

Annual Review of Biochemistry

The Initiation of Eukaryotic DNA Replication

Alessandro Costa¹ and John F.X. Diffley²

¹Macromolecular Machines Laboratory, The Francis Crick Institute, London, UK;
email: alessandro.costa@crick.ac.uk

²Chromosome Replication Laboratory, The Francis Crick Institute, London, UK;
email: john.diffley@crick.ac.uk

ANNUAL
REVIEWS **CONNECT**

www.annualreviews.org

- Download figures
- Navigate cited references
- Keyword search
- Explore related articles
- Share via email or social media

Annu. Rev. Biochem. 2022. 91:107–31

First published as a Review in Advance on
March 23, 2022

The *Annual Review of Biochemistry* is online at
biochem.annualreviews.org

<https://doi.org/10.1146/annurev-biochem-072321-110228>

Copyright © 2022 by Annual Reviews.
All rights reserved

Keywords

DNA replication, cell cycle, cryo–electron microscopy

Abstract

DNA replication in eukaryotic cells initiates from large numbers of sites called replication origins. Initiation of replication from these origins must be tightly controlled to ensure the entire genome is precisely duplicated in each cell cycle. This is accomplished through the regulation of the first two steps in replication: loading and activation of the replicative DNA helicase. Here we describe what is known about the mechanism and regulation of these two reactions from a genetic, biochemical, and structural perspective, focusing on recent progress using proteins from budding yeast.

Contents

1. INTRODUCTION	108
2. REPLICATION ORIGINS AND THE INITIATOR	110
2.1. Eukaryotic DNA Replication Origins	110
2.2. Mechanism of Origin Recognition by the Origin Recognition Complex	112
3. HELICASE LOADING	113
3.1. Structural Mechanism of First MCM Ring Loading	113
3.2. Structural Mechanism of Second MCM Hexamer Loading	115
3.3. Regulation of Helicase Loading	116
4. HELICASE ACTIVATION	117
4.1. Overview of Helicase Activation	117
4.2. Structural Mechanism of Helicase Activation and Replisome Assembly	119
5. CONCLUSIONS	123

1. INTRODUCTION

The emergence of complex, multicellular life on Earth was enabled by the evolution of molecular machines and regulatory pathways that bring about the accurate and rapid duplication of large genomes. In bacteria, like *Escherichia coli*, the entire genome is contained on a single circular chromosome and is replicated by two DNA replication machines (replisomes) traveling in opposite directions from a single initiation site (oriC) known as a replication origin (1). As a consequence, the time required to replicate the entire genome (the synthetic or S phase) in bacteria is proportional to the genome size: Addition of new genes or genetic pathways would come at the cost of a longer S phase. Eukaryotic cells initiate replication from many replication origins distributed across the genome. Because of this replication strategy, the length of S phase in eukaryotes is dependent upon the frequency of origins in the genome and is independent of overall genome size. Disconnecting genome size from S phase length was crucial for supporting the evolution of the large genomes—up to 150 billion base pairs (bp)—seen in modern eukaryotes. For such a strategy to be effective, however, the firing of individual origins must be tightly regulated and coordinated so that the genome is copied exactly once in each cell cycle—nothing can remain unreplicated, and nothing can be replicated more than once.

Eukaryotes arose from an organism related to extant archaeal prokaryotes, most likely from an organism closely related to current *Asgardarchaeota* (2). Orthologs of most of the key proteins that are involved in eukaryotic DNA replication initiation can be found in archaeal genomes (3). Moreover, one can see the beginnings of coordinated, multi-origin replication in Archaea. For example, archaeal species in the *Sulfolobus* genus contain three replication origins on a single circular chromosome. Each origin in *Sulfolobus* is flanked by a gene encoding its own dedicated initiator—a protein that binds specifically to that origin and promotes initiation. How Archaea regulate these origins to ensure they fire once per cell cycle is currently largely unknown.

At the heart of all replisomes is a DNA helicase, and the loading of the helicase by initiator proteins is a key step in initiating replication. Replicative DNA helicases are hexameric rings that encircle single-stranded DNA (ssDNA) at replication forks, using ATP hydrolysis to translocate, displacing the complementary, nontranslocation strand as they proceed (4). Consequently, to establish bidirectional replication forks at origins in double-stranded DNA, the two DNA strands in the origin must first be separated in order for two DNA helicases to be loaded. In *E. coli*, the DnaA

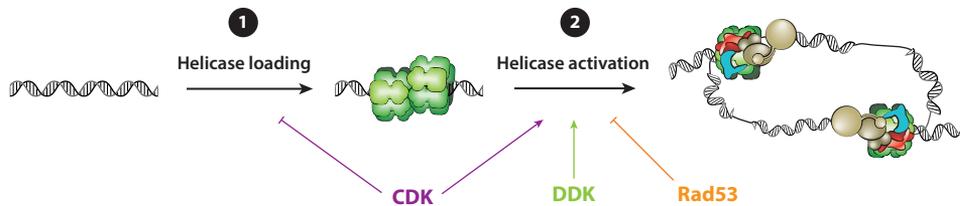


Figure 1

❶ Loading of the eukaryotic replicative helicase onto DNA replication origins leads to the formation of a head-to-head DH of MCM that encircles duplex DNA. ❷ Activation involves the assembly of two CMG holo-helicases that unwind the DNA, providing the single-stranded DNA template for the replicative polymerases. Upon cell-cycle transition into S phase, the activities of CDK and DDK promote CMG formation. At the same time, CDK blocks further DH loading to prevent rereplication. When DNA damage is detected during S phase, the Rad53 checkpoint kinase prevents late-origin firing. Abbreviations: CDK, cyclin dependent kinase; CMG, Cdc45–MCM–GINS; DDK, Dbf4 dependent kinase; DH, double hexamer; GINS, go-ichi-nii-san; MCM, minichromosome maintenance; S phase, synthetic phase.

initiator accomplishes this by binding to *oriC* and generating ssDNA in an adjacent DNA unwinding element; two DnaB helicase hexamers are then recruited via interactions with DnaA, with one hexamer encircling each of the two single strands of DNA. In eukaryotes, the origin recognition complex (ORC) initiator loads the minichromosome maintenance (MCM) helicase—comprising six related subunits, *Mcm2–7*—around duplex DNA as an inactive head-to-head double hexamer (DH) in a process we refer to as helicase loading (**Figure 1**). The DH then undergoes a series of structural changes leading to helicase activation. In this step, MCM generates untwisted DNA within its central channel and then is reconfigured so that each MCM ring encircles one strand of DNA. During helicase activation, two additional proteins—Cdc45 (cell division cycle 45) and the GINS heterotetramer (go-ichi-nii-san, which means 5-1-2-3 in Japanese), consisting of Sld5 (synthetic lethal with Dpb11), Psf1 (partner of Sld5 1), Psf2, and Psf3—become stably associated with each MCM ring to form the active replicative helicase called CMG (Cdc45–MCM–GINS).

The control of the helicase loading and activation steps by global cell cycle regulators ensures that genome duplication is appropriately coupled to cell growth, environmental cues, chromosome segregation, and cell division and also ensures that no origin fires more than once in any cell division cycle. The eukaryotic cell cycle is controlled by cyclin dependent kinase (CDK) and its opposing ubiquitin ligase, the anaphase promoting complex or cyclosome (APC/C). Bistable switches, driven by system-level feedback mechanisms, divide the cell cycle into two states: One state, which is characterized by low CDK and high APC/C activities, extends from the end of mitosis through to the end of G1 phase; the alternate state, which is characterized by high CDK and low APC/C activities, extends from S phase through mitosis (5). Helicase loading can occur only during the low CDK period, while helicase activation can occur only during the high CDK period. Large numbers of origins can have helicase loaded before entry into S phase; during S phase, as loaded helicases become activated, new helicases cannot be loaded, thus ensuring origins cannot fire more than once in a cell cycle. As a consequence, there is no urgency for all origins to fire immediately, and indeed, most cells have a temporal program of origin firing, with origins being activated throughout S phase in a predictable pattern (6).

In addition to being a fascinating and fundamental biological process, DNA replication also has significance for human health: Defects in replication initiation and its regulation are responsible for a series of congenital human diseases (7); moreover, oncogene expression can misregulate origin firing, contributing to genomic instability in cancer (8). Here, we review progress in understanding the mechanisms and regulation of helicase loading and activation in eukaryotes. We

focus on progress in understanding this in the budding yeast *Saccharomyces cerevisiae*, which has been enabled by incisive genetics and driven more recently by advances in single-particle cryo-electron microscopy (cryo-EM) and the reconstitution of helicase loading and activation with purified proteins (9–11).

2. REPLICATION ORIGINS AND THE INITIATOR

2.1. Eukaryotic DNA Replication Origins

The first eukaryotic DNA replication origins isolated and characterized were from the *S. cerevisiae* genome. These autonomously replicating sequences (ARSs) are both the genetic elements required to initiate replication and the biochemical sites where replication begins (12–18). All ARSs contain a T-rich, nonpalindromic ARS consensus sequence (ACS) in an A domain (19), which is essential but not sufficient for origin function (20–22). In a biochemical tour de force, Bell and Stillman (23) identified and purified ORC as a factor that specifically recognizes the ACS. ORC is a six-subunit complex that requires ATP binding but not ATP hydrolysis to recognize and stably bind the ACS. ORC protects the ACS from cleavage by DNase I and generates DNase I hypersensitive sites in adjacent sequences. This diagnostic cleavage signature was used to show ORC is bound at replication origins *in vivo* (24). The genes encoding all six ORC subunits are essential for viability in budding yeast, and hypomorphic *orc* mutants exhibit defects in initiating DNA replication (25–31). Homologs of the yeast ORC subunits are found in all of the main eukaryotic lineages (32), and protein complexes containing human, *Xenopus*, and *Drosophila* homologs of ORC subunits have been characterized (33–36). ORC is required for replication in *Xenopus* egg extracts (36), and ORC subunits are essential for replication in human cells (37), though some cancer cell lines appear to be able to proliferate without Orc1 and Orc2 (38). Taken together, these results show that ORC is the replication initiator protein in budding yeast, a role likely retained across Eukarya.

In addition to the ACS, a flanking sequence called the B domain is required for origin function. B domains vary somewhat in length, but in general, yeast origins are contained within 100–200-bp nucleosome-free regions (39). An early analysis of the ARS1 origin identified three sequence elements, B1, B2, and B3 (**Figure 2a**), within the B domain that are important, but not essential, for origin function (40). The B3 element is the binding site for the general transcription factor Abf1, which functions to position a nucleosome adjacent to the origin (41). The B1 element, which lies next to the ACS, contributes to ORC binding affinity (42, 43) and can thus be considered part of the extended ORC-binding site.

Understanding the function of the B2 element has been important in elucidating the mechanism of helicase loading. The B2 element of ARS1 contains a near match to the ACS in the opposite orientation to the A-domain ACS. ORC can bind to the B2 element, but this binding is greatly enhanced when the A element is mutated, indicating that two ORC molecules cannot bind simultaneously to ARS1 (23). Unbiased approaches have shown that the optimal B2 sequence corresponds to the ACS (44), suggesting that the function of B2 is to bind ORC. DNase I footprinting with purified ORC revealed additional protections and hypersensitive sites beyond the B2 element that are not accounted for by the A/B1 or B2 elements, indicating that there are likely to be additional, weaker ORC-binding sites in the B domain (23, 42). Indeed, the occurrence of multiple near matches to the ACS downstream and in the opposite orientation of the main ACS was a feature of origins noted as early as 1988 (45). More recently, two ORC-binding sites (A/B1) separated by random DNA sequences were shown to be functional as replication origins but only when placed in opposing orientations; functional ARSs were generated with sites anywhere between 25 and 400 bp apart (22), indicating that the distance between sites is less important than

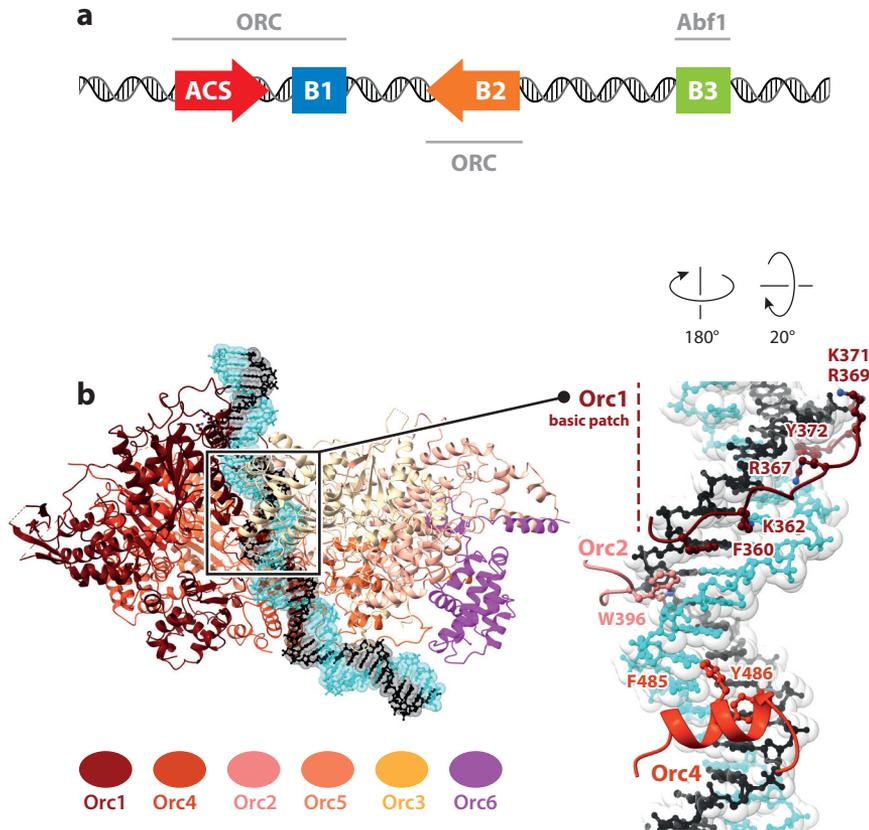


Figure 2

Linear structure of the ARS1 origin of replication and its recognition mechanism. (a) The ACS and B1 elements constitute one binding site for ORC. The B2 element is an inverted ORC-binding site. The B3 element is recognized by the general transcription factor, Abf1. (b) ORC wraps around and bends duplex DNA at the ACS and B1 sites. Base recognition is performed by residues clustered in a basic patch N terminal of the AAA+ domain in Orc1, the ISM element of Orc3, and an insertion in the WHD beta hairpin of Orc4. Abbreviations: AAA+, ATPase associated with various cellular activities; ACS, ARS consensus sequence; ARS, autonomously replicating sequence; ISM, initiator-specific motif; ORC, origin recognition complex; WHD, winged helix domain.

the orientation of the sites. Taken together, these results suggest that one function, perhaps the primary function, of B domains is to provide sites for ORC binding in the opposite orientation to the A/B1 ORC-binding site; the B2 element at ARS1 is an example of such a site.

B domains have additional features that likely contribute to origin function. Because of the asymmetric distribution of adenine and thymine residues in the Watson and Crick strands, they have an inherent tendency to exclude nucleosome occupancy (39). This base composition feature also means that B domains have lower helical stability; that is, they are inherently easier to unwind (46).

At approximately the same time that the first yeast replication origins were being identified, Harland and Laskey (47) showed that a variety of DNA molecules from bacteria, bacteriophages, and animal viruses could replicate efficiently when microinjected into *Xenopus laevis* eggs, indicating that a requirement for sequence-defined replication origins is not a universal feature of

eukaryotic replication. They further showed that these microinjected DNAs replicated just once in a cell cycle (47), demonstrating that specific DNA sequences were also not required for restricting replication to once per cell cycle. As described in Section 3, the system ensuring once per cell cycle replication in eukaryotes, including yeast and metazoans, does not operate at the level of individual origin sequences.

Nonetheless, replication in many metazoan cells begins at defined, specific sites. The methodology for identifying origins and the current view of what defines a metazoan replication origin have been reviewed extensively (48, 49). The only DNA element found thus far at many origins is a G-rich element that does not have a clear consensus sequence but rather shares a propensity for forming a noncanonical four-stranded helical structure called the G quadruplex (G4). In contrast to the budding yeast ORC, human ORC appears to bind DNA and promote initiation in a largely sequence-independent manner (50, 51); however, consistent with the origin mapping data, human ORC has been shown to exhibit preferential binding to G4 ssDNA (52). Metazoan replication origins tend to have decreased nucleosome densities, and a number of local histone modifications have been found to correlate with origin positions and/or ORC binding. Of these, H4K20me2 interacts directly with the bromo-adjacent homology (BAH) domain of Orc1 (53). ORC may also be recruited to specifically modified histones via interactions with the HBO1 acetyltransferase (54) and the LRWD1/ORCA protein (55, 56).

2.2. Mechanism of Origin Recognition by the Origin Recognition Complex

ORC is a hetero-hexameric assembly composed of five homologous proteins (Orc1–5), which form a crescent moon shape, and an Orc6 subunit bound in the periphery. Orc1–5 contain an N-terminal AAA+ (ATPase associated with various cellular activities) module and a C-terminal DNA-binding domain. Several ATPase sites of ORC have become inactivated during evolution, with only Orc1, Orc4, and Orc5 retaining the ability to bind ATP, while hydrolysis occurs only at one bipartite catalytic center (formed between the arginine finger in Orc4 and the ATP-binding site in Orc1, Orc4>Orc1) (57). Orc6 is structurally related to the transcription factor TFIIB, which is composed of a tandem repeat of cyclin-box folds (58, 59) with the C-terminal end docked onto Orc3 (60). Crystallographic and cryo-EM studies on yeast, *Drosophila*, and human complexes showed that DNA binds the central cavity of the Orc1–5 crescent (61–64). Structures of the metazoan complexes show that DNA binding can be impaired by ATPase autoinhibition. In the autoinhibited state, the AAA+ module of Orc1 is positioned away from the Orc4 subunit, which disrupts the Orc4>Orc1 active site and at the same time occludes the DNA-binding cavity of ORC (60, 63). Autoinhibition is released upon ATP binding and the interaction with additional initiators (61).

The structure of *S. cerevisiae* ORC bound to ACS DNA reveals how specific origin sequences are recognized at the origin of replication (62). DNA contacts are mediated by distinct ORC elements, whose structure was previously described for archaeal Orc1 orthologs in complex with origin DNA (65, 66). These elements include an insertion in the AAA+ domain known as the initiator-specific motif (ISM) (with Orc1, 4, 5, and 3 in the eukaryotic complex forming a spiral around DNA), as well as the winged helix domain (WHD). Additionally, in the *S. cerevisiae* ORC–DNA structure, a helical insertion in the Orc4 WHD beta hairpin engages in the recognition of thymine and adenine bases in the ACS (**Figure 2b**), and its mutation alters origin discrimination and firing patterns in cells (67). Other base-recognition elements are seen only in the *S. cerevisiae* ORC structure and engage a thymine-rich stretch in the ACS minor groove, all through hydrogen bonding with thymine O2. One is a tryptophan in the Orc2 ISM, and the second is a patch of basic residues that map onto a loop between an N-terminal BAH appendix and the AAA+ domain

of Orc1 (62). The Orc2 ISM tryptophan and the Orc4 WHD helical insertion exist only in a subgroup of *Saccharomyces*-related budding yeasts and evolved together with origin sequence specificity (67). Conversely, the Orc1 basic patch is universally conserved in eukaryotes.

Nucleosome interaction may also have a role in recruiting ORC onto origins. The crystal structure of the N-terminal domain of Orc1 bound to a nucleosome core particle explains how this engagement occurs, showing a BAH domain that binds H2A and H2B via a canonical acidic patch interaction (68). As described in Section 2.1, nucleosome recognition in higher eukaryotes may involve Orc1 BAH engagement of the H4K20me2 mark (53).

Duplex DNA threaded through the ORC central channel becomes bent and is bound by Orc6 (**Figure 2b**) as it emerges on the C-terminal side of the crescent (61, 62, 64). During the helicase-recruitment process, the Cdc6 protein becomes transiently associated with ORC through a mechanism controlled by ATP binding of Cdc6. To create the space required for Cdc6 recruitment, the WHD of Orc2 disengages from DNA, which in turn leads to the reconstitution of a new Orc1>Cdc6 active site and heptameric complex formation (61, 64). Cdc6 closes the ORC ring, topologically entrapping duplex DNA, while also maintaining a bent DNA conformation, as in the isolated ORC.

3. HELICASE LOADING

To explain how replication from nonspecific start sites could nonetheless be precisely regulated to once per cell cycle in metazoan cells, Harland and Laskey (47, 69) proposed that chromatin is licensed prior to S phase with a factor that binds frequently and nonspecifically to chromatin and is required for replication; this license is removed during DNA replication, so the replicated chromatin is unlicensed and therefore unable to rereplicate. Using *Xenopus* egg extracts, Blow and Laskey (70) showed that permeabilization of nuclei after replication, for example, by detergent treatment, allowed chromatin to be relicensed in egg extracts for another round of replication. This led them to propose that the licensing factor cannot cross the nuclear envelope and can access chromatin only when the nuclear envelope breaks down in mitosis. Exploiting the egg extract replication system, MCM and ORC were identified as essential components of the licensing reaction (36, 71–73).

Genomic footprinting and chromatin immunoprecipitation experiments using synchronized budding yeast cells showed that ORC remains bound to origins throughout the cell cycle (74–76). Genomic footprinting identified an extended protection across the B domain (74). This pre-replicative complex (pre-RC) is present at origins from the end of mitosis until S phase (74) and requires the Cdc6 protein (77). Subsequently, in both *Xenopus* and budding yeast, Cdc6 was found to be required for MCM loading onto chromatin (78, 79), suggesting that licensing in *Xenopus* and pre-RC assembly in yeast are likely to be the same thing. ORC and Cdc6 can be removed from chromatin by salt extraction without removing MCM (79), indicating that these proteins load MCM onto chromatin but, once MCM is loaded, are no longer required to retain MCM on chromatin. The reconstitution of this reaction with purified proteins (9, 10) has provided an approach to understand this loading reaction in detail, discussed in the following three sections.

3.1. Structural Mechanism of First MCM Ring Loading

The six distinct subunits in the MCM ring form two tiers, an N-terminal oligomerization domain and a AAA+ ATPase tier, embellished by C-terminal WHD extensions (80). In the loading-competent state, MCM exists as an open ring (81, 82) containing a discontinuity between Mcm2 and Mcm5, which is used as an entry gate for DNA threading (83). The structure and mechanism of the loading factor Cdt1 appear different in yeast and metazoans. In yeast, Cdt1 is stably bound

to MCM (10) and contains an inactive dioxygenase fold important for loading (81, 82). In higher eukaryotes, Cdt1 does not form a stable complex with MCM and, in place of the dioxygenase fold, it contains an N-terminal intrinsically disordered domain with a role in liquid–liquid phase separation (61, 84).

The middle (M) and C-terminal (C) WHD modules of Cdt1 are universally conserved and their essential role in helicase loading can be explained by the cryo-EM structure of yeast MCM–Cdt1. They bind to the side of MCM (wrapping around Mcm2, 6, and 4) and contribute to keeping the Mcm2–5 gate open, which is required for DNA access to the helicase central channel (83). Biochemical data indicate the Mcm3 WHD plays a key role in the recruitment and loading of MCM onto origin DNA and also triggers ATP hydrolysis by the ORC–Cdc6 complex (85). This is in line with recent cryo-EM studies where the MCM recruitment reaction was slowed by truncating the Mcm6 WHD domain (86). Here, the Mcm3 WHD, together with the Mcm7 WHD, was found engaged with the C-terminal face of the ORC–Cdc6 ring, contacting Orc1, Orc2, and Cdc6, in a so-called engagement complex. Upon binding of Mcm4 by the Orc1 WHD, a preinsertion state is stabilized in which the bent DNA emerging from the ORC–Cdc6 ring becomes aligned with the Mcm2–5 gate (Figure 3). In this configuration, the DNA appears ideally poised for threading into the MCM central channel (86). Robust ORC–Cdc6–Cdt1–MCM (OCCM) formation can be achieved with wild-type proteins by using the slowly hydrolyzable ATP analog ATP γ S (85, 86). Both the preinsertion complex visualized with the Mcm6 truncation mutant and the ATP γ S-stabilized OCCM complex closely resemble structural intermediates observed only in early time points of the ATP-dependent DH loading reaction from time-resolved two-dimensional cryo-EM

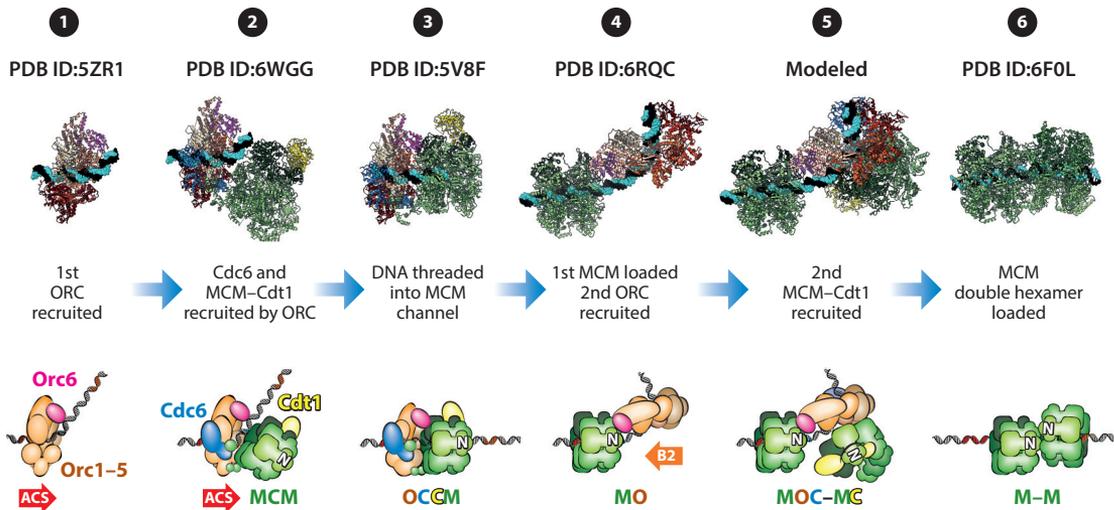


Figure 3

Structural mechanism of MCM DH loading onto origin DNA. Diagrams are based on experimentally determined or modelled atomic structures. Six steps are required for the sequential, quasi-symmetric loading of a DH. ❶ ORC binds and bends the ACS/B1 site. ❷ Cdc6 binds to ORC and recruits a first MCM–Cdt1 complex. In this preinsertion complex, bent DNA is aligned with the Mcm2–5 gate. ❸ DNA is threaded through the Mcm2–5 gate to form the OCCM complex, in which ORC binds the C-terminal side of the MCM ring. ❹ OCCM is disassembled and ORC binds to the N-terminal side of a first loaded MCM ring, forming the MO complex. ❺ Cdc6 engages ORC and recruits MCM–Cdt1, forming a second preinsertion complex. This structural intermediate is modelled by superposing the MO and preinsertion complex, and it recapitulates two-dimensional images obtained by time-resolved cryo-EM imaging of the MCM loading reaction. ❻ A DH of MCM encircles duplex DNA. Abbreviations: ACS, ARS consensus sequence; ARS, autonomously replicating sequence; cryo-EM, cryo-electron microscopy; DH, double hexamer; MCM, minichromosome maintenance; MO, MCM–ORC; OCCM, ORC–Cdc6–Cdt1–MCM; ORC, origin recognition complex; PDB ID, Protein Data Bank identifier.

analysis (59). This observation supports the notion that OCCM is a bona fide loading intermediate (**Figure 3**).

In the OCCM complex, the C-terminal WHD tier of ORC/Cdc6 sits on top of the MCM WHDs, so that the two rings are aligned and both harbor duplex DNA in their central channels. As part of the global reconfiguration from MCM–Cdt1 to OCCM, the C-terminal WHD of the Mcm6 becomes restrained to the side of the MCM through engagement with a linker region of Cdt1 that connects the M and C domains. The Mcm5 WHD also moves away from the pore center facilitating DNA access into the MCM channel (87). The MCM ring transitions from a lock-washer shape to a planar structure as it engages the ORC–Cdc6 complex. At the same time, several ORC–DNA contacts are lost, allowing bent DNA to straighten, pass through the Mcm2–5 gate, and enter into the helicase central channel (86). In the OCCM complex, the MCM ring clamps around duplex DNA; however, the 2–5 gate remains ajar, raising the question of how topological closure of the first MCM ring around duplex DNA is achieved (87). Single-molecule analyses provide important insights into OCCM disassembly, which is required to complete MCM loading. First, Cdc6 is released upon ATP hydrolysis in the Orc1>Cdc6 active site. Then Cdt1 is ejected, and the Mcm2–5 gate is shut around DNA (88, 89). Analysis of ATPase-defective mutants has indicated that ATP hydrolysis by the Orc1>Cdc6 and Orc4>Orc1 sites is not absolutely required for MCM loading; however, ATP hydrolysis by MCM subunits is essential for loading (90, 91). It remains to be established which MCM ATPase sites need to fire and in what order to lock a first loaded MCM ring around duplex DNA. Release of ORC from the C-terminal face of MCM is understood to be promoted by hydrolysis at the Orc4>Orc1 ATPase site (57), and single-molecule data indicate that this is important for the completion of DH loading onto origin DNA (88). ATP hydrolysis by Orc1>Cdc6 is also important for disassembling aberrant or incomplete intermediates and has been proposed to act as a licensing proofreader (85, 90, 91).

3.2. Structural Mechanism of Second MCM Hexamer Loading

In vitro reconstitution using mutant MCM variants established that the loading of the first and second MCM rings employ the same Mcm3 WHD element, indicating that both helicases are loaded via the OCCM mechanism (92). Two ORC binding events, to inverted sites on origin DNA, were found to be required for efficient DH formation to occur (22). This suggested that the orientation of ORC binding to DNA is important for the engagement of the N-terminal domains of two MCM rings that leads to DH formation (22). Single-molecule spectroscopy studies indicated that helicase loading is sequential, suggesting that the first loaded MCM ring plays a role in the recruitment of the second ring (88, 89). Recent time-resolved cryo-EM experiments explained how sequential helicase loading occurs. When imaging the MCM loading reaction reconstituted on the ARS1 origin of replication, a new loading intermediate was visualized, containing MCM bound to ORC (MO) on DNA, after the release of Cdt1 and Cdc6 that comes with OCCM disassembly (**Figure 3**). Single-particle reconstruction showed that the N-terminal domain of Orc6 bridges across the closed Mcm2–5 gate, which can occur only when ATP is hydrolyzed (88). Therefore, MO contains an ATP-hydrolysis-dependent, fully loaded single MCM ring encircling duplex DNA (59). Single-loaded rings were frequently seen in single-molecule experiments (93), suggesting they may not be uncommon. The interaction between MCM and ORC in MO is completely different from that of OCCM, as ORC no longer engages the C-terminal WHD elements in MCM. Rather, it contacts the N-terminal tier and binds DNA in an inverted configuration, which supports the formation of a symmetric DH (22).

While fluorescence spectroscopy experiments established that two Cdt1 and two Cdc6 molecules are involved in loading a DH onto origins of replication, the data support a mechanism whereby one ORC molecule could load both the first and the second MCM rings (89). Cryo-EM

experiments describe two distinct and inverted binding sites for ORC on MCM before loading of the first MCM ring (C-terminal MCM-interacting in OCCM) and after loading (N-terminal MCM-interacting in MO) (**Figure 3**). As ORC employs distinct functional elements for OCCM and MO formation, one could envisage that ORC swaps its position on the same MCM ring, while never losing contacts with the single-loaded helicase particle. However, single-particle analysis identified a molecular species in which one ORC particle engages N-terminal MCM while a second ORC particle engages the C-terminal end, indicating that helicase loading can occur using two distinct ORC molecules (59). This scenario is consistent with ORC titration experiments showing that the increase in MCM loading efficiency at increasing ORC concentrations is not linear (22). It is thus likely that DH formation can be accomplished by either one or two ORC molecules (22, 89).

In the fully loaded MCM DH, the two helicases become interlocked via six B domains (five of which are zinc fingers) in the N-terminal module, as well as a helical extension in Mcm7 that engages Mcm5 in *trans* (80) (**Figure 3**). All MCM subunits contact DNA, via the zinc finger elements and a set of beta hairpins that project inside the MCM central channel in the N-terminal module, as well as a helix-2 insertion and a presensor 1 beta hairpin that embellish the AAA+ module (94, 95). The two rings are slightly offset in this complex, resulting in a cockeyed configuration that causes a constriction in the central channel, at the interface between the two MCM rings. Duplex DNA can snake through this constriction, becoming bent but not melted (94, 95). Both DNA strands in the double helix are touched by MCM subunits within each helicase ring. These observations provide a rationale for why only passive sliding and not unidirectional translocation occurs in the loaded MCM DH (9, 10). Helicase activation that occurs upon CMG formation in S phase is required to achieve DNA melting, translocation strand selection, and replication-fork establishment.

3.3. Regulation of Helicase Loading

The timing of pre-RC appearance (74) correlates closely with the destruction of the mitotic cyclin Clb2 and expression of the CDK inhibitor Sic1, which suggested CDK may be an inhibitor of pre-RC assembly. This was demonstrated by showing that inactivation of CDK in G2 could drive the inappropriate assembly of pre-RCs, even in the absence of functional APC/C or other mitotic regulators (96, 97).

CDK directly inhibits helicase loading by phosphorylating ORC, Cdc6, and MCM. CDK phosphorylation of degron sequences in Cdc6 targets it for degradation via the Skp1–cullin–F box protein (SCF)^{Cdc4} E3 ubiquitin ligase (98–102). CDK phosphorylation of a nuclear import-export cassette shared between Mcm2 and 3 causes relocation of the MCM complex, along with associated Cdt1, from the nucleus to the cytoplasm (103–107). CDK phosphorylates the Orc2 and 6 subunits of ORC and directly inhibits ORC's ability to load MCM in vitro (85, 108) (**Figure 1**). Moreover, Clb5–CDK can bind to a short sequence motif [Arg-X-Leu (RXL)] in Orc6, sterically interfering with MCM loading (109). Cdc6 degradation and MCM nuclear exclusion can be triggered by the late G1 cyclins (Cln1–3) (99, 103) as well as the S phase and mitotic cyclins (Clb1–6). Cln1–3, however, cannot directly trigger helicase activation; that is done by the downstream S phase cyclins, Clb5 and 6. This generates a mechanism analogous to an airlock, in which the cell shuts down helicase loading before, not concomitant with, helicase activation, ensuring that relicensing of fired origins cannot accidentally occur during the transition between cell states.

Only when all three inhibitory pathways (Cdc6, MCM, and ORC phosphorylation) were bypassed in G2 arrested cells was substantial rereplication detected by flow cytometry (110), indicating that these three inhibitory pathways are at least partially overlapping. The mechanisms

involved in preventing MCM loading outside of G1 phase evolve rapidly (111). This is partly because the loss of any single pathway is not lethal and partly because regulatory mechanisms can be swapped between pre-RC components without compromising the block to rereplication (111).

As a consequence of this rapid evolution, the block to MCM loading is quite different in metazoans. Firstly, the APC/C plays a direct role by targeting the Cdt1 inhibitor Geminin for degradation (112), and secondly, Cdt1 is targeted for degradation during S phase by interaction with DNA-bound proliferating cell nuclear antigen (PCNA) (113, 114). CDK also directly inhibits pre-RC components; for example, CDK targets ORC1 for degradation during S phase (115). Recently it has been shown that *Drosophila* ORC, Cdc6, and as previously mentioned, Cdt1 have intrinsically disordered domains that drive liquid–liquid phase separation, and CDK phosphorylation of Orc1 and Cdt1 prevents this (84).

4. HELICASE ACTIVATION

4.1. Overview of Helicase Activation

Many of the factors required for helicase activation (firing factors) were identified by Araki (116) and colleagues, beginning with the gene encoding the largest subunit of DNA polymerase ϵ (Pol ϵ), *POL2*. Pol ϵ is the leading strand DNA polymerase (117) but is also essential for CMG formation after MCM loading (11). The elegant genetic screens used to discover the remaining firing factors, Sld3–7, Sld2, DNA polymerase B 11 (Dpb11), GINS, and Mcm10, have been described previously (116). In addition to these components, two protein kinases—CDK and Dbf4 dependent kinase (DDK)—play direct, essential roles in helicase activation (118).

Helicase activation begins with the stepwise assembly of a preinitiation complex (pre-IC) (119) that includes the firing factors described in the previous paragraph (**Figure 4**). CDK and DDK are inactive during G1 phase because they each have a regulatory subunit (cyclin and Dbf4, respectively) that is targeted for degradation by APC/C. When APC/C is inactivated at the end of G1 phase, these regulatory subunits can accumulate and activate their respective kinases. The substrates for DDK are the Mcm6 and Mcm4 subunits of the MCM DH. Mutation of DDK phosphorylation sites to alanine in these subunits greatly reduces DNA replication (120, 121), and there are numerous mutants in MCM subunits, including phosphomimicking mutants, that can bypass the requirement for DDK in vivo (120–123). DDK can efficiently phosphorylate MCM, but only when it is loaded into DHs on DNA (124), and phosphorylation of MCM alone is necessary to activate the helicase with purified proteins (11). MCM phosphorylation by DDK generates binding sites for the Sld3 protein (120); Sld3, in complex with Sld7, then recruits Cdc45 to the MCM DH (125). The substrates for CDK are Sld2 and Sld3 (126–128). Sld2 contains multiple CDK phosphorylation sites; most of these sites are required to promote CDK phosphorylation of a single essential site, Thr84 (127, 129). Two CDK sites in Sld3, Thr600, and Ser622, are essential for replication in yeast (126, 128) and are conserved in the human ortholog of Sld3, Treslin/TICRR (130–134). Phosphorylation of Sld2 and Sld3 is necessary and sufficient to activate the MCM helicase with purified proteins (11). The phosphorylation of Sld2 and Sld3 generates binding sites for pairs of BRCA1 C-terminal (BRCT) repeats in Dpb11 (TopBP1 in humans); BRCT repeats 1 and 2 of Dpb11 bind phospho-Sld3, and repeats 3 and 4 bind phospho-Sld2. These interactions can be rendered constitutive by phosphomimicking mutants (Sld2) (126, 128) and direct protein fusion (Sld3–Dpb11) (128), and when these bypass mutants are combined, DNA replication can be induced in the absence of CDK activity (126, 128). Thus, phosphorylation of MCM by DDK recruits Sld3/7 and Cdc45; CDK phosphorylation of Sld2 and 3 then recruits Sld2 to the MCM DH via Dpb11. Sld2 in turn recruits Pol ϵ and GINS (135).

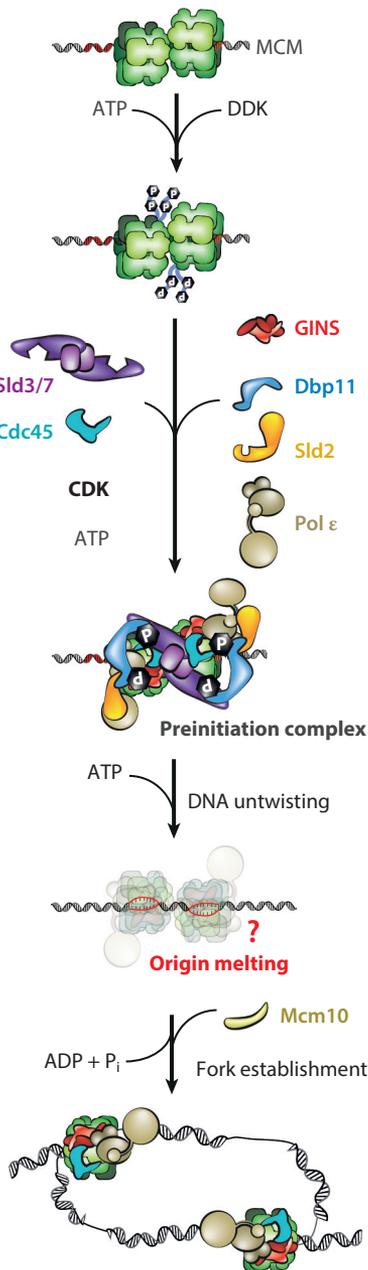


Figure 4

Helicase activation. DNA-loaded DHs are phosphorylated by DDK. Phospho-Mcm4 and -Mcm6 are recognized by Sld3, which exists in complex with homo-dimeric Sld7 and recruits Cdc45 onto MCM. CDK targets Sld3 as well as Sld2. Phospho-Sld2 and phospho-Sld3 bind to Dpb11, which also binds Pol ϵ and GINS. Sld2, Dpb11, GINS, Pol ϵ , Sld3/7, and Cdc45 binding to a DH forms the preinitiation complex. Release of ADP and binding of ATP leads to stable CMG formation, concomitant with DH interface disruption and origin DNA untwisting. Addition of Mcm10 switches on ATPase-powered DNA unwinding by MCM, causing two CMG particles to cross their paths, which establishes bidirectional replication forks. Abbreviations: CDK, cyclin dependent kinase; CMG, Cdc45–MCM–GINS; DDK, Dbf4 dependent kinase; DH, double hexamer; GINS, go-ichi-nii-san; Mcm, minichromosome maintenance; Pol ϵ , DNA polymerase ϵ .

Several of the firing factors have been shown to be present in cells at levels well below the number of origins used in S phase, and overexpression of subsets of these firing factors can accelerate initiation from late-firing origins (125, 136). This suggests a model for replication timing in which sequential action of limiting firing factors, driven by some underlying difference in origin context, results in an ordered firing of origins during S phase (137). Late origin firing is inhibited in response to DNA replication stress or DNA damage by the DNA damage checkpoint kinase Rad53 (138, 139), and this occurs via phosphorylation and inactivation of Sld3 and Dbf4 (140, 141).

Generation of this pre-IC has an important effect on the MCM DH: It triggers release of the ADP that was generated during MCM loading and remained stably bound to MCM until this point (142). This allows MCM to bind ATP; it is ATP binding, without hydrolysis, that drives the next transformation. In this step, a number of things appear to happen concomitantly: The DH separates into individual hexamers, Cdc45 and GINS become stably bound in a CMG complex, and the first DNA untwisting of 6–7 bp per CMG is seen (142). Because this is a relatively short region of untwisting, it is likely that the lagging strand template has not yet been evicted from the MCM central channel (142) (**Figure 4**). In the final step, an additional protein, Mcm10, together with ATP hydrolysis, converts this inactive CMG into an active helicase, translocating in a 3′-to-5′ direction on the leading-strand template with the N-terminal domain of MCM at the front of the helicase, implying that the two CMGs must pass each other by translocating on opposing strands in the origin (142, 143). DNA sequence appears to play little or no role in the helicase activation step since DHs pushed away from their position of loading by RNA polymerase are still competent for firing (144).

MCM subunits, along with Cdc45, are required for both the initiation and elongation phases of DNA replication (145–147) and MCM, along with Cdc45, moves with replication forks during elongation (75). MCM exhibits weak DNA helicase activity (148, 149), but the CMG has robust helicase activity (150–152) and forms the core of the replisome at replication forks (153). Cdc45 is one of the firing factors limiting for origin firing (125), so its recycling after replication termination may play some role in coordinating the completion of replication from early-firing origins with the activation of late-firing origins.

4.2. Structural Mechanism of Helicase Activation and Replisome Assembly

A structural understanding of the origin firing process remains to be achieved. Several firing factors transiently bind to the DNA-loaded MCM DH in the pre-IC on the path to CMG assembly. It is unclear whether these factors engage individual MCM rings in the DH or rather bind across the MCM DH structure, recognizing the symmetry of loaded helicases poised to support bidirectional replication. As discussed in Section 4.1, in the cascade of molecular events leading to helicase activation, the first step involves DDK phosphorylation of DNA-loaded MCM (124, 154). DDK docks onto Mcm2 and phosphorylates Mcm4 and Mcm6 (155). Whether docking and phosphorylation occur on the same ring or across the DH remains to be established. It is also unclear whether or not conversion of one DH into two CMGs is cooperative, with the binding of firing factors to one MCM ring promoting the recruitment of firing factors on the second ring of the DH. Crystallographic characterization of the Sld3/7 complex, which recruits Cdc45 to the DNA-loaded DH, provides tantalizing evidence in support of a cooperative mechanism. In particular, the structure of the Sld3/7 oligomerization domain reveals that Sld7 forms a homodimer, which would place two copies of Sld3 in a position suitable to read Mcm4/6 phosphorylation symmetrically in both helicase rings across the MCM DH (156). Single-molecule work shows that Sld3-mediated recruitment of Cdc45 is sequential (157), arguing against a purely symmetric mechanism for the simultaneous recruitment of two sets of helicase activators by symmetric firing factors. Likewise,

the structural mechanism for the recruitment of GINS to loaded MCMs remains to be explored. An endogenous complex containing phospho-Sld2, Dpb11, Pol ϵ , and GINS has been purified from yeast cells (135). It is unknown whether this complex contains any homo-dimerization domain like Sld3/7, which would support a symmetric mechanism for the recruitment of two GINS molecules.

Mutations that destabilize the MCM DHs still support CMG formation and initial DNA unwinding but block DNA unwinding and origin firing (158). These data invite a model whereby CMG formation occurs independently on two loaded MCM rings. On the path to CMG conversion, the firing factors that recruit Cdc45 and GINS to the MCM might still serve a bridging role between the two DNA-loaded helicases, even when the DH interface is disrupted. Despite their differences, one unifying theme of all proposed helicase activation mechanisms is that building two CMGs from one symmetric DH is required to establish bidirectional replication.

The structural analysis of the *in vitro* conversion of MCM DHs to CMGs at origins of replication is limited to low-resolution negative stain electron microscopy (142). This is due to the complexity and relatively low efficiency of CMG formation *in vitro*, which poses exceptional challenges to high-resolution imaging. However, overexpression approaches led to describing the three-dimensional structure of the CMG using electron microscopy at increasingly higher resolution. A first inspection of the CMG structure immediately suggests a mechanism whereby GINS and Cdc45 recruitment would cause disruption of the MCM DH interface. GINS and Cdc45 latch across the side of the helicase ring, engaging the Mcm3, 5, and 2 subunits and connecting the N-terminal and ATPase tiers (150, 159, 160). The GINS subunit Psf2, in particular, engages the A domain of Mcm5. This contact is incompatible with an interaction that stabilizes MCM dimerization in the DH, involving a helical extension of Mcm7 that engages the A domain of Mcm5 in *trans* (80). This Mcm7 element is shorter in Metazoa and might not play the same DH-stabilization function, suggesting that some differences in the structural mechanism of helicase activation might exist in yeast compared to those in higher eukaryotes.

Co-overexpression of the 11 members of the CMG complex has also allowed a structural and functional understanding of the replicative helicase at the replication fork. Cryo-EM structures of yeast, *Drosophila*, and human assemblies all indicate that the CMG engages a DNA fork substrate by encircling ssDNA (143, 159, 161–164), and biochemical data show that unwinding involves translocation along the leading-strand template via a steric exclusion mechanism (151, 165–168). Here, the lagging-strand template is kept outside of the helicase central channel, most efficiently in the presence of the ssDNA-binding protein replication protein A (RPA), which stimulates unwinding (167). Several findings support the notion that a replisome reconstituted by combining the overexpressed CMG and Pol ϵ complexes on an artificial DNA fork has the same structure and function as a replisome built *in vitro* by recapitulating the origin activation cellular pathway (169–171). First, the same Pol ϵ elements that are essential for (a) cell viability (172, 173), (b) GINS recruitment to the MCM (174), and (c) DNA replication *in vitro* have been shown to bridge MCM and GINS in the cryo-EM structure (164). These findings provide insights into the noncatalytic role of Pol ϵ as a firing factor that is essential for helicase activation. Second, in both systems, the directionality of DNA translocation positions the N-terminal MCM domain in the CMG at the leading edge of the advancing helicase. As introduced before, this implies that two separated CMG particles originating from the same DH must cross their paths to achieve replication-fork establishment (142, 143, 161). During this process, the ejected strand of one ring becomes the translocation strand of the opposite ring, ensuring that origin firing occurs only if both MCM helicases transition from a duplex-DNA to a ssDNA interaction. Third, Mcm10 stimulates DNA unwinding in both systems (142, 165, 175). Mcm10 facilitates the transition between a duplex DNA-interacting to a single-strand-interacting form of the CMG, which may

itself stimulate ATP hydrolysis (142, 175, 176). Origin firing by Mcm10 therefore requires the ejection of the lagging strand template from the MCM central channel; however, it is unknown what subunit interface in the MCM ring might function as a DNA gate for ssDNA eviction during replication-fork establishment. Two observations suggest that the DNA-ejection gate might not be the same gate used for helicase loading (Mcm2–5). First, in the cryo-EM structure of CMG–Pol ϵ , which is formed before lagging-strand exclusion, the Mcm2–5 gate appears securely shut, with GINS engaging Mcm5 and Cdc45 plugging the Mcm2/Mcm5 opening in the N-terminal half of the helicase ring (150). On the C-terminal half, the Pol ϵ subunits Pol2 and Dpb2 straddle the Mcm2 and Mcm5 ATPase domains, respectively. Collectively, GINS, Cdc45, and Pol ϵ binding to MCM seals off both ends of the Mcm2–5 DNA-entry gate, making it a less likely candidate for the lagging-strand-ejection gate (164). Second, a CMG variant containing a rapamycin-inducible linkage between Mcm2 and Mcm5 efficiently transitions from a duplex-interacting to a ssDNA-interacting state, indicating that one or more exit routes for lagging strand ejection may exist that differ from the Mcm2–5 gate (176). How Mcm10 acts at the molecular level to promote lagging strand ejection and stimulate DNA unwinding still needs to be established. It is, for example, unknown whether a transition from duplex- to ssDNA-interacting helicases occurs before two helicases meet, allowing both CMGs to cross their paths and establishing no new physical interaction. Alternatively, lagging strand ejection could occur in the moment when the two helicases translocate past one another.

The structural mechanism of DNA unwinding after replication-fork establishment is starting to be elucidated by cryo-EM imaging of the actively translocating CMG, or a supercomplex that also includes chromosome transmission fidelity 4 (Ctf4) and the fork-stabilization complex, Csm3–Tof1–Mrc1 (163) that stimulates DNA unwinding and replication-fork progression (177, 178). Duplex DNA is split by a set of beta hairpin loops that project into the helicase central channel from the N-terminal oligonucleotide/oligosaccharide-binding fold oligomerization domain of MCM. Here, a conserved Mcm7 Phe363 element engages in π – π interactions with a paired base at the fork nexus, and the beta hairpins of Mcm3 divert the lagging strand toward a gap between the Mcm3 and Mcm5 B domains (**Figure 5**) (161–163). Different CMG conformers observed in both the yeast and *Drosophila* helicases provide insights into how ATP binding in different AAA+ active sites around the MCM ring stabilizes different modes of DNA engagement in the ATPase tier in a manner that suggests N terminus-to-C terminus, 3'-to-5' movement (161, 163). This translocation would involve substrate rotation within the inner perimeter of the MCM ring. The same rotary cycling mechanism has been proposed for AAA+ protein translocases (179), as well as viral bacterial and archaeal helicases that have the opposite polarity of DNA unwinding, suggesting a conserved mechanism of movement along peptide and nucleic acid polymers for hexameric ATPases (180–183).

The ssDNA exposed through the unwinding action of the CMG serves as a template for priming by Pol α and lagging and leading DNA synthesis by Pol δ and ϵ (184). An RNA primer can be synthesized by Pol α on the lagging-, but not the leading-, strand template, meaning that the orientation of Pol α relative to the CMG helicase is critical in this process. Pol α can be linked to the CMG via the replisome-organizing factor Ctf4, a homo-trimeric complex that is capable of concomitantly binding core replisome components, as well as other client proteins, via conserved recognition motifs (**Figure 5**) (185–188). When engaged with Ctf4, Pol α is positioned at the front of the advancing replisome (188) and in proximity to the Mcm3–5 interface, through which the lagging-strand tail is channeled at the DNA-fork nexus (162, 163). This architecture suggests a simple structural mechanism for lagging-strand priming. Ctf4 links the CMG to up to two copies of Pol α (185, 188), which recapitulates stoichiometries observed in single-molecule studies on the replisome during DNA synthesis (189). Alternatively, Ctf4 has been observed to bridge two CMG

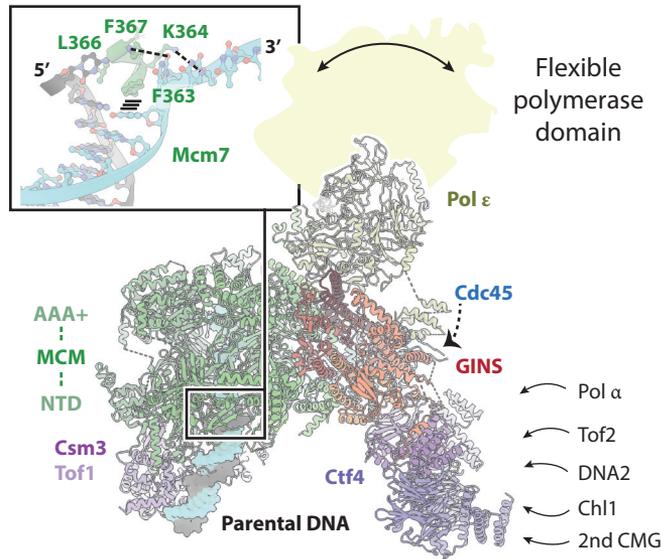


Figure 5

Composite atomic model that combines the CMG complex bound to DNA, Csm3–Tof1, and Ctf4 (PDB ID: 6SKL) and a structure of CMG bound to DNA and Pol ϵ (PDB ID: 6HV9). GINS and Cdc45 bind to the side of an MCM ring, here viewed from the side. Parental duplex DNA enters the N-terminal face of the MCM ring, where Csm3–Tof1 and Ctf4 are located. Ctf4 forms a homo-trimeric disk and functions as a molecular hub onto which multiple factors can dock. Pol ϵ engages the AAA+ face of the MCM ring and is poised to capture the leading-strand template as it exits from the helicase central channel. Abbreviations: AAA+, ATPase associated with various cellular activities; CMG, Cdc45–MCM–GINS; Ctf, chromosome transmission fidelity; GINS, go-ichi-nii-san; MCM, minichromosome maintenance; NTD, N-terminal domain; PDB ID, Protein Data Bank identifier; Pol, DNA polymerase.

helicases, which led to the proposal that diverging helicases might remain linked during bidirectional replication (188). These tantalizing observations predict a key role for Ctf4 in replisome function; however, DNA replication establishment and fork progression rates *in vitro* are unaffected by the presence or absence of Ctf4 (169), and single-molecule experiments have shown that interactions between divergent replisomes are not required for replication in *Xenopus* egg extracts (190). Thus, the main role of Ctf4 must be linking the replisome with factors that play genome maintenance functions distinct from DNA replication (**Figure 5**). In fact, deletion of Ctf4 shows a strong sister chromatid cohesion defect (191), and cellular studies implicated Ctf4, together with noncatalytic elements in Mcm2 and Pol α , in recycling of parental histones onto the duplicated lagging strand (192). These are frontier topics in chromosome biology that fall outside the scope of this article.

The crystal structure of the catalytic domain of Pol ϵ trapped on a DNA substrate (193) and the cryo-EM structure of the ternary complex of Pol δ bound to DNA and PCNA (194) explain the structural mechanism of processive DNA synthesis by the leading- and lagging-strand polymerases. However, how these functions are coordinated at the fork remains largely unknown. Biochemical reconstitution indicates that Pol δ plays an important role during the establishment of leading strand synthesis, suggesting that substrate handoff between Pol δ and Pol ϵ must occur before processive DNA replication is established (178). Structural work suggests that, while Pol ϵ is anchored to the CMG, the catalytic domain of Pol2 is free to move with respect to the replisome core (195), visiting configurations that appear distinct from the rigid structure recently described

for full length Pol ϵ (196). Less is known about Pol δ incorporation into the replisome, although single-molecule studies indicate that Pol δ is recycled for multiple rounds of Okazaki fragment synthesis and is tethered to the replisome directly through an interaction of the Pol32 subunit with the catalytic subunit of Pol α (189, 197).

5. CONCLUSIONS

Our understanding of helicase loading and activation in budding yeast, summarized in this review, has proceeded at a breathtaking rate in recent years, and a relatively complete understanding of these processes is clearly on the horizon. The proteins involved in initiating replication are, for the most part, highly conserved across Eukarya. Understanding whether the mechanisms described in this review are also conserved requires further work. Biochemical systems using proteins from other organisms, including humans, will play a major part in resolving this. We are also just beginning to appreciate how misregulation of these processes may affect genome stability and contribute to human diseases including cancer. Interactions between biochemists and cancer biologists will hopefully help drive a deeper understanding of these connections.

DISCLOSURE STATEMENT

The authors are not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.

ACKNOWLEDGMENTS

We are grateful to Joe Brock for help with illustrations. Work in the authors' laboratories is funded jointly by the Wellcome Trust, the Medical Research Council, and Cancer Research UK at the Francis Crick Institute (FC001065 to A.C. and FC001066 to J.F.X.D.). A.C. receives funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (grant agreement no. 820102). This work was also funded by a Wellcome Trust Senior Investigator Award (106252/Z/14/Z) and an ERC Advanced Grant (669424-CHROMOREP) to J.F.X.D.

LITERATURE CITED

1. Mott ML, Berger JM. 2007. DNA replication initiation: mechanisms and regulation in bacteria. *Nat. Rev. Microbiol.* 5:343–54
2. Nasir A, Mughal F, Caetano-Anollés G. 2021. The tree of life describes a tripartite cellular world. *BioEssays* 43:e2000343
3. Greci MD, Bell SD. 2020. Archaeal DNA replication. *Annu. Rev. Microbiol.* 74:65–80
4. Singleton MR, Dillingham MS, Wigley DB. 2007. Structure and mechanism of helicases and nucleic acid translocases. *Annu. Rev. Biochem.* 76:23–50
5. Kapuy O, He E, López-Avilés S, Uhlmann F, Tyson JJ, Novák B. 2009. System-level feedbacks control cell cycle progression. *FEBS Lett.* 583:3992–98
6. Vouzas AE, Gilbert DM. 2021. Mammalian DNA replication timing. *Cold Spring Harb. Perspect. Biol.* In press. <https://doi.org/10.1101/cshperspect.a040162>
7. Schmit M, Bielinsky AK. 2021. Congenital diseases of DNA replication: clinical phenotypes and molecular mechanisms. *Int. J. Mol. Sci.* 22:911
8. Hills SA, Diffley JFX. 2014. DNA replication and oncogene-induced replicative stress. *Curr. Biol.* 24:R435–44
9. Evrin C, Clarke P, Zech J, Lurz R, Sun J, et al. 2009. A double-hexameric MCM2–7 complex is loaded onto origin DNA during licensing of eukaryotic DNA replication. *PNAS* 106:20240–45

10. Remus D, Beuron F, Tolun G, Griffith JD, Morris EP, Diffley JFX. 2009. Concerted loading of Mcm2–7 double hexamers around DNA during DNA replication origin licensing. *Cell* 139:719–30
11. Yeeles JT, Deegan TD, Janska A, Early A, Diffley JFX. 2015. Regulated eukaryotic DNA replication origin firing with purified proteins. *Nature* 519:431–35
12. Stinchcomb DT, Struhl K, Davis RW. 1979. Isolation and characterisation of a yeast chromosomal replicator. *Nature* 282:39–43
13. Hsiao CL, Carbon J. 1979. High-frequency transformation of yeast by plasmids containing the cloned yeast *ARG4* gene. *PNAS* 76:3829–33
14. Struhl K, Stinchcomb DT, Scherer S, Davis RW. 1979. High-frequency transformation of yeast: autonomous replication of hybrid DNA molecules. *PNAS* 76:1035–39
15. Tschumper G, Carbon J. 1980. Sequence of a yeast DNA fragment containing a chromosomal replicator and the *TRP1* gene. *Gene* 10:157–66
16. Brewer BJ, Fangman WL. 1987. The localization of replication origins on ARS plasmids in *S. cerevisiae*. *Cell* 51:463–71
17. Huberman JA, Spotila LD, Nawotka KA, El Assouli SM, Davis LR. 1987. The in vivo replication origin of the yeast 2 micron plasmid. *Cell* 51:473–81
18. Saffer LD, Miller OL Jr. 1986. Electron microscopic study of *Saccharomyces cerevisiae* rDNA chromatin replication. *Mol. Cell. Biol.* 6:1148–57
19. Chang F, Theis JF, Miller J, Nieduszynski CA, Newlon CS, Weinreich M. 2008. Analysis of chromosome III replicators reveals an unusual structure for the ARS318 silencer origin and a conserved WTW sequence within the origin recognition complex binding site. *Mol. Cell. Biol.* 28:5071–81
20. Kearsey S. 1984. Structural requirements for the function of a yeast chromosomal replicator. *Cell* 37:299–307
21. Srien F, Bailey JE, Campbell JL. 1985. Effect of *ARS1* mutations on chromosome stability in *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* 5:1676–84
22. Coster G, Diffley JFX. 2017. Bidirectional eukaryotic DNA replication is established by quasi-symmetrical helicase loading. *Science* 357:314–18
23. Bell SP, Stillman B. 1992. ATP-dependent recognition of eukaryotic origins of DNA replication by a multiprotein complex. *Nature* 357:128–34
24. Diffley JFX, Cocker JH. 1992. Protein-DNA interactions at a yeast replication origin. *Nature* 357:169–72
25. Foss M, McNally FJ, Laurenson P, Rine J. 1993. Origin recognition complex (ORC) in transcriptional silencing and DNA replication in *S. cerevisiae*. *Science* 262:1838–44
26. Bell SP, Kobayashi R, Stillman B. 1993. Yeast origin recognition complex functions in transcription silencing and DNA replication. *Science* 262:1844–49
27. Bell SP, Mitchell J, Leber J, Kobayashi R, Stillman B. 1995. The multidomain structure of Orc1p reveals similarity to regulators of DNA replication and transcriptional silencing. *Cell* 83:563–68
28. Loo S, Fox CA, Rine J, Kobayashi R, Stillman B, Bell S. 1995. The origin recognition complex in silencing, cell-cycle progression, and DNA replication. *Mol. Biol. Cell* 6:741–56
29. Micklem G, Rowley A, Harwood J, Nasmyth K, Diffley JFX. 1993. Yeast origin recognition complex is involved in DNA replication and transcriptional silencing. *Nature* 366:87–89
30. Li JJ, Herskowitz I. 1993. Isolation of ORC6, a component of the yeast origin recognition complex by a one-hybrid system. *Science* 262:1870–74
31. Hardy C. 1996. Characterisation of an essential Orc2p-associated factor that plays a role in DNA replication. *Mol. Cell. Biol.* 16:1832–41
32. Ocana-Pallares E, Vergara Z, Desvoyes B, Tejada-Jimenez M, Romero-Jurado A, et al. 2020. Origin recognition complex (ORC) evolution is influenced by global gene duplication/loss patterns in eukaryotic genomes. *Genome Biol. Evol.* 12:3878–89
33. Siddiqui K, Stillman B. 2007. ATP-dependent assembly of the human origin recognition complex. *J. Biol. Chem.* 282:32370–83
34. Vashee S, Simancek P, Challberg MD, Kelly TJ. 2001. Assembly of the human origin recognition complex. *J. Biol. Chem.* 276:26666–73

35. Chesnokov I, Gossen M, Remus D, Botchan M. 1999. Assembly of functionally active *Drosophila* origin recognition complex from recombinant proteins. *Genes Dev.* 13:1289–96
36. Rowles A, Chong JP, Brown L, Howell M, Evan GI, Blow JJ. 1996. Interaction between the origin recognition complex and the replication licensing system in *Xenopus*. *Cell* 87:287–96
37. Chou HC, Bhalla K, Demerdesh OE, Klingbeil O, Hanington K, et al. 2021. The human origin recognition complex is essential for pre-RC assembly, mitosis, and maintenance of nuclear structure. *eLife* 10:e61797
38. Shibata E, Kiran M, Shibata Y, Singh S, Kiran S, Dutta A. 2016. Two subunits of human ORC are dispensable for DNA replication and proliferation. *eLife* 5:e19084
39. Eaton ML, Galani K, Kang S, Bell SP, MacAlpine DM. 2010. Conserved nucleosome positioning defines replication origins. *Genes Dev.* 24:748–53
40. Marahrens Y, Stillman B. 1992. A yeast chromosomal origin of DNA replication defined by multiple functional elements. *Science* 255:817–23
41. Lipford JR, Bell SP. 2001. Nucleosomes positioned by ORC facilitate the initiation of DNA replication. *Mol. Cell* 7:21–30
42. Rowley A, Cocker JH, Harwood J, Diffley JFX. 1995. Initiation complex assembly at budding yeast replication origins begins with the recognition of a bipartite sequence by limiting amounts of the initiator, ORC. *EMBO J.* 14:2631–41
43. Rao H, Stillman B. 1995. The origin recognition complex interacts with a bipartite DNA binding site within yeast replicators. *PNAS* 92:2224–28
44. Wilmes GM, Bell SP. 2002. The B2 element of the *Saccharomyces cerevisiae* *ARS1* origin of replication requires specific sequences to facilitate pre-RC formation. *PNAS* 99:101–6
45. Palzkill TG, Newlon CS. 1988. A yeast replication origin consists of multiple copies of a small conserved sequence. *Cell* 53:441–50
46. Umek RM, Kowalski D. 1988. The ease of DNA unwinding as a determinant of initiation at yeast replication origins. *Cell* 52:559–67
47. Harland RM, Laskey RA. 1980. Regulated replication of DNA microinjected into eggs of *Xenopus laevis*. *Cell* 21:761–71
48. Prioleau MN, MacAlpine DM. 2016. DNA replication origins—where do we begin? *Genes Dev.* 30:1683–97
49. Mechali M, Yoshida K, Coulombe P, Pasero P. 2013. Genetic and epigenetic determinants of DNA replication origins, position and activation. *Curr. Opin. Genet. Dev.* 23:124–31
50. Vashee S, Cvetic C, Lu W, Simanek P, Kelly TJ, Walter JC. 2003. Sequence-independent DNA binding and replication initiation by the human origin recognition complex. *Genes Dev.* 17:1894–908
51. Remus D, Beall EL, Botchan MR. 2004. DNA topology, not DNA sequence, is a critical determinant for *Drosophila* ORC–DNA binding. *EMBO J.* 23:897–907
52. Hoshina S, Yura K, Teranishi H, Kiyasu N, Tominaga A, et al. 2013. Human origin recognition complex binds preferentially to G-quadruplex-preferable RNA and single-stranded DNA. *J. Biol. Chem.* 288:30161–71
53. Kuo AJ, Song J, Cheung P, Ishibe-Murakami S, Yamazoe S, et al. 2012. The BAH domain of ORC1 links H4K20me2 to DNA replication licensing and Meier-Gorlin syndrome. *Nature* 484:115–19
54. Iizuka M, Stillman B. 1999. Histone acetyltransferase HBO1 interacts with the ORC1 subunit of the human initiator protein. *J. Biol. Chem.* 274:23027–34
55. Shen Z, Sathyan KM, Geng Y, Zheng R, Chakraborty A, et al. 2010. A WD-repeat protein stabilizes ORC binding to chromatin. *Mol. Cell* 40:99–111
56. Bartke T, Vermeulen M, Xhemalce B, Robson SC, Mann M, Kouzarides T. 2010. Nucleosome-interacting proteins regulated by DNA and histone methylation. *Cell* 143:470–84
57. Bowers JL, Randell JC, Chen S, Bell SP. 2004. ATP hydrolysis by ORC catalyzes reiterative Mcm2–7 assembly at a defined origin of replication. *Mol. Cell* 16:967–78
58. Chesnokov IN, Chesnokova ON, Botchan M. 2003. A cytokinetic function of *Drosophila* ORC6 protein resides in a domain distinct from its replication activity. *PNAS* 100:9150–55
59. Miller TCR, Locke J, Greiwe JF, Diffley JFX, Costa A. 2019. Mechanism of head-to-head MCM double-hexamer formation revealed by cryo-EM. *Nature* 575:704–10

60. Bleichert F, Botchan MR, Berger JM. 2015. Crystal structure of the eukaryotic origin recognition complex. *Nature* 519:321–26
61. Schmidt JM, Bleichert F. 2020. Structural mechanism for replication origin binding and remodeling by a metazoan origin recognition complex and its co-loader Cdc6. *Nat. Commun.* 11:4263
62. Li N, Lam WH, Zhai Y, Cheng J, Cheng E, et al. 2018. Structure of the origin recognition complex bound to DNA replication origin. *Nature* 559:217–22
63. Jaremko MJ, On KF, Thomas DR, Stillman B, Joshua-Tor L. 2020. The dynamic nature of the human origin recognition complex revealed through five cryoEM structures. *eLife* 9:e58622
64. Feng X, Noguchi Y, Barbon M, Stillman B, Speck C, Li H. 2021. The structure of ORC-Cdc6 on an origin DNA reveals the mechanism of ORC activation by the replication initiator Cdc6. *Nat. Commun.* 12:3883
65. Gaudier M, Schuwirth BS, Westcott SL, Wigley DB. 2007. Structural basis of DNA replication origin recognition by an ORC protein. *Science* 317:1213–16
66. Dueber EL, Corn JE, Bell SD, Berger JM. 2007. Replication origin recognition and deformation by a heterodimeric archaeal Orc1 complex. *Science* 317:1210–13
67. Hu Y, Tareen A, Sheu YJ, Ireland WT, Speck C, et al. 2020. Evolution of DNA replication origin specification and gene silencing mechanisms. *Nat. Commun.* 11:5175
68. De Ioannes P, Leon VA, Kuang Z, Wang M, Boeke JD, et al. 2019. Structure and function of the Orc1 BAH-nucleosome complex. *Nat. Commun.* 10:2894
69. Harland R. 1981. Initiation of DNA replication in eukaryotic chromosomes. *Trends Biochem. Sci.* 6:71–74
70. Blow JJ, Laskey RA. 1988. A role for the nuclear envelope in controlling DNA replication within the cell cycle. *Nature* 332:546–48
71. Kubota Y, Mimura S, Nishimoto S-I, Takisawa H, Nojima H. 1995. Identification of the yeast MCM3-related protein as a component of *Xenopus* DNA replication licensing factor. *Cell* 81:601–9
72. Chong JPJ, Mahbubani HM, Khoo C-Y, Blow JJ. 1995. Purification of an MCM-containing complex as a component of the DNA replication licensing system. *Nature* 375:418–21
73. Madine MA, Khoo CY, Mills AD, Laskey RA. 1995. MCM3 complex required for cell cycle regulation of DNA replication in vertebrate cells. *Nature* 375:421–24
74. Diffley JFX, Cocker JH, Dowell SJ, Rowley A. 1994. Two steps in the assembly of complexes at yeast replication origins in vivo. *Cell* 78:303–16
75. Aparicio OM, Weinstein DM, Bell SP. 1997. Components and dynamics of DNA replication complexes in *S. cerevisiae*: redistribution of MCM proteins and Cdc45p during S phase. *Cell* 91:59–69
76. Tanaka T, Knapp D, Nasmyth K. 1997. Loading of an Mcm protein onto DNA-replication origins is regulated by Cdc6p and CDKs. *Cell* 90:649–60
77. Cocker JH, Piatti S, Santocanale C, Nasmyth K, Diffley JFX. 1996. An essential role for the Cdc6 protein in forming the pre-replicative complexes of budding yeast. *Nature* 379:180–82
78. Coleman TR, Carpenter PB, Dunphy WG. 1996. The *Xenopus* Cdc6 protein is essential for the initiation of a single round of DNA replication in cell-free extracts. *Cell* 87:53–63
79. Donovan S, Harwood J, Drury LS, Diffley JFX. 1997. Cdc6p-dependent loading of Mcm proteins onto pre-replicative chromatin in budding yeast. *PNAS* 94:5611–16
80. Li N, Zhai Y, Zhang Y, Li W, Yang M, et al. 2015. Structure of the eukaryotic MCM complex at 3.8 Å. *Nature* 524:186–91
81. Zhai Y, Cheng E, Wu H, Li N, Yung PY, et al. 2017. Open-ringed structure of the Cdt1-Mcm2–7 complex as a precursor of the MCM double hexamer. *Nat. Struct. Mol. Biol.* 24:300–8
82. Frigola J, He J, Kinkelin K, Pye VE, Renault L, et al. 2017. Cdt1 stabilizes an open MCM ring for helicase loading. *Nat. Commun.* 8:15720
83. Samel SA, Fernandez-Cid A, Sun J, Riera A, Tognetti S, et al. 2014. A unique DNA entry gate serves for regulated loading of the eukaryotic replicative helicase MCM2–7 onto DNA. *Genes Dev.* 28:1653–66
84. Parker MW, Bell M, Mir M, Kao JA, Darzacq X, et al. 2019. A new class of disordered elements controls DNA replication through initiator self-assembly. *eLife* 8:e48562
85. Frigola J, Remus D, Mehanna A, Diffley JFX. 2013. ATPase-dependent quality control of DNA replication origin licensing. *Nature* 495:339–43

86. Yuan Z, Schneider S, Dodd T, Riera A, Bai L, et al. 2020. Structural mechanism of helicase loading onto replication origin DNA by ORC-Cdc6. *PNAS* 117:17747–56
87. Yuan Z, Riera A, Bai L, Sun J, Nandi S, et al. 2017. Structural basis of Mcm2–7 replicative helicase loading by ORC-Cdc6 and Cdt1. *Nat. Struct. Mol. Biol.* 24:316–24
88. Ticau S, Friedman LJ, Champasa K, Correa IR Jr., Gelles J, Bell SP. 2017. Mechanism and timing of Mcm2–7 ring closure during DNA replication origin licensing. *Nat. Struct. Mol. Biol.* 24:309–15
89. Ticau S, Friedman LJ, Ivica NA, Gelles J, Bell SP. 2015. Single-molecule studies of origin licensing reveal mechanisms ensuring bidirectional helicase loading. *Cell* 161:513–25
90. Coster G, Frigola J, Beuron F, Morris EP, Diffley JFX. 2014. Origin licensing requires ATP binding and hydrolysis by the MCM replicative helicase. *Mol. Cell* 55:666–77
91. Kang S, Warner MD, Bell SP. 2014. Multiple functions for Mcm2–7 ATPase motifs during replication initiation. *Mol. Cell* 55:655–65
92. Frigola J, Remus D, Mehanna A, Diffley JFX. 2013. ATPase-dependent quality control of DNA replication origin licensing. *Nature* 495:339–43
93. Sanchez H, McCluskey K, van Laar T, van Veen E, Asscher FM, et al. 2021. DNA replication origins retain mobile licensing proteins. *Nat. Commun.* 12:1908
94. Noguchi K, Vassilev A, Ghosh S, Yates JL, DePamphilis ML. 2006. The BAH domain facilitates the ability of human Orc1 protein to activate replication origins *in vivo*. *EMBO J.* 25:5372–82
95. Abid Ali F, Douglas ME, Locke J, Pye VE, Nans A, et al. 2017. Cryo-EM structure of a licensed DNA replication origin. *Nat. Commun.* 8:2241
96. Dahmann C, Diffley JFX, Nasmyth KA. 1995. S-phase-promoting cyclin-dependent kinases prevent re-replication by inhibiting the transition of replication origins to a pre-replicative state. *Curr. Biol.* 5:1257–69
97. Noton E, Diffley JFX. 2000. CDK inactivation is the only essential function of the APC/C and the mitotic exit network proteins for origin resetting during mitosis. *Mol. Cell* 5:85–95
98. Drury LS, Perkins G, Diffley JFX. 1997. The Cdc4/34/53 pathway targets Cdc6p for proteolysis in budding yeast. *EMBO J.* 16:5966–76
99. Drury LS, Perkins G, Diffley JFX. 2000. The cyclin-dependent kinase Cdc28p regulates distinct modes of Cdc6p proteolysis during the budding yeast cell cycle. *Curr. Biol.* 10:231–40
100. Elsasser S, Chi Y, Yang P, Campbell JL. 1999. Phosphorylation controls timing of Cdc6p destruction: a biochemical analysis. *Mol. Biol. Cell* 10:3263–77
101. Perkins G, Drury LS, Diffley JFX. 2001. Separate SCF^{CDC4} recognition elements target Cdc6 for proteolysis in S phase and mitosis. *EMBO J.* 20:4836–45
102. Piatti S, Bohm T, Cocker JH, Diffley JFX, Nasmyth K. 1996. Activation of S-phase-promoting CDKs in late G1 defines a “point of no return” after which Cdc6 synthesis cannot promote DNA replication in yeast. *Genes Dev.* 10:1516–31
103. Labib K, Diffley JFX, Kearsley SE. 1999. G1-phase and B-type cyclins exclude the DNA-replication factor Mcm4 from the nucleus. *Nat. Cell Biol.* 1:415–22
104. Nguyen VQ, Co C, Irie K, Li JJ. 2000. Clb/Cdc28 kinases promote nuclear export of the replication initiator proteins Mcm2–7. *Curr. Biol.* 10:195–205
105. Liku ME, Nguyen VQ, Rosales AW, Irie K, Li JJ. 2005. CDK phosphorylation of a novel NLS-NES module distributed between two subunits of the Mcm2–7 complex prevents chromosomal rereplication. *Mol. Biol. Cell* 16:5026–39
106. Tanaka S, Diffley JFX. 2002. Interdependent nuclear accumulation of budding yeast Cdt1 and Mcm2–7 during G1 phase. *Nat. Cell Biol.* 4:198–207
107. Hennessy KM, Clark CD, Botstein D. 1990. Subcellular localization of yeast CDC46 varies with the cell cycle. *Genes Dev.* 4:2252–63
108. Chen S, Bell SP. 2011. CDK prevents Mcm2–7 helicase loading by inhibiting Cdt1 interaction with Orc6. *Genes Dev.* 25:363–72
109. Wilmes GM, Archambault V, Austin RJ, Jacobson MD, Bell SP, Cross FR. 2004. Interaction of the S-phase cyclin Clb5 with an ‘RXL’ docking sequence in the initiator protein Orc6 provides an origin-localized replication control switch. *Genes Dev.* 18:981–91

110. Nguyen VQ, Co C, Li JJ. 2001. Cyclin-dependent kinases prevent DNA re-replication through multiple mechanisms. *Nature* 411:1068–73
111. Drury LS, Diffley JFX. 2009. Factors affecting the diversity of DNA replication licensing control in eukaryotes. *Curr. Biol.* 19:530–35
112. McGarry TJ, Kirschner MW. 1998. Geminin, an inhibitor of DNA replication, is degraded during mitosis. *Cell* 93:1043–53
113. Arias EE, Walter JC. 2005. Replication-dependent destruction of Cdt1 limits DNA replication to a single round per cell cycle in *Xenopus* egg extracts. *Genes Dev.* 19:114–26
114. Takeda DY, Parvin JD, Dutta A. 2005. Degradation of Cdt1 during S phase is Skp2-independent and is required for efficient progression of mammalian cells through S phase. *J. Biol. Chem.* 280:23416–23
115. Mendez J, Zou-Yang XH, Kim SY, Hidaka M, Tansey WP, Stillman B. 2002. Human origin recognition complex large subunit is degraded by ubiquitin-mediated proteolysis after initiation of DNA replication. *Mol. Cell* 9:481–91
116. Araki H. 2011. Initiation of chromosomal DNA replication in eukaryotic cells; contribution of yeast genetics to the elucidation. *Genes Genet. Syst.* 86:141–49
117. Burgers PMJ, Kunkel TA. 2017. Eukaryotic DNA replication fork. *Annu. Rev. Biochem.* 86:417–38
118. Labib K. 2010. How do Cdc7 and cyclin-dependent kinases trigger the initiation of chromosome replication in eukaryotic cells? *Genes Dev.* 24:1208–19
119. Zou L, Stillman B. 2000. Assembly of a complex containing Cdc45p, replication protein A, and Mcm2p at replication origins controlled by S-phase cyclin-dependent kinases and Cdc7p-Dbf4p kinase. *Mol. Cell. Biol.* 20:3086–96
120. Deegan TD, Yeeles JT, Diffley JFX. 2016. Phosphopeptide binding by Sld3 links Dbf4-dependent kinase to MCM replicative helicase activation. *EMBO J.* 35:961–73
121. Randell JC, Fan A, Chan C, Francis LI, Heller RC, et al. 2010. Mec1 is one of multiple kinases that prime the Mcm2–7 helicase for phosphorylation by Cdc7. *Mol. Cell* 40:353–63
122. Hardy CF, Dryga O, Seematter S, Pahl PM, Sclafani RA. 1997. *Mcm5/Cdc46-bob1* bypasses the requirement for the S phase activator Cdc7p. *PNAS* 94:3151–55
123. Sheu YJ, Stillman B. 2010. The Dbf4-Cdc7 kinase promotes S phase by alleviating an inhibitory activity in Mcm4. *Nature* 463:113–17
124. Francis LI, Randell JC, Takara TJ, Uchima L, Bell SP. 2009. Incorporation into the prereplicative complex activates the Mcm2–7 helicase for Cdc7-Dbf4 phosphorylation. *Genes Dev.* 23:643–54
125. Tanaka S, Nakato R, Katou Y, Shirahige K, Araki H. 2011. Origin association of Sld3, Sld7, and Cdc45 proteins is a key step for determination of origin-firing timing. *Curr. Biol.* 21:2055–63
126. Tanaka S, Umemori T, Hirai K, Muramatsu S, Kamimura Y, Araki H. 2007. CDK-dependent phosphorylation of Sld2 and Sld3 initiates DNA replication in budding yeast. *Nature* 445:328–32
127. Masumoto H, Muramatsu S, Kamimura Y, Araki H. 2002. S-Cdk-dependent phosphorylation of Sld2 essential for chromosomal DNA replication in budding yeast. *Nature* 415:651–65
128. Zegerman P, Diffley JFX. 2007. Phosphorylation of Sld2 and Sld3 by cyclin-dependent kinases promotes DNA replication in budding yeast. *Nature* 445:281–85
129. Tak YS, Tanaka Y, Endo S, Kamimura Y, Araki H. 2006. A CDK-catalysed regulatory phosphorylation for formation of the DNA replication complex Sld2-Dpb11. *EMBO J.* 25:1987–96
130. Kumagai A, Shevchenko A, Dunphy WG. 2010. Treslin collaborates with TopBP1 in triggering the initiation of DNA replication. *Cell* 140:349–59
131. Kumagai A, Shevchenko A, Dunphy WG. 2011. Direct regulation of Treslin by cyclin-dependent kinase is essential for the onset of DNA replication. *J. Cell Biol.* 193:995–1007
132. Sansam CL, Cruz NM, Danielian PS, Amsterdam A, Lau ML, et al. 2010. A vertebrate gene, *ticrr*, is an essential checkpoint and replication regulator. *Genes Dev.* 24:183–94
133. Boos D, Sanchez-Pulido L, Rappas M, Pearl LH, Oliver AW, et al. 2011. Regulation of DNA replication through Sld3-Dpb11 interaction is conserved from yeast to humans. *Curr. Biol.* 21:1152–57
134. Sanchez-Pulido L, Diffley JFX, Ponting CP. 2010. Homology explains the functional similarities of Treslin/Ticrr and Sld3. *Curr. Biol.* 20:R509–10
135. Muramatsu S, Hirai K, Tak YS, Kamimura Y, Araki H. 2010. CDK-dependent complex formation between replication proteins Dpb11, Sld2, Pol ϵ , and GINS in budding yeast. *Genes Dev.* 24:602–12

136. Mantiero D, Mackenzie A, Donaldson A, Zegerman P. 2011. Limiting replication initiation factors execute the temporal programme of origin firing in budding yeast. *EMBO J.* 30:4805–14
137. Douglas ME, Diffley JFX. 2012. Replication timing: The early bird catches the worm. *Curr. Biol.* 22:R81–82
138. Santocane C, Diffley JFX. 1998. A Mec1- and Rad53-dependent checkpoint controls late-firing origins of DNA replication. *Nature* 395:615–18
139. Shirahige K, Hori Y, Shiraishi K, Yamashita M, Takahashi K, et al. 1998. Regulation of DNA-replication origins during cell-cycle progression. *Nature* 395:618–21
140. Lopez-Mosqueda J, Maas NL, Jonsson ZO, Defazio-Eli LG, Wohlschlegel J, Toczyski DP. 2010. Damage-induced phosphorylation of Sld3 is important to block late origin firing. *Nature* 467:479–83
141. Zegerman P, Diffley JFX. 2010. Checkpoint-dependent inhibition of DNA replication initiation by Sld3 and Dbf4 phosphorylation. *Nature* 467:474–78
142. Douglas ME, Ali FA, Costa A, Diffley JFX. 2018. The mechanism of eukaryotic CMG helicase activation. *Nature* 555:265–68
143. Georgescu R, Yuan Z, Bai L, de Luna Almeida Santos R, Sun J, et al. 2017. Structure of eukaryotic CMG helicase at a replication fork and implications to replisome architecture and origin initiation. *PNAS* 114:E697–706
144. Gros J, Kumar C, Lynch G, Yadav T, Whitehouse I, Remus D. 2015. Post-licensing specification of eukaryotic replication origins by facilitated Mcm2–7 sliding along DNA. *Mol. Cell* 60:797–807
145. Sinha P, Chang V, Tye BK. 1986. A mutant that affects the function of autonomously replicating sequences in yeast. *J. Mol. Biol.* 192:805–14
146. Labib K, Tercero JA, Diffley JFX. 2000. Uninterrupted MCM2–7 function required for DNA replication fork progression. *Science* 288:1643–47
147. Tercero JA, Labib K, Diffley JFX. 2000. DNA synthesis at individual replication forks requires the essential initiation factor Cdc45p. *EMBO J.* 19:2082–93
148. Ishimi Y. 1997. A DNA helicase activity is associated with an MCM4, -6, and -7 protein complex. *J. Biol. Chem.* 272:24508–13
149. Bochman ML, Schwacha A. 2008. The Mcm2–7 complex has in vitro helicase activity. *Mol Cell* 31:287–93
150. Costa A, Ilves I, Tamberg N, Petojevic T, Nogales E, et al. 2011. The structural basis for MCM2–7 helicase activation by GINS and Cdc45. *Nat. Struct. Mol. Biol.* 18:471–77
151. Ilves I, Petojevic T, Pesavento JJ, Botchan MR. 2010. Activation of the MCM2–7 helicase by association with Cdc45 and GINS proteins. *Mol. Cell* 37:247–58
152. Moyer SE, Lewis PW, Botchan MR. 2006. Isolation of the Cdc45/Mcm2–7/GINS (CMG) complex, a candidate for the eukaryotic DNA replication fork helicase. *PNAS* 103:10236–41
153. Gambus A, Jones RC, Sanchez-Diaz A, Kanemaki M, van Deursen F, et al. 2006. GINS maintains association of Cdc45 with MCM in replisome progression complexes at eukaryotic DNA replication forks. *Nat. Cell Biol.* 8:358–66
154. Deegan TD, Diffley JFX. 2016. MCM: one ring to rule them all. *Curr. Opin. Struct. Biol.* 37:145–51
155. Sheu YJ, Stillman B. 2006. Cdc7-Dbf4 phosphorylates MCM proteins via a docking site-mediated mechanism to promote S phase progression. *Mol. Cell* 24:101–13
156. Itou H, Shirakihara Y, Araki H. 2015. The quaternary structure of the eukaryotic DNA replication proteins Sld7 and Sld3. *Acta Crystallogr. D Biol. Crystallogr.* 71:1649–56
157. De Jesus-Kim L, Friedman LJ, Looke M, Ramsomair CK, Gelles J, Bell SP. 2021. DDK regulates replication initiation by controlling the multiplicity of Cdc45-GINS binding to Mcm2–7. *eLife* 10:e65471
158. Champasa K, Blank C, Friedman LJ, Gelles J, Bell SP. 2019. A conserved Mcm4 motif is required for Mcm2–7 double-hexamer formation and origin DNA unwinding. *eLife* 8:e45538
159. Abid Ali F, Renault L, Gannon J, Gahlon HL, Kotecha A, et al. 2016. Cryo-EM structures of the eukaryotic replicative helicase bound to a translocation substrate. *Nat. Commun.* 7:10708
160. Yuan Z, Bai L, Sun J, Georgescu R, Liu J, et al. 2016. Structure of the eukaryotic replicative CMG helicase suggests a pumpjack motion for translocation. *Nat. Struct. Mol. Biol.* 23:217–24
161. Eickhoff P, Kose HB, Martino F, Petojevic T, Abid Ali F, et al. 2019. Molecular basis for ATP-hydrolysis-driven DNA translocation by the CMG helicase of the eukaryotic replisome. *Cell Rep.* 28:2673–88.e8

162. Yuan Z, Georgescu R, Bai L, Zhang D, Li H, O'Donnell ME. 2020. DNA unwinding mechanism of a eukaryotic replicative CMG helicase. *Nat. Commun.* 11:688
163. Baretic D, Jenkyn-Bedford M, Aria V, Cannone G, Skehel M, Yeeles JTP. 2020. Cryo-EM structure of the fork protection complex bound to CMG at a replication fork. *Mol. Cell* 78:926–40.e13
164. Goswami P, Abid Ali F, Douglas ME, Locke J, Purkiss A, et al. 2018. Structure of DNA-CMG-Pol epsilon elucidates the roles of the non-catalytic polymerase modules in the eukaryotic replisome. *Nat. Commun.* 9:5061
165. Langston LD, Mayle R, Schauer GD, Yurieva O, Zhang D, et al. 2017. Mcm10 promotes rapid isomerization of CMG-DNA for replisome bypass of lagging strand DNA blocks. *eLife* 6:e29118
166. Kose HB, Larsen NB, Duxin JP, Yardimci H. 2019. Dynamics of the eukaryotic replicative helicase at lagging-strand protein barriers support the steric exclusion model. *Cell Rep.* 26:2113–25.e6
167. Kose HB, Xie S, Cameron G, Strycharska MS, Yardimci H. 2020. Duplex DNA engagement and RPA oppositely regulate the DNA-unwinding rate of CMG helicase. *Nat. Commun.* 11:3713
168. Kang YH, Farina A, Bermudez VP, Tappin I, Du F, et al. 2013. Interaction between human Ctf4 and the Cdc45/Mcm2-7/GINS (CMG) replicative helicase. *PNAS* 110:19760–65
169. Yeeles JTP, Deegan TD, Janska A, Early A, Diffley JFX. 2015. Regulated eukaryotic DNA replication origin firing with purified proteins. *Nature* 519:431–35
170. Georgescu RE, Langston L, Yao NY, Yurieva O, Zhang D, et al. 2014. Mechanism of asymmetric polymerase assembly at the eukaryotic replication fork. *Nat. Struct. Mol. Biol.* 21:664–70
171. Langston LD, Zhang D, Yurieva O, Georgescu RE, Finkelstein J, et al. 2014. CMG helicase and DNA polymerase ϵ form a functional 15-subunit holoenzyme for eukaryotic leading-strand DNA replication. *PNAS* 111:15390–95
172. Kesti T, Flick K, Keränen S, Syväoja JE, Wittenberg C. 1999. DNA polymerase ϵ catalytic domains are dispensable for DNA replication, DNA repair, and cell viability. *Mol. Cell* 3:679–85
173. Dua R, Levy DL, Campbell JL. 1998. Role of the putative zinc finger domain of *Saccharomyces cerevisiae* DNA polymerase ϵ in DNA replication and the S/M checkpoint pathway. *J. Biol. Chem.* 273:30046–55
174. Sengupta S, van Deursen F, de Piccoli G, Labib K. 2013. Dpb2 integrates the leading-strand DNA polymerase into the eukaryotic replisome. *Curr. Biol.* 23:543–52
175. Looke M, Maloney MF, Bell SP. 2017. Mcm10 regulates DNA replication elongation by stimulating the CMG replicative helicase. *Genes Dev.* 31:291–305
176. Wasserman MR, Schauer GD, ME O'Donnell, Liu S. 2019. Replication fork activation is enabled by a single-stranded DNA gate in CMG helicase. *Cell* 178:600–11.e16
177. Lewis JS, Spenkelink LM, Schauer GD, Hill FR, Georgescu RE, et al. 2017. Single-molecule visualization of *Saccharomyces cerevisiae* leading-strand synthesis reveals dynamic interaction between MTC and the replisome. *PNAS* 114:10630–35
178. Yeeles JT, Janska A, Early A, Diffley JFX. 2017. How the eukaryotic replisome achieves rapid and efficient DNA replication. *Mol. Cell* 65:105–16
179. Puchades C, Sandate CR, Lander GC. 2020. The molecular principles governing the activity and functional diversity of AAA+ proteins. *Nat. Rev. Mol. Cell Biol.* 21:43–58
180. Thomsen ND, Berger JM. 2009. Running in reverse: the structural basis for translocation polarity in hexameric helicases. *Cell* 139:523–34
181. Enemark EJ, Joshua-Tor L. 2006. Mechanism of DNA translocation in a replicative hexameric helicase. *Nature* 442:270–75
182. Meagher M, Epling LB, Enemark EJ. 2019. DNA translocation mechanism of the MCM complex and implications for replication initiation. *Nat. Commun.* 10:3117
183. Gao Y, Cui Y, Fox T, Lin S, Wang H, et al. 2019. Structures and operating principles of the replisome. *Science* 363:eaav7003
184. Attali I, Botchan MR, Berger JM. 2021. Structural mechanisms for replicating DNA in eukaryotes. *Annu. Rev. Biochem.* 90:77–106
185. Simon AC, Zhou JC, Perera RL, van Deursen F, Evrin C, et al. 2014. A Ctf4 trimer couples the CMG helicase to DNA polymerase α in the eukaryotic replisome. *Nature* 510:293–97
186. Vella F, Simon AC, Ortiz Bazan MA, Kilkenny ML, Wirthensohn D, et al. 2016. Ctf4 is a hub in the eukaryotic replisome that links multiple CIP-box proteins to the CMG helicase. *Mol. Cell* 63:385–96

187. Samora CP, Saksouk J, Goswami P, Wade BO, Singleton MR, et al. 2016. Ctf4 links DNA replication with sister chromatid cohesion establishment by recruiting the Chl1 helicase to the replisome. *Mol. Cell* 63:371–84
188. Yuan Z, Georgescu R, Santos RLA, Zhang D, Bai L, et al. 2019. Ctf4 organizes sister replisomes and Pol α into a replication factory. *eLife* 8:e47405
189. Lewis JS, Spenkelink LM, Schauer GD, Yurieva O, Mueller SH, et al. 2020. Tunability of DNA polymerase stability during eukaryotic DNA replication. *Mol. Cell* 77:17–25.e5
190. Yardimci H, Loveland AB, Habuchi S, van Oijen AM, Walter JC. 2010. Uncoupling of sister replisomes during eukaryotic DNA replication. *Mol. Cell* 40:834–40
191. Petronczki M, Chwalla B, Siomos MF, Yokobayashi S, Helmhart W, et al. 2004. Sister-chromatid cohesion mediated by the alternative RF-C^{Ctf18/Dcc1/Ctf8}, the helicase Chl1 and the polymerase- α -associated protein Ctf4 is essential for chromatid disjunction during meiosis II. *J. Cell Sci.* 117:3547–59
192. Gan H, Serra-Cardona A, Hua X, Zhou H, Labib K, et al. 2018. The Mcm2-Ctf4-Pol α axis facilitates parental histone H3-H4 transfer to lagging strands. *Mol. Cell* 72:140–51.e3
193. Hogg M, Osterman P, Bylund GO, Ganai RA, Lundstrom EB, et al. 2014. Structural basis for processive DNA synthesis by yeast DNA polymerase ϵ . *Nat. Struct. Mol. Biol.* 21:49–55
194. Lancey C, Tehseen M, Raducanu VS, Rashid F, Merino N, et al. 2020. Structure of the processive human Pol δ holoenzyme. *Nat. Commun.* 11:1109
195. Zhou JC, Janska A, Goswami P, Renault L, Abid Ali F, et al. 2017. CMG–Pol epsilon dynamics suggests a mechanism for the establishment of leading-strand synthesis in the eukaryotic replisome. *PNAS* 114:4141–46
196. Yuan Z, Georgescu R, Schauer GD, O'Donnell ME, Li H. 2020. Structure of the polymerase ϵ holoenzyme and atomic model of the leading strand replisome. *Nat. Commun.* 11:3156
197. Johansson E, Garg P, Burgers PM. 2004. The Pol32 subunit of DNA polymerase delta contains separable domains for processive replication and proliferating cell nuclear antigen (PCNA) binding. *J. Biol. Chem.* 279:1907–15