Mechanisms of DNA replication
Megan J Davey* and Mike O’Donnell†

Advances in recent years have led to exciting new ideas about the initiation, regulation and coordination of DNA replication. Structural studies have yielded fascinating glimpses of replisome action. In addition, the involvement of replication proteins in other cellular processes has blurred the lines between replication, repair and recombination.

Introduction
The replication field has seen numerous exciting discoveries in the past few years. New factors that function with origin recognition complex (ORC) in the activation of yeast origins have been identified. Among the most notable are the minichromosome maintenance (MCM) proteins, which appear to be the long-awaited replicative helicase. A master regulator, the Cdc6 protein, may load an MCM ring onto DNA. Origin events also appear to be regulated by cell-cycle kinases. Several novel mechanisms have been discovered in prokaryotic and eukaryotic systems by which replication systems coordinate their actions. Cell localization studies suggest that numerous replication proteins may act in one large super assembly, called a factory, to initiate replication, copy both DNA strands, and possibly segregate the chromosomes. Advances in structural analysis have essentially caught DNA polymerase in the act of synthesis and tell a tale of unexpectedly large conformational transitions during the catalytic cycle. Also, many proteins that function on DNA are revealed to have a toroidal shape, expanding the use of rings from polymerase processivity factors to ligase, topoisomerases, nucleases and recombinases. Finally, the areas of replication, repair, and recombination are beginning to merge; some proteins and machines are used in all these processes. Soon it will be difficult to talk of these areas separately. Herein, we review a handful of important findings over the past five years.

Replication origins and moving forks
Recent years have seen remarkable advances in the understanding of replication initiation in eukaryotic cells. The groundwork was laid about 10 years ago with the characterization of yeast origin structure and discovery of the six-subunit ORC [1]. ORC is present at origins throughout the cell cycle [2] and interacts with other proteins, invoking action as a replication ‘landing pad’ [3]. In the past few years, new factors have been discovered that function with ORC to activate origins. These factors include the six MCM proteins, which act as a complex. The MCMs are conserved in all eukaryotes [4]. These essential proteins are part of a pre-replicative complex (preRC) that assembles at yeast origins in late G1 [5]. The six MCM proteins form a ring-shaped heterohexamer [6]. Prokaryotic replicative helicases are also ring-shaped hexamers. Surprisingly, two each of Mcm4,6,7 form a hexamer with 5′ to 3′ helicase activity [7,8]. The hexamer structure, activity, and localization to origins in G1, suggest the MCMs may be the chromosomal helicase. The modest activity of the Mcm4,6,7 helicase (50 bp/hour [8]) provokes concern that it is not sufficiently active to power unwinding of chromosomes. However, the replicative helicase of Escherichia coli, DnaB, is also inefficient on its own, but becomes rapid (~1 kb/s) when combined with the chromosomal replicase, DNA polymerase III holoenzyme (Pol III HE) [9]. Perhaps MCM helicase will be activated by its cognate replicase.

It is interesting to note that the MCM heterohexamer does not display helicase activity. In fact, Mcm2, as well as an Mcm3,5 complex interfere with Mcm4,6,7 helicase [7,8]. Hence, Mcm2,3, and 5 may perform a regulatory role. Alternatively, other combinations of MCM proteins may form helicases when present with certain specific proteins.

Another important origin-activating factor is the Cdc6 (cdc18) protein, which appears to be a master regulator [10]. Cdc6 is homologous to Orc1, Orc4 and Orc5 [11] and appears to interact directly with ORC. Recent studies show that Cdc6 localizes to origins early, and is required for the association of MCMs at origins [12]. Whereas the MCMs appear to move with the DNA polymerase at replication forks [13], Cdc6 does not and its action at origins may be transient, perhaps functioning in a similar manner to E. coli DnaC (explained below). Other factors that interact with origins are still being identified. Chief among these is Cdc45, important in origin firing [14**].

The function of these replication factors in eukaryotes bears a striking resemblance to the function of their counterparts in prokaryotes. Like ORC, several molecules of the DnaA initiation protein recognize the E. coli origin, oriC. DnaA utilizes ATP to destabilize AT-rich repeats within the origin (Figure 1a; [15]). The helicase, DnaB, is...
a ring-shaped homohexamer [16,17] that surrounds DNA on the lagging strand [18]. Two molecules of DnaB assemble onto the region of oriC unwound by DnaA. Assembly of DnaB onto the origin requires DnaC, which uses ATP to place DnaB onto DNA. DnaC does not remain in this complex [15]. Recent work shows that primase cannot act until DnaB has begun to move from the origin [19]. Primase then targets the moving helicase for RNA synthesis, explaining why the initiation sites have no unique position, and are sometimes outside the origin. Heterogeneous positioning of initiation sites generalizes to numerous systems, procaryotic and eukaryotic [19,20].

Recent studies have shown that the activity of E. coli DnaB is functionally coupled to the replicate through direct interaction [9]. Hence, although DnaB continues to unwind DNA before and after primer synthesis at the origin, it does not move far, as it progresses slowly by itself. Only after the connection to Pol III HE is established can DnaB proceed to unwind DNA at 700–1000 bp/s.

**Regulation of over-replication**

There are numerous origins along each linear chromosome in eukaryotic cells. Chromosome duplication is complete when replication forks from these origins meet head on. It is important that origins do not fire repeatedly, however, but instead are limited to once per cell division. Cdc6 appears central to this prevention of origin re-firing. Cdc6 is competent for origin activation in a limited section of the cell cycle, from the end of M phase to late G1, before DNA synthesis is activated by cyclin-dependent kinases (CDKs) [21]. Formation of the preRC is prevented by high CDK activity in the rest of the cell cycle [22]. This negative regulation seems to be mediated through Cdc6, as demonstrated by a Cdc6 mutant in Saccharomyces cerevisiae that causes persistent association of MCM proteins to chromatin and over-replication, thus bypassing the negative regulation by CDKs [2].

Recent studies on E. coli replication illustrate novel mechanisms to prevent over-replication. The ability of DnaA to promote origin activation is dependent on its nucleotide-bound state; DnaA–ATP can activate the origin, whereas DnaA–ADP cannot [15]. Interestingly, the conversion of DnaA to the inactive, ADP-bound state is promoted by an actively replicating helicase. This negative feedback is communicated through the β subunit clamp of the Pol III HE along with a second, unknown factor, IdaB [23**].

Over-replication is also regulated by titrating the DnaA initiator at sites distant from the origin. This idea is based on the intriguing observation that deletion of a high-affinity DnaA binding site, datA (located away from the origin), results in over-replication of the chromosome [24].

Yet another regulatory mechanism involves SeqA, a factor involved in the sequestration of hemimethylated, and therefore newly replicated, origin DNA into the membrane [25], preventing reinitiation on the hemimethylated origin. Such sequestration prevents DnaA-mediated unwinding at the origin, thereby preventing assembly of the DnaB helicase and replication forks [26]. Helix unwinding at the origin may also be regulated in eukaryotic cells since the entry of replication protein A (RPA; eukaryotic single strand DNA-binding protein) into the origin in S. cerevisiae requires a cell-cycle kinase [27].

**Replicase action**

Formation of RNA primers and their extension by the replicase requires the cooperation of many proteins. Pol α (primase) synthesizes the primer, and remains attached to it via contact with the single-strand DNA-binding protein, RPA [28,29**]. Next, the replication factor C (RFC) clamp loader must gain access to this site to load the proliferating cell nuclear antigen (PCNA) ring onto the primer. Recent studies show that RFC displaces Pol α by binding RPA and competing with Pol α for RPA contact, thereby releasing Pol α from DNA [29**] (Figure 2). After RFC assembles PCNA onto the primed site, it remains there and prevents heterologous DNA polymerase action on the primer terminus. How then does Pol δ (the functional DNA polymerase on the leading strand of the eukaryotic DNA replication fork) gain access to PCNA and the primer terminus? This exchange is also based in competing contacts wherein Pol δ specifically competes with RFC for RPA, thereby moving RFC out of the way. These RPA-based competition switches underlie the original observation that Pol α and Pol δ switch places in an RFC-mediated reaction [30].
The subunit structure of the eukaryotic Pol δ has recently been expanded to three subunits in *S. cerevisiae* [31**] and four subunits in *Schizosaccharomyces pombe* [32]. Pol δ is a dimeric polymerase, dimerization being mediated by one of the small subunits (Pol32 in *S. cerevisiae*). This subunit appears to be conserved in eukaryotes, and contains the consensus sequence for binding the PCNA sliding clamp. This dimeric polymerase structure generalizes to prokaryotes. The *E. coli* replicase, DNA Pol III HE, has long been known to contain two copies of the DNA polymerase [33], and the T4 polymerase appears to function as a dimer at the replication fork [34].

The overall similarity in prokaryotic and eukaryotic replication is a recurring theme. The competition switch is also conserved. *E. coli* primase contacts single-strand DNA-binding protein (SSB) in order to remain bound to its primed site, protecting it from nucleases and heterologous polymerases. The γ complex clamp loader binds SSB in a competitive fashion with primase, thereby displacing primase from DNA [35].

The replication factory

Studies that locate the position of molecules in cells have recently yielded fascinating results. Localization of *Bacillus subtilis* PolC, the chromosomal replicase, shows it remains stationary at mid cell [36**]. However, the DNA origins move toward the cell poles [37]. These studies support a factory model of replication in which the replicating apparatus remains stationary while the DNA moves through the factory (Figure 3), maybe even segregating the new chromosomes into daughter cells. Initiation factors may also be part of the factory. For example, *E. coli* SeqA displays a cellular localization pattern similar to polymerase [38]. In addition, *dnaA* mutants suppress a *dnaX* mutant (encodes DNA Pol III HE subunits), suggesting that the initiator at least transiently associates with the replisome [39]. Footprinting studies of the two DnaB helicases at *oriC* suggest that they are positioned head-to-head so as to unwind DNA between them, an action that fits nicely with the notion of a fixed factory [19]. A double hexamer of helicase has also been observed for SV40 T antigen [15] and the recently characterized archaean MCM helicase [40**].

Replication factories (or foci) have also been observed in metazoan cells, initially identified as foci of newly replicated DNA [41]. Recent visualisation of these foci in living mammalian cells demonstrates that they too are stationary [42]. These foci contain not only proteins involved in DNA replication such as PCNA, Pol α and DNA ligase [41], but also other proteins such as CDKs, repair proteins and methyltransferase [42,43], suggesting that replication factories may be organisation centers for many cellular processes.

Structural insights into polymerase action

Advances in structural analysis have resulted in astonishing advances in our understanding of polymerase action. Despite significant differences in sequence between families of polymerases, structural studies reveal that they share the structure of a U shaped DNA-binding cleft, resembling a partially opened right hand with fingers, thumb and palm subdomains [44]. Crystallographic
studies on mammalian Pol β, T7 DNA polymerase, and in particular Taq DNA polymerase I (Klentaq) with and without template DNA and nucleotide, have essentially captured polymerases in different stages of a catalytic cycle [44]. Upon binding substrate, Klentaq undergoes a remarkable conformational change, whereby the fingers subdomain moves by as much as 46°, folding over the primed template and the nucleotide-binding site [45••].

This closed structure for polymerase allows conserved, functionally important residues to contact the sugar and phosphate moieties of the incoming nucleotide, essentially forming the polymerase active site for nucleotide incorporation. The resulting tight fit is thought to preclude a mismatch at this site.

The ring-shaped processivity factors of chromosomal polymerases illustrate an elegant use of the toroidal form in DNA enzymology. These sliding clamps bind their respective polymerases and encircle DNA to form a topological link. This mobile tether keeps polymerase associated to DNA while it synthesize several thousand nucleotides [46,47]. In recent years, yet more proteins that function on DNA have been found to be toroids, underscoring the utility of this shape in DNA enzymology (Figure 4). These include hexameric helicases, such as DnaB and MCM proteins [6,16,17,40••], which possibly utilize the toroidal form for processivity. λ exonuclease also uses its ring shape for processivity [48].

Topoisomerases assume a toroidal form, which creates an enclosed space to trap the DNA intermediates of their reactions [49]. DNA ligases and recombinases have also been shown recently to have a ring shaped [50,51]. The diverse functions of these proteins illustrate the usefulness of the toroidal form in many aspects of DNA metabolism and other uses will probably come to light as more examples of toroids are discovered.

**Figure 3**

![Polymerase does not move (gray ellipses) while the DNA threads through it, as seen by the migration of the origin (black circles) towards the pole. Adapted from [36••]. Schematic of how a replisome may be structured within the factory, with a double hexamer helicase. For simplicity, the replisomes are drawn before the initiation of lagging strand synthesis.](current-opinion-in-chemical-biology)

**Figure 4**

![Ring structures of the E. coli β clamp [46], E. coli DnaB [17], λ exonuclease [48], E. coli topoisomerase III [49], and human Dmc1, a recombination protein [50]. Figures have been reproduced with permission from the respective journals.](current-opinion-in-chemical-biology)
Convergence of scientific fields

Advances in the replication, repair and recombination fields are teaching us that these processes are intertwined. DNA damage that is not repaired before the replication fork arrives will halt synthesis, perhaps even resulting in a chromosome. In the latter case, RecBCD and RecA combine their action with Holliday-resolving enzymes to repair the break, and then coordinate with replication proteins to restart replication forks [52]. Exciting new studies reveal that the role of the prismsome is to assemble DnaB on recombination intermediates in the restart process [53••]. Replication forks, stalled at a lesion, coordinate actions with several repair and recombination proteins in the RecF pathway to restart replication [54]. Alternatively, the damaged DNA is simply bypassed by the recently discovered novel polymerases, Pol IV (DinB) and Pol V (UmuC) [55••, 56••]. For example, UmuC is a specialized polymerase that functions in complex with UmuD’ to target the β clamp, which enables it to bypass lesions in a RecA-dependent process [55]. Once across the lesion, Pol III resumes its place with β for rapid synthesis. Eukaryotic homologues of this novel class of DNA polymerases include polymerases η and ζ [58].

Reactions that were once thought to occur in isolation are now seen to be connected. As new connections continue to be discovered, a true appreciation of a given reaction, for example DNA repair, may only be appreciated within the greater context of replication and recombination.

Conclusions

This review has summarized a few of the many exciting advances in the DNA replication field. The coming years will see new discoveries and important advancements on the latest findings. We anticipate the identification of new origin activation factors, co-ordination of protein action in macro assemblies, and an exciting and deeper understanding of how proteins combine their functions to achieve new goals.

Acknowledgements

We thank Manju Hingorani for useful discussions. We are grateful to D Jeruzalmi, E Egelman, BW Matthews and A Mondragon, for use of the images in Figure 4. This work was supported National Institutes of Health grant GM38839 to MO.

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

A PolC-green fluorescent protein fusion provided the ability to directly
polymerase: evidence for a factory model of replication.

25. Lu M, Campbell JL, Boye E, Kleckner N: 

26. Torheim NK, Skarstad K: Escherichia coli SeqA protein affects DNA

27. Tanaka T, Nasmyth K: Association of RPA with chromosomal replication
origins requires an Mcm protein, and is regulated by

28. Domreiter I, Erdle LF, Gilbert IU, von Winkler D, Kelly TJ, Fanning E:

29. Yuzhakov A, Kelman Z, Hurwitz J, O'Donnell M: 

30. Tsurimoto T, Stillman B: 

31. Burgers PM, Gerik KJ: Saccharomyces cerevisiae
replication to the PCNA ring. One of the subunits dimerizes Pol δ, possibly for
simultaneous replication of both strands of chromosomal DNA.

Hurwitz J: DNA polymerase δ from Schizosaccharomyces pombe

33. Kelman Z, O'Donnell M: DNA polymerase III holoenzyme: structure

34. Young MC, Reddy MK, von Hippel PH: Structure and function of the

This report outlines the discovery of a three-subunit structure for S. cerevisiae Pol δ and shows the need for three subunits to observe full coupling to the PCNA ring. One of the subunits dimerizes Pol δ, possibly for simultaneous replication of both strands of chromosomal DNA.

35. Yuzhakov A, Kelman Z, O'Donnell M: Trading places on DNA — a
three-point switch underlies primer handoff from primase to the

36. Lemon KP, Grossman AD: Localization of bacterial DNA

A PolC-green fluorescent protein fusion provided the ability to directly visualize the location of PolC in actively growing cells. The signal remained relatively stationary at mid cell.

37. Webb CD, Teleana A, Gordon S, Straight A, Belmont A, Lin DC, 


coli dnaA which suppress a dnaX(Ts) polymerization mutation and are dominant when located in the chromosomal allele and recessive on plasmids. J Bacteriol 1995, 177:705-715.

hexamer archaeal minichromosome maintenance protein is an

The single MCM of an archaeabacterium was expressed and purified. The MCM definitively contains helicase activity. Weight measurements and detailed electron micrographs indicate it exists as a double hexamer.


42. Leonhardt H, Rahn HP, Weinzierl P, Sobotort A, Cremer T, Zirk D,

43. Yan H, Chen CY, Kobayashi R, Newport J: Replication focus-forming

44. Doublet S, Sawaara MR, Ellenberger T: An open and closed case for

45. Li Y, Korolev S, Wakeman G: Crystal structures of open and closed
forms of binary and ternary complexes of the large fragment of
Thermus aquaticus DNA polymerase I: structural basis for
nucleotide incorporation. EMBO J 1998, 17:7514-7525.

Clever crystal soaking methods were used to obtain the structure of Klentaq at different stages of the reaction cycle. The three crystal structures provided views showing the large conformational changes that occur before and during catalysis.

46. Kong XP, Onrust R, O'Donnell M, Kuriyan J: Three-dimensional
structure of the β subunit of E. coli DNA polymerase III


49. Mondragon A, DiGate R: The structure of Escherichia coli DNA


51. Lee JY, Chang C, Song HK, Moon J, Yang JK, Kim HK, Kwon ST, 

52. Kowalczykowski SC: Initiation of genetic recombination and

53. Lu J, Liewei X, Sandler SJ, Marians KJ: Replication fork assembly at

This study showed the ability of PriA to assemble the replicative DnaB helicase onto a D-loop. The DnaB then supports primase action followed by assembly of Pol III HE to ‘restart’ a replication fork at a recombination intermediate.

54. Cox MM, Goodman MF, Kreuzer KN, Sherratt DJ, Sandler SJ, 

55. Tang M, Shen X, Frank EG, O'Donnell M, Woodgate R, Goodman MF:
** UmuD' C is an error-prone DNA polymerase. Escherichia coli

The report demonstrates that UmuC is a novel DNA polymerase. It represents a new polymerase class because its sequence is unlike Pol I, II or III, bringing the total number of DNA polymerases in E. coli to five. UmuC functions in lesion bypass, and requires many other factors including the β sliding clamp.

** Nohmi T: The dinB gene encodes a novel E. coli DNA

This study identifies the fifth polymerase in E. coli, DinB, a UmuC homolog.

57. Goodman MF: Coping with replication ‘train wrecks’ in Escherichia

58. Johnson RE, Washington MT, Prakash S, Prakash L: Bridging the
gap: a family of novel DNA polymerases that replicate faulty DNA.