic DNA metabolism. Annual Review of Biochemis‐

1993. The DNA-binding domain of simian virus 40 tumor antigen has multiple func‐


merases in simian virus 40 DNA replication in vitro. Proceedings of the National 
Academy of Sciences of the United States of America 84:950-954.

merases in simian virus 40 DNA replication in vitro. Proceedings of the National 
Academy of Sciences of the United States of America 84:950-954.

Mcm4-Mcm6-mcm7 DNA helicase activity. Molecular and Cellular Biology 
19:8003-8015.

reactions for RPA order the assembly of the DNA polymerase delta holoenzyme. 
Embo J 18:6189-6199.

Fanning, and X. S. Chen. 2012. Structural basis for the interaction of a hexameric rep‐
licative helicase with the regulatory subunit of human DNA polymerase alpha-prin‐
mase. The Journal of Biological Chemistry.

Kornbluth, and A. Dutta. 2007. Mcm10 and And-1/CTF4 recruit DNA polymerase al‐
pha to chromatin for initiation of DNA replication. Genes & Development 
21:2288-2299.

Fanning, and T. Nethanel. 1996. DNA polymerase epsilon may be dispensable for 