Loading of the replicative DNA helicase at origins of replication is of central importance in DNA replication. As the first of the replication fork proteins assemble at chromosomal origins of replication, the loaded helicase is required for the recruitment of the rest of the replication machinery. In this work, we review the current knowledge of helicase loading at *Escherichia coli* and eukaryotic origins of replication. In each case, this process requires both an origin recognition protein as well as one or more additional proteins. Comparison of these events shows intriguing similarities that suggest a similar underlying mechanism, as well as critical differences that likely reflect the distinct processes that regulate helicase loading in bacterial and eukaryotic cells.

Replicative DNA helicases are ring-shaped molecules with a central cavity through which DNA passes as they unwind DNA. Their loading at replication origins is a critical and highly regulated event in chromosomal replication. The DNA helicase is the first of the replication fork proteins recruited to and loaded onto origins of replication, and the loaded helicase is required for the recruitment of the rest of the replication machinery (Remus and Diffley 2009; Kaguni 2011). Indeed, the replicative DNA helicase links the replication machinery to the parental DNA (O’Donnell 2006). In *Escherichia coli* cells, the DnaB replicative helicase binds to primase (DnaG) and the sliding clamp loader, which in turn binds the DNA polymerases. Although the polymerases are also linked to the template DNA by sliding clamps, when these interactions are broken, the polymerases’ association with the sliding clamp loader and the helicase keeps them at the site of replication. The interactions that tether the DNA polymerases to the eukaryotic replication fork are less clear but very likely involve direct and indirect interactions with the Mcm2–7 replicative helicase (Calzada et al. 2005).

Helicase loading is carefully regulated to control the location and frequency of replication initiation. In eukaryotic cells, helicase loading is tightly restricted to the G1 phase of the cell cycle. This constraint is a key part of the mechanisms that ensure that no origin can initiate more than once per cell cycle (Siddiqui et al. 2013; Zielke et al. 2013). In addition, the sites

---

**Editors:** Stephen D. Bell, Marcel Méchali, and Melvin L. DePamphilis

Additional Perspectives on DNA Replication available at www.cshperspectives.org

Copyright © 2013 Cold Spring Harbor Laboratory Press; all rights reserved; doi: 10.1101/cshperspect.a010124

Cite this article as *Cold Spring Harb Perspect Biol* 2013;5:a010124
of eukaryotic replicative helicase loading define the potential sites of replication initiation in the cell (but not all loaded helicases are used during a given S phase [Rhind and Gilbert 2012]). Although the central regulated event in bacterial chromosome duplication is the recruitment of the ATP-bound initiator protein DnaA (Skarstad and Katayama 2013), the loading of the replicative helicase represents a key committed step during initiation.

Here we will discuss the mechanism of helicase loading in bacteria and eukaryotic cells. Much of the discussion will focus on studies in the bacterium *E. coli*, the yeast *Saccharomyces cerevisiae*, and the frog *Xenopus laevis*, in which the events of helicase loading are best understood. Comparison of these mechanisms shows important similarities and differences between the domains of life. In both bacteria and eukaryotic cells, multiple AAA+ proteins use ATP binding and hydrolysis to direct helicase loading and both helicases are initially loaded in an inactive form. On the other hand, the eukaryotic helicase is loaded around double-stranded DNA (dsDNA) and as a double hexamer, whereas the bacterial helicase is loaded around single-stranded DNA (ssDNA) as a single hexamer. These distinctions are very likely due to the very different regulatory mechanisms of DNA replication in bacteria and eukaryotic cells.

Before describing the process of helicase loading, it is relevant to know that the essential function of replicative helicases is to unwind the parental duplex DNA using the energy provided by the hydrolysis of a nucleoside triphosphate (Table 1) (reviewed in Patel et al. 2011). Replicative DNA polymerases then copy each parental DNA strand to duplicate the genome. That replicative DNA helicases are hexameric (and sometimes heptameric) structures that have a positively charged, central channel provides a framework for how these molecular machines work. Several models that describe the mechanism of unwinding have been considered. One is the steric or strand-exclusion model, in which the helicase excludes one strand of DNA while the other passes through the central cavity as the enzyme moves. Because these enzymes can bind and translocate on duplex DNA without unwinding (Kaplan et al. 2003), two additional models have been proposed. In the ploughshare model, duplex DNA enters the central cavity and exits as unwound DNA by virtue of a domain/protein that acts as a ploughshare or pin, which disrupts the hydrogen bonds of DNA as it is pumped through the enzyme (reviewed in Takahashi et al. 2005). A second model, the DNA-pumping model, was proposed on the basis of the double-hexameric forms of SV40, Mcm2–7, and other replicative helicases (Mastrangelo et al. 1989; Remus et al. 2009). This model proposes that the two helicases pump duplex DNA toward one another, resulting in torsional strain that forces the two strands apart, at which point they exit the central channel as two ssDNA loops. Current evidence supports the steric exclusion model (Jezewska et al. 1998b; Kaplan 2000; Galletto et al. 2004; Fu et al. 2011). Of interest, the *E. coli* enzyme moves in the 5’-to-3’ direction relative to the engaged ssDNA that passes through the central cavity, whereas archaeal and eukaryotic enzymes move in the 3’-to-5’ direction.

**HELICASE LOADING IN *E. coli***

The *E. coli* replication origin (oriC) has two essential functions. One is to serve as a site where

| Table 1. Replicative DNA helicases of free-living organisms are hexameric |
|-----------------|-----------------|-----------------|-----------------|
| Domain         | Model organism(s) | Helicase*       | Direction of movement |
| Bacteria       | *Escherichia coli* | DnaB            | 5’ → 3’          |
| Archaea        | *Sulfolobus solfataricus* | MCM            | 3’ → 5’          |
| Eukarya        | *Saccharomyces cerevisiae*  | Mcm2–7         | 3’ → 5’          |
|                | *Drosophila melanogaster*  |                |                  |
|                | *Xenopus laevis*          |                |                  |

* DNA B and *S. solfataricus* MCM are homohexamers, whereas eukaryotic Mcm2–7 is composed of six nonidentical subunits.
the replication fork machinery assembles. DNA sequence motifs in this chromosomal locus and their roles are described in detail elsewhere (Leonard and Méchali 2013). The second is to act as a site where the process of DNA replication is controlled. As described in Skarstad and Katayama (2013), several separate mechanisms modulate the frequency of initiation of DNA replication. At this stage, a critical event is the loading of the replicative helicase named DnaB (Bell and Botchan 2013), which must be in complex with its partner, DnaC. A summary of the functions of DnaA, DnaB, and DnaC serves as the foundation for a description of the helicase loading process.

DnaA IS THE REPLICATION INITIATOR IN BACTERIA

E. coli DnaA performs a central role in the initiation of DNA replication at oriC, recognizing specific DNA sequences named the DnaA box, I-site, τ-site, and C-site within oriC (Leonard and Méchali 2013). Studies reveal that this protein, which is highly conserved among bacteria, can be divided into four domains designated 1 to 4 from amino to carboxyl terminus (http://www.molgen.mpg.de/~messer/) (reviewed in Mott and Berger 2007; Ozaki and Katayama 2009; Leonard and Grimwade 2011). The interactions of these domains with specific proteins, DNA sequences, and nucleotides are summarized below.

The amino-terminal domain 1 (amino acids 1–90 of E. coli DnaA) interacts with a variety of proteins, including DnaB (Sutton et al. 1998), and DnaA itself in the assembly of the DnaA oligomer at oriC (Simmons et al. 2003; Felczak and Kaguni 2004; Abe et al. 2007). These interactions are critical for helicase loading and are described in more detail below. Other proteins that interact with domain 1 are HU (Chodavarapu et al. 2008a), DiaA (Keyamura et al. 2007, 2009), Dps (Chodavarapu et al. 2008b), ribosomal protein L2 (Chodavarapu et al. 2011), and RNA polymerase (Flätten et al. 2009). The interaction with RNA polymerase correlates with the role of DnaA as a transcriptional activator in stabilizing the binding of RNA polymerase at the gidA and λP_R promoters (Szalewska-Pałasz et al. 1998; Flätten et al. 2009), and possibly at the glpD and fliC promoters (Mizushima et al. 1994; Messer and Weigel 1997). With HU or DiaA, their respective binding stabilizes DnaA oligomerized at oriC to stimulate initiation in vitro (Keyamura et al. 2007, 2009; Chodavarapu et al. 2008a). By comparison, L2 inhibits initiation by impairing the formation of a DnaA oligomer at oriC (Chodavarapu et al. 2011). Like L2, Dps bound to domain 1 interferes with strand opening of oriC in vitro, which correlates with less frequent initiation in vivo when the Dps level is elevated (Chodavarapu et al. 2008b). These observations suggest that domain 1 acts as a sensor, responding to proteins that modulate the initiation process by affecting the assembly of DnaA at oriC.

Among DnaAs of bacteria (http://www.molgen.mpg.de/~messer/), domain 2 (residues 90–130 of E. coli DnaA) is not conserved in amino acid sequence or length. For E. coli DnaA, the removal of residues 96–120 in domain 2 leads to reduced activity in DNA replication, but deletion of other consecutive sequences from this region has no apparent effect on activity (Nozaki and Ogawa 2008; Molt et al. 2009). These observations suggest that domain 2 acts as a flexible link between domains 1 and 3.

The third domain of DnaA (residues 130–347 of E. coli DnaA) is responsible for ATP binding and hydrolysis (Duderstadt and Berger 2008). Like other members of the AAA+ family of ATPases, this region of DnaA carries the Walker A (P-loop) and B boxes that are involved in ATP binding. The sensor I, II (box VII), and box VIII motifs are believed to coordinate ATP hydrolysis with a change in conformation (reviewed in Erzberger and Berger 2006). The adenine nucleotide bound to a shortened variant of Aquifex aeolicus DnaA lacking domains 1 and 2 has a striking effect on its X-ray crystal structure (Erzberger et al. 2002, 2006). With ADP, six monomers of this DnaA assemble as a closed ring. With ATP instead, an open ring with a right-handed pitch forms by virtue of a distended conformation of domain 3.

Acidic phospholipids in a fluid bilayer bind to a segment spanning the junction between
domains 3 and 4 (reviewed in Boeneman and Crooke 2005). Intriguingly, this interaction promotes nucleotide release from DnaA (Sekimizu and Kornberg 1988). As the concentration of ATP is higher than that of ADP in vivo, this effect may favor the ATP-bound state of DnaA and is the basis for speculation that this interaction may regulate the activity of DnaA. Although fluorescence microscopy indicates that DnaA associates with the inner membrane in vivo (Boeneman et al. 2009) as well as with the oriC region (Nozaki et al. 2009), it is not known if interactions with the inner membrane promote nucleotide exchange in living cells.

Domain 4 (residues 347–467) binds to the DnaA box. Nuclear magnetic resonance and X-ray crystallographic analysis of this domain reveal a helix-turn-helix amino acid motif and a basic loop (Erzberger et al. 2002; Fujikawa et al. 2003). Combined with the biochemical characterization of mutant DnaAs bearing specific amino acid substitutions that impair the recognition of the DnaA box, we have an atomic understanding of how DnaA binds to the DnaA box (Sutton and Kaguni 1997; Blaesing et al. 2000).

**DnaB is the replicative helicase, and interacts with DnaC to form the DnaB–DnaC complex**

In *E. coli*, DnaB is the replicative helicase whose native structure is a toroid of six identical subunits oriented in the same direction (Rehakrantz and Hurwitz 1978; Arai et al. 1981; Donate et al. 2000; Bailey et al. 2007). Whereas Mg$^{2+}$ ion stabilizes this native structure, removal of the metal ion by dialysis or chelation causes the DnaB hexamer to dissociate into trimers and monomers (Bujalowski et al. 1994). X-ray crystallography of the homologous proteins from *Geobacillus kaustophilus* and *Geobacillus stearothermophilus* provides a detailed view of this helicase and the relative arrangement of its smaller amino-terminal and larger carboxy-terminal domains (Bailey et al. 2007; Lo et al. 2009).

Studies of *E. coli* DnaB and others in the DnaB helicase family indicate that the enzyme unwinds DNA by translocating in the 5′-to-3′ direction on the ssDNA to which it is bound (Venkatesan et al. 1982; Matson et al. 1983; Lebowitz and McMacken 1986; Lee et al. 1989; Richardson and Nossal 1989). Fluorescent energy transfer experiments suggest that 20 ± 3 nucleotides of ssDNA pass through the central cavity of DnaB during translocation (Jeziowska et al. 1996; Kaplan 2000). DnaB is uniquely oriented on a forked DNA molecule with a 5′ ssDNA tail (Jeziowska et al. 1998a,b). On the basis of this orientation, the smaller amino-terminal domain is upstream of the carboxy-terminal domain, which is nearer the apex of the replication fork (Fig. 1, step 3).

To participate in helicase loading at oriC, DnaB must be in a complex with its partner, DnaC. Early studies suggested that ATP bound to DnaC is necessary for DnaC to form an isolable complex with DnaB and to protect DnaC from inactivation by *N*-ethylmaleimide (Wickner and Hurwitz 1975; Kobori and Kornberg 1982). More recent studies indicate that ATP is not needed for DnaC to interact with DnaB (Davey et al. 2002; Galletto et al. 2003; Mott et al. 2008). Both cryoelectron microscopy of the DnaB–DnaC complex in comparison with DnaB and fluorescence energy transfer experiments indicate that DnaC is bound to the larger (carboxy-terminal) end of the DnaB toroid (Barcena et al. 2001; Galletto et al. 2003). Relative to the ensuing direction of helicase movement, DnaC in the DnaB–DnaC complex is positioned proximal to the replication fork junction (Fig. 1).

Because amino acid substitutions within residues 10–44 near the amino terminus of DnaC impair the interaction between DnaC and DnaB, this region of DnaC apparently binds directly to DnaB to form the DnaB–DnaC complex (Ludlam et al. 2001). Initial studies showed that six DnaC monomers bind per DnaB hexamer (Wickner and Hurwitz 1975; Kobori and Kornberg 1982; Lanka and Schuster 1983). Subsequent sedimentation velocity analysis to determine the affinity of DnaC for DnaB in combination with estimates of the in vivo levels of DnaC and DnaB suggest that the cellular DnaB–DnaC complex is heterogeneous in composition, with a fraction containing less than the...
Helicase Loading at Chromosomal Origins

Figure 1. Helicase loading at the *E. coli* replication origin (*oriC*) is a stepwise process. Near the top of the figure, DNA sequence elements in *E. coli* *oriC* are shown (see Leonard and Méchali 2013 for details). Binding sites for Fis and IHF are shown, which are believed to alter the architecture of *oriC*, and their roles are described in detail in Leonard and Méchali (2013). The four domains of DnaA shown in green are represented by the overlaid numbers. In step 1, DnaA complexed to ATP binds to the DnaA boxes, I-, τ-, and C-sites to form a DnaA oligomer. Following the unwinding of the region of *oriC* containing the 13-mers, DnaA then loads a DnaB–DnaC complex on each of the separated strands (step 2). In step 3, primase (DnaG) interacts with the amino-terminal region of DnaB. Primer formation by primase (steps 4 and 5) leads to the dissociation of DnaC from the carboxy-terminal domain of DnaB, which is necessary to activate DnaB as a DNA helicase (see Bell and Botchan 2013). Although not shown, the primers displayed in black are then extended by the cellular replicase, DNA polymerase III holoenzyme, for the continuous synthesis of each leading strand at each replication fork (see Johansson and Dixon 2013). A dimer of this DNA polymerase is believed to be at each replication fork. DnaB is proposed to interact with one unit of a dimer of DNA polymerase III holoenzyme, coordinating the unwinding of the parental DNA with DNA synthesis. As this DNA helicase translocates in the 5′-to-3′ direction to support replication fork movement, primase occasionally interacts with DnaB to synthesize additional primers. These primers are used by the unit of the DNA polymerase dimer that synthesizes Okazaki fragments.
maximum of six DnaC monomers per DnaB hexamer (Galletto et al. 2003).

**DnaC CONTROLS THE ACTIVITY OF DnaB**

DnaC, like DnaA, is a member of the AAA\(^+\) family of ATPases (Makarova and Koonin 2013). Despite the presence in DnaC of motifs shared by AAA\(^+\) family members and the high affinity of other AAA\(^+\) proteins for ATP; DnaC binds weakly to this nucleotide; the \(K_d\) is about 8 \(\mu\)M (Davey et al. 2002; Galletto et al. 2003; Biswas et al. 2004). By itself, DnaC is also a feeble ATPase. The influence of ATP on DnaC function has been an enigma, but new evidence suggests that its ATPase activity is activated at a specific step of the initiation process (see below). Stimulated by ATP, DnaC also interacts weakly with ssDNA (Learn et al. 1997; Davey et al. 2002; Mott et al. 2008), supporting the idea that this activity may be involved in helicase loading. Because DnaC complexed to DnaB inhibits its ATPase and helicase activities (Wahle et al. 1989a,b; Allen and Kornberg 1991; Davey et al. 2002; Mott et al. 2008), DnaC must dissociate from DnaB for helicase activation to occur.

**DnaA UNWINDS A REGION NEAR THE LEFT BORDER OF oriC**

With a negatively supercoiled plasmid carrying oriC, DnaA bound to ATP unwinds an AT-rich region containing the M and R 13-mers as judged by potassium permanganate sensitivity (Fig. 1) (Sekimizu et al. 1987; Bramhill and Kornberg 1988; Gille and Messer 1991). Considering that isolated plasmids have a negative superhelix density ranging from −0.05 to −0.08, as much as 330 to 530 nucleotides can be single-stranded if these DNAs are supercoiled solely as a result of unwinding of the duplex DNA instead of the writhing of the DNA upon itself (Cozzarelli 1980; Bliska and Cozzarelli 1987). These observations suggest that DnaA confines some of the single-stranded character of the plasmid to this AT-rich region of oriC. To provide a biochemical mechanism of unwinding, recent studies on the assembly of a DnaA oligomer as a right-handed filament suggest that positively charged and hydrophobic amino acids that line the interior of the DnaA filament interact with the negatively charged ssDNA (Erzberger et al. 2006; Ozaki et al. 2008; Duderstadt et al. 2010, 2011; Ozaki and Katayama 2011). Consistent with the ATP requirement for DnaA-dependent origin unwinding, the formation of this DnaA filament requires the ATP-bound form of the enzyme.

**HELICASE LOADING AT oriC BY DnaA**

After unwinding of oriC, DnaA loads the DnaB–DnaC complex onto this unwound DNA, which apparently becomes enlarged to include the left 13-mer and additional sequences beyond it (Fig. 1, step 2) (Davey et al. 2002). Footprinting studies suggest that the DnaB–DnaC complex protects a region on the top strand near the left border of oriC. On the bottom strand the area protected is to the left of DnaA box R1 that overlaps the right 13-mer. Quantitative analysis of this protein complex assembled at oriC supports the conclusion from footprinting studies that two DnaB–DnaC complexes are bound (Fang et al. 1999; Carr and Kaguni 2001).

Helicase loading appears to involve two regions of DnaA. One region is within domain 3, based on studies of a monoclonal antibody that interferes with the interaction between DnaA and DnaB measured in solid-phase binding assays (Marszalek and Kaguni 1994; Sutton et al. 1998). This antibody recognizes an epitope within residues 111–148 of DnaA. Deletion analysis localized the interacting region to amino acids 135–148 (Seitz et al. 2000). A region within domain 1 is also needed because an alanine substitution for phenylalanine 46 abrogates this interaction (Keyamura et al. 2009).

On the basis that DnaC from *A. aeolicus* can interact with DnaA, and like DnaA forms a helical filament as determined by X-ray crystallography, a specific proposal is that DnaA oligomerized at oriC loads one DnaB–DnaC complex on the top strand via the interaction between DnaA molecules at the left end of the DnaA filament and DnaC in the DnaB–DnaC complex (Mott et al. 2008). The second DnaB–DnaC complex loads on the bottom strand.

---

S.P. Bell and J.M. Kaguni
through the interaction between the left end of the DnaA oligomer and DnaB complexed to DnaC. During loading, it is interesting to consider that one of the interfaces between DnaB protomers must open in this ring-shaped protein in order for the helicase to encircle the ssDNA. The role of DnaA or DnaC in this ring opening remains unclear.

After DnaA loads one DnaB–DnaC complex on each of the separated strands of oriC, DnaC must dissociate for DnaB to be active as a helicase. Of interest, recent studies suggest that three DnaC monomers are bound to each DnaB hexamer (Makowska-Grzyska and Kaguni 2010), whereas earlier experiments suggested that the DnaB–DnaC complex contains six DnaC monomers per DnaB (Wickner and Hurwitz 1975; Kobori and Kornberg 1982; Lank and Schuster 1983; Wahl et al. 1989a). Perhaps the different forms of the DnaB–DnaC complex function at the initiation stage of DNA replication and in restarting stalled or collapsed replication forks.

ATP or ATPγS bound to DnaC supports the DnaA-dependent loading of the DnaB–DnaC complex at oriC (Davey et al. 2002; Makowska-Grzyska and Kaguni 2010). That DnaC remains bound to DnaB in the presence of either ATP or the analog suggests that the act of helicase loading does not stimulate the hydrolysis of ATP bound to DnaC or its release from DnaB (Makowska-Grzyska and Kaguni 2010). In the presence of ATP and other ribonucleotides, the inclusion of primase (DnaG), which interacts with the amino-terminal domain of DnaB as it forms primers for DNA replication, induces the release of DnaC from the carboxy-terminal (large) domain of DnaB (Fig. 1, steps 3–5). Apparently the interaction of primase with DnaB while it synthesizes a primer alters the conformation of DnaB that causes DnaC to dissociate. Other results with mutant forms of DnaC that are speculated to be defective in ATP hydrolysis suggest that this interaction of primase with DnaB stimulates the hydrolysis of ATP bound by DnaC. This step of helicase activation is comparable in function to the activation of the Mcm2–7 complex at a eukaryotic replication origin (Labib 2010).

HELICASE LOADING IN EUKARYOTIC CELLS

Origin Licensing, Pre-RC Formation, and Helicase Loading

Early work on eukaryotic DNA replication revealed that the events of replication initiation were separated into separate cell-cycle stages. Mammalian cell fusion studies suggested the existence of positive factors in S-phase cells that cause G1 nuclei to initiate DNA replication (Johnson and Rao 1970). Analyses of cycling Xenopus egg extracts suggested the existence of a “licensing factor” that gained access to the DNA only after nuclear envelope breakdown (during mitosis) and was consumed during S phase (Blow and Laskey 1988). Finally, in vivo footprinting studies at S. cerevisiae origins showed a distinct and more extensive protection pattern during G1 (called the prereplicative complex or pre-RC) than in S, G2, and M phase (Diffley et al. 1994).

We now know that replication origin licensing and pre-RC formation are both related to the helicase loading event that only occurs in G1 cells (Bell and Dutta 2002; Sclafani and Holzen 2007). The licensing factor that is excluded from the nucleus is probably not identical among different cell types (Arias and Walter 2007), but origin licensing is clearly equivalent to helicase loading. The pre-RC footprint that is associated with S. cerevisiae origins during G1 is due to binding of helicase loading factors to the origin DNA (Perkins and Diffley 1998; Speck et al. 2005). Much of this footprint is attributed to origin recognition complex (ORC) and Cdc6, but it additionally reflects helicase loading. For this reason the term “pre-RC formation” has become synonymous with helicase loading. We will refer to the G1 events at origins as helicase loading and the events that occur in S phase as helicase activation and replisome assembly because these terms most accurately reflect the events that are occurring at the origin.

Mcm2–7 HELICASE

The eukaryotic replicative DNA helicase is the Mcm2–7 complex. Each Mcm2–7 complex contains one copy of the six essential and related
Mcm2, Mcm3, Mcm4, Mcm5, Mcm6, and Mcm7 proteins (reviewed in Bochman and Schwacha 2009). Although high-resolution structural data are not available, electron microscopy (EM) and subunit interaction studies indicate that the Mcm2–7 proteins interact in a defined order (Davey et al. 2003; Yu et al. 2004; Bochman et al. 2008) to form either a ring (Remus et al. 2009) or a gapped ring (Costa et al. 2011) with a positively charged central channel (Fig. 2). The carboxy-terminal half of each Mcm2–7 subunit contains a conserved AAA$^+$ domain that includes insertions specific to the Mcm2–7 proteins (Iyer et al. 2004) that are predicted to form β-hairpins that may interact with the ssDNA during unwinding (Bochman and Schwacha 2009). As with other AAA$^+$ multimers, the interface between subunits in the Mcm2–7 complex forms an ATPase active site. The amino-terminal half of each subunit contains an OB-fold-related motif that is found in many ssDNA-binding proteins, and several of the subunits also contain a Zn-finger motif. Although the AAA$^+$ domain is highly conserved among Mcm2–7 subunits, the characteristic amino- and carboxy-terminal extensions allow homologs for each of the six Mcm2–7 proteins to be readily identified in all eukaryotes studied. As described above, as a $3’ \to 5’$ helicase the Mcm2–7 helicase moves on the leading-strand DNA template at a replication fork.

**EUKARYOTIC HELICASE LOADING PROTEINS**

Eukaryotic helicase loading is directed by three proteins: ORC, Cdc6, and Cdt1. ORC is a heterohexamer with five of the six subunits (Orc1, -2, -3, -4, and -5) showing homology with AAA$^+$ proteins, although only two of these subunits (Orc1 and Orc5) are known to bind ATP (Klemm et al. 1997). Four ORC subunits (Orc1, -2, -4, and -5) also contain potential winged-helix domains at their carboxyl termini. The smallest ORC subunit, Orc6, is related to the RNA Pol II auxiliary factor TFIIIB in metazoans (Liu et al. 2011), although fungal Orc6 lacks this similarity. Like Mcm2–7 and ORC, Cdc6 is an AAA$^+$ protein in the same initiator clade as the AAA$^+$ ORC subunits, DnaA and DnaC (Iyer et al. 2004). In addition, the carboxy-terminal portion of Cdc6 folds into a winged-helix domain (Liu et al. 2000). Both Orc6 and Cdt1 are poorly conserved across eukaryotes, but all Cdt1

Figure 2. Structure of the Mcm2–7 complex. (A) The Mcm2–7 proteins assemble into a hexameric toroid. Each Mcm2–7 complex includes one copy of the Mcm2–7 proteins that are arranged in the indicated order around a central channel. (B) Directionality of Mcm2–7 movement. The Mcm2–7 proteins move in a $3’ \to 5’$ direction along ssDNA. By analogy with the archaea homologs, the carboxy-terminal AAA$^+$ motif is proximal to the replication fork.
proteins carry a pair of winged-helix domains in their carboxy-terminal half (Lee et al. 2004; Khayrutdinov et al. 2009; Jee et al. 2010) and the most carboxy-terminal winged helix binds to Mcm2–7 (Yanagi et al. 2002; Teer and Dutta 2008; You and Masai 2008; Jee et al. 2010). Interestingly, structural studies suggest that the winged-helix domains associated with ORC, Cdc6, and Cdt1 are related (Khayrutdinov et al. 2009; Jee et al. 2010). Mcm8 and Mcm9 are Mcm2–7-related proteins found in metazoan cells, but are not found to associate with any of the Mcm2–7 proteins either in or out of the Mcm2–7 complex (Maiorano et al. 2005; Lutzmann and Méchali 2008). Although these proteins have been implicated in helicase loading (Volkening and Hoffmann 2005; Lutzmann and Méchali 2008), this view remains controversial (Blanton et al. 2005; Hartford et al. 2011) and a direct mechanistic involvement in helicase loading has not been shown.

INITIAL RECRUITMENT OF THE HELICASE TO THE ORIGIN DNA

Before helicase loading, ORC, Cdc6, Cdt1, and Mcm2–7 are recruited to the origin DNA. This complex is short-lived and is only detected in vitro when Cdc6 ATP hydrolysis is inhibited (Randell et al. 2006). In vivo, chromatin immunoprecipitation assays only detect ORC and Mcm2–7 at specific origins (Aparicio et al. 1997; Tanaka et al. 1997), presumably because of the rapid release of Cdc6 and Cdt1 after helicase loading.

ORIGIN RECOGNITION

The first step of helicase loading is the binding of ORC to origin DNA. In the yeast S. cerevisiae, ORC recognizes a conserved sequence within the origin of replication and this event is ATP-dependent (Bell and Stillman 1992). This sequence is not sufficient to direct ORC binding in vivo. An important additional determinant is an asymmetric nucleosome-free region surrounding the consensus sequence (Eaton et al. 2010). S. pombe ORC shows limited sequence specificity that is mediated by the AT-rich DNA-binding specificity of an AT-hook domain within the Orc4 subunit (Chuang and Kelly 1999; Kong and DePamphilis 2001; Lee et al. 2001).

The determinants for ORC binding in metazoan cells are less clear. ORC from metazoan cells does not exhibit sequence-specific DNA binding in vitro (Vashee et al. 2003; Remus et al. 2004), but ORC is localized to specific sites along metazoan chromosomes, indicating the existence of other determinants for ORC localization (Karnani et al. 2010; MacAlpine et al. 2010). Genome-wide studies of metazoan ORC binding support a number of determinants including nucleosome-depleted regions, promoter proximity, and dynamic nucleosomes (reviewed in Ding and MacAlpine 2011). In addition, DNA-binding studies show that ORC prefers to bind negatively supercoiled DNA (Remus et al. 2004; Houchens et al. 2008), which may be related to the wrapping of DNA around ORC (Clarey et al. 2006, 2008). Perhaps the most intriguing potential mechanism of ORC localization is the recognition of specifically modified nucleosomes. Human ORC binds a protein called ORCA or LRWD1 that interacts with methylated nucleosomes that are associated with heterochromatin, suggesting an alternative method to localize ORC to compacted chromatin (Bartke et al. 2010; Shen et al. 2010; Vermeulen et al. 2010). In addition, the amino terminus of the Orc1 protein contains a conserved bromo-adjacent homology (BAH) domain. In metazoan cells this domain interacts with histone H4 tails methylated at lysine 20, suggesting another mechanism for ORC to be localized along the chromatin (Kuo et al. 2012).

INTERACTION OF Cdc6, Cdt1, AND Mcm2–7

As cells enter G1 phase, ORC recruits Cdc6, Cdt1, and Mcm2–7 to the origin DNA. Biochemical studies support a model in which ORC first interacts with Cdc6 and this complex then recruits Cdt1 and Mcm2–7. Loss of Cdt1 does not prevent Cdc6 chromatin association in vivo or origin recruitment in vitro (Maiorano et al. 2000; Nishitani et al. 2000; Tsuyama et al. 2005; Randell et al. 2006; Remus et al. 2009). In contrast,
elimination of Cdc6 interferes with Cdt1 origin binding in vitro (Randell et al. 2006; Remus et al. 2009). Studies in Xenopus extracts have shown that Cdt1 associates with chromatin in the absence of Cdc6 (Gillespie et al. 2001; Tsuyama et al. 2005); however, only Cdt1 associated in the presence of Cdc6 is able to contribute to Mcm2–7 loading (Tsuyama et al. 2005). Consistent with a robust interaction between Cdc6 and ORC, a complex between the proteins has been structurally characterized (Speck et al. 2005; Sun et al. 2012) and origin-bound ORC stimulates Cdc6 ATP hydrolysis (Randell et al. 2006).

Mcm2–7 and Cdt1 are recruited to the origin as a complex. The carboxy-terminal winged-helix domain of Cdt1 binds Mcm2–7 (Yanagi et al. 2002; Ferenbach et al. 2005; Teer and Dutta 2008; You and Masai 2008; Zhang et al. 2010; Takara and Bell 2011). Cdt1 mutants lacking this region prevent or dominantly inhibit Mcm2–7 loading (Ferenbach et al. 2005; Zhang et al. 2010; Takara and Bell 2011). In S. cerevisiae cells, Cdt1 and Mcm2–7 are imported to the nucleus as a complex (Tanaka and Diffley 2002), and formation of this complex is necessary for the recruitment of either protein to the origin (Takara and Bell 2011). The primary Cdt1-binding site on Mcm2–7 is the carboxy-terminal region of Mcm6, and structural studies suggest that this region also folds into a winged-helix domain (Yanagi et al. 2002; Wei et al. 2010).

Characterization of the helicase loading intermediate composed of ORC, Cdc6, Cdt1, and Mcm2–7 suggests how a single asymmetric ORC protein can direct the assembly of a bi-directional replication fork. Binding of Cdt1/ Mcm2–7 to the origin requires the smallest ORC subunit, Orc6, which includes two binding sites for Cdt1 (Chen et al. 2007). Analysis of the protein complex formed at the origin before helicase loading (when Cdc6 ATP hydrolysis is inhibited) indicates that multiple Cdt1 proteins and their associated Mcm2–7 complexes are initially recruited to each ORC–Cdc6 complex (Takara and Bell 2011). It is most likely that this complex contains two Cdt1 molecules, one bound to each of the two Cdt1-binding sites on Orc6. Importantly, interfering with the formation of this multi-Cdt1 intermediate prevents subsequent helicase loading, and it is likely that this intermediate facilitates the simultaneous loading of an Mcm2–7 double hexamer (see below).

How Cdc6 contributes to Cdt1 and Mcm2–7 recruitment is less well understood. It is possible that there are direct interactions between Cdc6 and Cdt1 and/or Mcm2–7. Alternatively, Cdc6 could alter the conformation of ORC to allow Cdt1 and Mcm2–7 to bind. The latter hypothesis would explain why ORC does not bind either Cdt1 or Mcm2–7 in the absence of Cdc6 (Takara and Bell 2011) but isolated Orc6 can bind Cdt1 (Chen et al. 2007). Consistent with this model, recent EM studies of ORC bound to DNA in the presence and absence of Cdc6 suggested that Orc6 becomes exposed upon Cdc6 binding (Sun et al. 2012).

**LOADING OF RECRUITED HELICASES**

After recruitment of the helicase loading factors and Mcm2–7 to the origin, loading of the recruited Mcm2–7 requires ATP hydrolysis and involves significant changes of both the proteins associated with the origin and the structure of the Mcm2–7 ring. Loaded Mcm2–7 complexes form double-hexameric pairs that encircle dsDNA. Thus, helicase loading necessarily requires the establishment of strong interactions between the amino termini of two Mcm2–7 complexes and opening and closing of the Mcm2–7 ring to allow DNA access to the Mcm2–7 central channel. Finally, to ensure the proper regulation of replication initiation, as initially loaded, the Mcm2–7 complexes are inactive for unwinding the parental duplex.

**STRUCTURE OF LOADED Mcm2–7**

The reconstitution of Mcm2–7 loading using four purified proteins has led to important advances in the understanding of the architecture of loaded Mcm2–7. EM of loaded Mcm2–7 shows head-to-head double hexamers, with the amino termini of the Mcm2–7 subunits mediating hexamer interactions (Evrin et al. 2009; Remus et al. 2009). This structure is...
similar to the double hexamers observed for the homohexameric archaea Mcm complexes (Brewster and Chen 2010). Unlike the archaea homologs, Mcm2–7 double hexamers are only detected after loading. Mcm2–7 double hexamers survive treatment with DNase and gel filtration, indicating that once formed, these complexes are very stable (Evrin et al. 2009; Gambus et al. 2011). Based on structural studies of the archaea Mcm complex, these interactions likely involve the Zn-finger domains in the amino termini of Mcm2–7 subunits (Fletcher et al. 2003). The inter-Mcm2–7 contacts within the double hexamer are unknown but must be largely heterotypic owing to their head-to-head interaction. A lack of loaded single hexamers and the presence of multiple Cdt1/Mcm2–7 complexes before Mcm2–7 loading supports a model in which both hexamers are loaded in a concerted fashion (Remus et al. 2009; Takara and Bell 2011). Importantly, both EM and topological linkage studies indicate that loaded Mcm2–7 complexes encircle and slide non-directionally on dsDNA (Evrin et al. 2009; Remus et al. 2009). Thus, loaded Mcm2–7 is topologically linked to the DNA but is neither active as a helicase nor tightly engaged with the DNA.

ROLE OF ATP DURING HELICASE LOADING

ATP binding and hydrolysis play critical roles during helicase loading. As described above, ATP binding but not hydrolysis is required for the initial recruitment of the helicase to the origin (Gillespie et al. 2001; Randell et al. 2006). ATP hydrolysis by Cdc6 and ORC plays distinct and ordered roles in helicase loading. Cdc6 is activated to bind and hydrolyze ATP when it associates with origin-bound ORC, and this hydrolysis event is required to observe Mcm2–7 loading (Perkins and Diffley 1998; Randell et al. 2006). ORC ATP hydrolysis is also essential but functions after Cdc6 hydrolysis and appears to regulate repeated Mcm2–7 loading (Bowers et al. 2004; Randell et al. 2006). One likely role for ORC ATP hydrolysis is to cause the release of the other helicase loading factors from the origin (Tsakraklides and Bell 2010). Evidence from studies of S. cerevisiae suggests that a second consequence of ATP hydrolysis is to disable ORC DNA binding until it is able to bind another molecule of ATP. An important unanswered question is how many Mcm2–7 hexamers are loaded per round of ATP hydrolysis by Cdc6 and ORC. Although current evidence suggests two hexamers are loaded coordinate-ly, it remains possible that Cdc6 or ORC could require multiple rounds of ATP hydrolysis to accomplish the loading of an Mcm2–7 double hexamer. Overall, the coordinated and ordered set of ATP hydrolysis reactions ensures that Cdc6 only acts on Cdt1/Mcm2–7 when bound to ORC at the origin and a new round of loading cannot occur until the previous round is complete.

Mcm2–7 GATE FUNCTION AND IMPLICATIONS FOR LOADING

The Mcm2–7 ring structure must be opened during loading to provide access to the central DNA-binding channel. Binding of purified Mcm2–7 to ssDNA circles suggests that the interface between Mcm2 and Mcm5 acts as a “gate” to the central channel (Bochman and Schwacha 2008). EM studies of purified Drosophila Mcm2–7 support this conclusion (Costa et al. 2011). Although ATP binding is believed to close the Mcm2/5 gate, it is likely that some combination of Cdt1 binding and Cdc6 ATP hydrolysis modulates Mcm2–7 ring opening during helicase loading. Because Cdt1 is released from Mcm2–7 after loading (Randell et al. 2006), it is tempting to speculate that Cdt1 binding opens the Mcm2/5 gate and Cdc6 ATP hydrolysis separates Cdt1 from Mcm2–7, leaving behind a closed form of the Mcm2–7 ring. The role of ATP binding and hydrolysis by Mcm2–7 during helicase loading remains uncertain despite studies to address this issue. Mutations in the Mcm6 and Mcm7 ATPase active sites (reviewed in Bochman and...
Schwacha 2009) suggest that ATP bound to other sites could be required for helicase loading.

REGULATION AND DYNAMICS OF HELICASE LOADING

Helicase loading is tightly restricted to G1 phase to ensure that no origin of replication can initiate DNA replication more than once per cell cycle (reviewed in Arias and Walter 2007; Siddiqui et al. 2013; Zielke et al. 2013). This regulation is mediated by the inhibition of helicase loading outside of G1 phase. The mechanisms of inhibition are diverse and vary among organisms. In budding yeast, inhibition is primarily mediated by cyclin-dependent kinase (CDK) phosphorylation of Cdc6, which leads to its degradation (Drury et al. 2000), Mcm2–7 nuclear export (Labib et al. 1999; Nguyen et al. 2000), and the inhibition of Cdt1/Mcm2–7 binding to ORC (Chen and Bell 2011). In metazoan organisms, Cdt1 is a primary target for inhibition. Cdt1 is inhibited by geminin binding outside of G1 (McGarry and Kirschner 1998; Wohlschlegel et al. 2000) and degraded in a proliferating cell nuclear antigen (PCNA)/DNA-dependent (Arias and Walter 2005) or CDK-dependent (Li et al. 2003; Liu et al. 2004) fashion in S phase. In addition, in many metazoans, CDK activity inhibits ORC DNA binding during G2/M phase (reviewed in DePamphilis 2005).

Although it is tempting to think that the multiple mechanisms that inhibit helicase loading are redundant, analysis of mutants that are defective for a subset of mechanisms shows partial replication (Green et al. 2006). Importantly, even limited replication is lethal to most dividing cells and can induce gene amplification associated with cancer (Green et al. 2010).

The role of the helicase loading proteins after loading is unclear. In vitro studies show that after loading, Mcm2–7 no longer requires ORC, Cdc6, or Cdt1 to associate with origin DNA (Donovan et al. 1997; Randell et al. 2006). Consistent with this observation, in vitro studies show that the helicase loading proteins are released from the origin upon Mcm2–7 loading (Randell et al. 2006; Tsakraklides and Bell 2010). In contrast to these findings, in vivo activation of S. cerevisiae ORC or Cdc6 in late G1 (after Mcm2–7 loading) results in loss of Mcm2–7 origin association (Aparicio et al. 1997; Semple et al. 2006; Chen et al. 2007). This suggests either that these proteins are required to maintain Mcm2–7 association with the origin DNA or that there is an activity that removes loaded Mcm2–7 from the DNA in G1 cells and the helicase loading proteins are required to restore Mcm2–7 to the origin. In either case, this reliance on helicase loading proteins changes upon entry into S phase, as Mcm2–7 is tightly associated with origins at this stage (unless initiation has occurred) even though Cdc6 is eliminated from S phase cells by degradation. It is noteworthy that studies in mammalian G1 cells suggest that there is little exchange between chromatin-associated and free Mcm2–7 (Kuiipers et al. 2011), arguing for a model in which helicase loading proteins stabilize loaded Mcm2–7 rather than a model in which Mcm2–7 release is balanced with new loading.

A MODEL FOR HELICASE LOADING

Based on current knowledge, we propose the following model for the events of eukaryotic helicase loading (Fig. 3). ORC recruitment to origin DNA is directed by a combination of DNA affinity (sequence-specific in budding yeast), local chromatin structure, and interaction with other proteins (e.g., ORC). Upon entry into G1 phase, ORC recruits Cdc6. In S. cerevisiae cells, we propose that this interaction reveals Cdt1-binding sites on Orc6 that are bound by two Cdt1/Mcm2–7 complexes. The distinct structure of Orc6 in metazoans raises the possibility that the initial recruitment is mediated by different interactions in these organisms. The connection of these events to Mcm2–7 ring opening is unclear; however, it seems most likely that the Mcm2/5 gate is opened at some point during these events. Alternatively, the ring could be opened before recruitment (e.g., after Cdt1 binds to Mcm2–7), although this would risk Mcm2–7 interacting with nonorigin DNA. Assuming two Mcm2–7 complexes are loaded coordinately, we suggest that the Mcm2/5 gates of the two helicases are aligned to form a single continuous
opening/gate. By analogy to the function of the sliding clamp loader (Kelch et al. 2011), we speculate that the Mcm2–7 complexes can encircle the DNA at this stage but remain open. The next stage of loading is triggered by Cdc6 ATP hydrolysis. This event leads to the release of Cdc6 and Cdt1 from the helicase, which is coupled with enrapment of the DNA within the Mcm2–7 central channel and closure of the Mcm2/5 gate. Finally, we propose that ATP hydrolysis by ORC leads to the release of the loaded helicases from ORC to prepare for a new round of loading after ADP → ATP exchange.

Once loaded, the Mcm2–7 complex awaits activation. Biochemical studies show that activation requires the association of two helicase activating proteins (Cdc45 and GINS) with Mcm2–7 (Ilves et al. 2010). Association of these proteins is triggered by the action of two kinases: S-phase CDK (S-CDK) and the Dbf4-dependent Cdc7 kinase (DDK). DDK phosphorylates Mcm2–7 (Sheu and Stillman 2006, 2010; Randell et al. 2010). As a result of Mcm2–7 phosphorylation, Cdc45 associates with Mcm2–7 (Heller et al. 2011; Tanaka et al. 2011). Interestingly, this event occurs at a subset of origins before entry into S phase (Aparicio et al. 1999; Kanemaki and Labib 2006; Heller et al. 2011; Tanaka et al. 2011). S-CDK targets two additional proteins, Sld2 and Sld3, causing them to associate with Dpb11 and eventually resulting in the recruitment of the GINS proteins as well as DNA Pol ε (Tanaka et al. 2007; Zegerman and Diffley 2007; Muramatsu et al. 2010). The subsequent activation of DNA unwinding is required for the recruitment of the lagging-strand DNA polymerases (DNA Pol α-primase and DNA Pol δ (Mimura et al. 2000; Heller et al. 2011).

COMPARISON OF HELICASES AND HELICASE LOADING IN E. coli AND EUKARYOTIC CELLS

There are many similarities between helicase loading at E. coli and eukaryotic chromosomal origins, suggesting that the fundamental mechanism of these events is related. Both E. coli DnaB and eukaryotic Mcm2–7 are composed...
of six subunits that assemble into ring-shaped molecules. Second, both events rely on origin recognition proteins (DnaA and ORC) that contact the origin DNA over an extended region. Third, loading of each DNA helicase requires at least one additional helicase loading factor that binds to the helicase. In E. coli, DnaC is required for loading and also restrains DnaB helicase activity. In eukaryotes, both Cdc6 and Cdt1 share properties with DnaC. Like DnaC, Cdc6 is an AAA$^+$ protein and is a weak ATPase on its own, whereas Cdt1 shares DnaC’s association with its cognate helicase. Finally, both helicases must be activated after loading, either by release of an inhibitor (DnaC in E. coli) or by recruiting helicase activators (Cdc45 and GINS in eukaryotic cells).

There are also a number of important differences. First, although similar in overall structure, DnaB and the Mcm2–7 complexes have opposite polarities: DnaB moves on the lagging-strand template, whereas Mcm2–7 moves on the leading-strand template. DnaB is a homohexamer that has a RecA-like ATPase domain (Bailey et al. 2007), whereas Mcm2–7 is a heterohexamer that has AAA$^+$–like ATPase domains (Iyer et al. 2004). Second, although DnaA origin binding leads to local DNA melting, ORC DNA binding does not. Third, two separate single hexamers of DnaB are loaded around ssDNA after loading, whereas a head-to-head double hexamer of Mcm2–7 is loaded around dsDNA.

These differences are most likely related to the rate of helicase activation after loading. Loading of DnaB around ssDNA at oriC is consistent with the rapid activation of DnaB after loading. In contrast, the need to restrain helicase activation until S phase in eukaryotic cells suggests the need for a loaded helicase that is less easily activated. Recent findings indicate that Mcm2–7 surrounds ssDNA and acts as a single hexamer at the replication fork (Yardimci et al. 2010; Fu et al. 2011). This indicates that Mcm2–7 activation requires two major changes in the loaded double hexamer. First, each Mcm2–7 complex must dissociate from its double-hexamer partner. Second, the Mcm2–7 ring originally bound to duplex DNA must open to expel the lagging-strand ssDNA template before DNA unwinding. Although we have an initial understanding of how DnaC expulsion leads to activation of DnaB (Makowska-Grzybska and Kaguni 2010), how the helicase activating proteins Cdc45 and GINS and the proteins that direct them to the origin (Sld2, Sld3, and Dpb11) catalyze these events remains unknown.

Although 25 years of research has provided an increasingly detailed understanding of the events of helicase loading, fundamental questions remain. How is the opening and closing of the DNA helicase ring controlled? How is the origin DNA properly positioned to ensure that the helicase encircles the appropriate form of the DNA at the correct position? How are these events regulated by the AAA$^+$ ATPases that are conserved among the different realms of life? These and other questions propel studies of this central event in replication initiation.

ACKNOWLEDGMENTS

We thank members of our labs for helpful discussions. This work is supported by the Michigan Agricultural Station and by grant R01 GM090063 from the National Institutes of Health to J.M.K. and by the Howard Hughes Medical Institute and grant R01 GM052339 from the National Institutes of Health to S.P.B.

REFERENCES

Helicase Loading at Chromosomal Origins


Maiorano D, Cuvier O, Danis E, Méchali M. 2005. MCM8 is an MCM2-7-related protein that functions as a DNA helicase during replication elongation and not initiation. *Cell* **120**: 315–328.


Helicase Loading at Chromosomal Origins


Takara TJ, Bell SP. 2011. Multiple Cdt1 molecules act at each origin to load replication-competent Mcm2–7 helicases. EMBO J 30: 4885–4896.


