Annu. Rev. Biochem. 2002. 71:333–74
DOI: 10.1146/annurev.biochem.71.110601.135425

First published as a Review in Advance on February 12, 2002



# DNA Replication in Eukaryotic Cells

Stephen P. Bell[1] and Anindya Dutta[2]
[1]*Howard Hughes Medical Institute, Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts, 02139; e-mail: spbell@mit.edu*
[2]*Department of Pathology, Brigham and Women's Hospital, Harvard Medical School, Boston, Massachusetts, 02115; e-mail: adutta@rics.bwh.harvard.edu*

**Key Words** cell cycle, origin recognition complex (ORC), CDC6, MCM proteins, MCM10, CDT1, CDC7, DBF4, CDC45, pre-replicative complex (pre-RC), cyclin-dependent kinases (CDK)

■ **Abstract** The maintenance of the eukaryotic genome requires precisely coordinated replication of the entire genome each time a cell divides. To achieve this coordination, eukaryotic cells use an ordered series of steps to form several key protein assemblies at origins of replication. Recent studies have identified many of the protein components of these complexes and the time during the cell cycle they assemble at the origin. Interestingly, despite distinct differences in origin structure, the identity and order of assembly of eukaryotic replication factors is highly conserved across all species. This review describes our current understanding of these events and how they are coordinated with cell cycle progression. We focus on bringing together the results from different organisms to provide a coherent model of the events of initiation. We emphasize recent progress in determining the function of the different replication factors once they have been assembled at the origin.

## CONTENTS

## INTRODUCTION

Complete and accurate DNA replication is integral to the maintenance of the genetic integrity of all organisms. In eukaryotic cells, this event is initiated at hundreds, if not thousands, of chromosomal elements called origins of replication. These sequences direct the assembly of multiprotein machines that eventually form two replication forks at each origin. Recent studies of the assembly and activation of these complexes have provided important new insights into eukaryotic DNA replication and how it is coordinated with other events of the cell cycle.

We focus on the events occurring at eukaryotic origins of replication. The sequences required for an origin of replication vary significantly between different eukaryotic organisms. In the unicellular eukaryote *Saccharomyces cerevisiae*, three to four sequences of 10–15 base pairs (bp) spread over 100–150 bp are sufficient to act as an origin. These sequences include the highly conserved and essential A-element or autonomously replicating sequence (ARS) consensus sequence (ACS) and less well conserved elements called B elements, which among other things are likely to provide a DNA unwinding element to the origin (reviewed in 1). In other organisms the *cis*-acting sequences required to direct the initiation of DNA replication are more complex. In the fission yeast, *Schizosaccharomyces pombe*, sequences spread over at least 800 to 1000 bp direct initiation (2–4). Detailed analysis of these sequences identified several AT-rich sequences of 20–50 bp that are important for origin function, but they do not exhibit the strong sequence similarity observed for the *S. cerevisiae* ACS. These sequences also appear to exhibit substantial internal redundancy, complicating their analysis. Metazoan origins are still less well defined and can extend over thousands of base pairs of DNA (reviewed in 1, 5). In addition, the sites of initiation are not always tightly linked within these regions. Although a handful of origins have been identified, the definition of *cis*-acting sequences required for their function has been difficult. The origins controlling the amplification of chorion DNA in *Drosophila* follicle cells and the replication of the *β-globin* gene cluster have been defined at the level of specific *cis*-acting sequences, although these sequences are relatively large compared to those defined in yeast. A number of other origins have been identified by physical analysis of replication intermediates, but less is known about specific sequences required for their function.

Finally, the early embryos of *Drosophila melanogaster* and *Xenopus laevis* are at the other extreme of origin sequence definition. Origins in these early embryos appear to require little or no sequence specificity, presumably to allow for extremely rapid S phase (reviewed in 6). Thus, although it is clear that some sites consistently act as origins of replication in the majority of eukaryotic cells, the mechanisms that select these sites and the sequences that determine their location remain elusive in many cell types. For a full discussion of eukaryotic origin structure see the reviews referenced above.

Eukaryotic origins of replication direct the formation of a number of protein complexes leading to the assembly of two bidirectional DNA replication forks. These events are initiated by the formation of the pre-replicative complex (pre-RC) at origins of replication during G1 (Figure 1). Pre-RC formation involves the ordered assembly of a number of replication factors including ORC, Cdc6p, Cdt1p, and Mcm2–7p. The regulation of pre-RC formation is a key element of the mechanisms coordinating DNA replication with the cell cycle. Once formed, this complex awaits activation by at least two kinases that trigger the transition to DNA replication. As with the formation of the pre-RC, the transition to replication involves the ordered assembly of additional replication factors that facilitate unwinding of the DNA at the origin (Figure 2) and culminate in the association of the multiple eukaryotic DNA polymerases with the unwound DNA (Figure 3). At the end of the review, we briefly address the control of replication timing and intra-S-phase checkpoint control as they affect replication initiation. We have not covered a number of other important areas of DNA replication function and control including the action of replication fork proteins during the elongation stages of DNA replication (reviewed in 7) and an extensive discussion of the proteins involved in checkpoint control (reviewed in 8, 9).

Throughout this review we have generally used a single name for each replication factor described. In numerous cases, functionally related factors from different species have been given distinct names prior to the determination that they were related. In general, we have chosen to use the name of the first factor identified. To distinguish between analogous factors from different species we have added a prefix composed of the first letters of the genus and species name of the organism that the factor is derived from (e.g. *Xenopus laevis* Cdc6p is indicated as XlCdc6p).

## COMPONENTS OF THE PRE-REPLICATIVE COMPLEX

### The Origin Recognition Complex

The origin recognition complex (ORC) is a six-subunit complex that acts as the initiator (the protein that selects the sites for subsequent initiation of replication) at eukaryotic origins of replication. Although identified in *S. cerevisiae* as

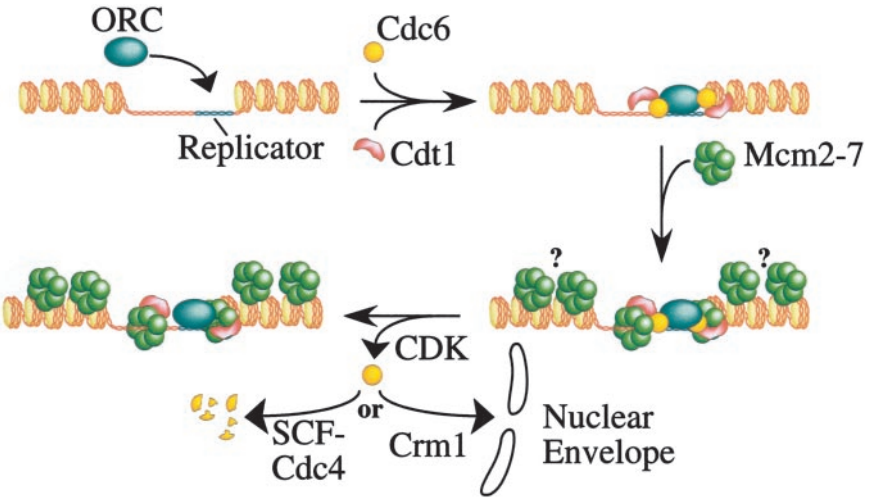


**Figure 1** A model for pre-replicative complex formation accommodating much but not all of the current information concerning pre-RC formation in eukaryotes. The stoichiometry of the different components is unknown. The apparent overabundance of Mcm2–7p relative to other components is illustrated as additional Mcm2–7p complexes associated with adjacent chromatin; however, the location relative to nucleosomes and the distribution adjacent to the origin have not been determined. The fate of Cdc6p after pre-RC formation is distinct in different organisms; both possibilities are illustrated. See text for further details.



binding to the conserved ACS (10), subsequent studies have found that this complex is a conserved feature of chromosomal replication in all eukaryotes studied. Studies in *Xenopus* egg extracts have demonstrated that an analogous XlORC is required for initiation of replication (11–13). Similarly, recessive lethal mutations in multiple *Drosophila melanogaster* ORC subunits (*DmORC2*, *DmORC3*, *DmORC5*) each show dramatic reductions of BrdU incorporation in third instar larva (14, 15). Hypomorphic alleles of *DmORC2* have defects in chorion gene amplification which requires multiple replication initiation events from a specific subset of origins without intervening M phases (16). The demonstration that both purified and recombinant DmORC can be used to reconstitute DNA replication in *Drosophila* extracts immunodepleted for DmORC provides biochemical support for these genetic findings (17). Although direct evidence that mammalian ORC functions in the initiation of replication is still lacking, recent studies indicate that replication from OriP of Epstein-Barr virus (EBV) requires ORC function in human cells (18).

In the last five years considerable progress has been made in our understanding of ORC, particularly in the identification and characterization of analogs derived from metazoans. In addition, the function of ORC in the formation of the

pre-RC has been refined. We first discuss the basic functions of ORC during DNA replication, including DNA recognition and its control by ATP. We then describe recent studies that indicate ORC function may be more dynamic in metazoans than previously thought and that additional factors may assist ORC in localizing to origins of replication in vivo.

DNA BINDING The best-understood activity of ORC is its ability to bind DNA. Because of the availability of a known binding site, the characterization of ORC binding to DNA is most advanced for *S. cerevisiae.* ScORC interacts specifically with both the A and B1 elements of yeast origins of replication, spanning a region of ~30 bp (10, 19, 20). Interestingly, binding to these sequences requires ATP (see below). More precise studies of this interaction suggest ScORC interacts primarily with the A-rich strand of this region (21). The binding sites for other ORC complexes are less clear. Both in vivo and in vitro studies indicate that DmORC binds the critical elements of the amplification control elements on Chromosome III (*ACE3* and *ori-β*) (22, 23). It has been hypothesized that the sequence specificity of DmORC is limited to AT-rich DNA (23), although this proposal has not been rigorously tested and sequences of very similar AT-content show significant differences in affinity (22). Similar studies have identified binding sites for the ORC at the amplification origin (ori) II/9A in the fly *Sciara coprophila* in vivo (24). In vitro studies using DmORC show binding to a specific 80-bp region adjacent to the start site of DNA replication (24). Despite this progress, a precise DNA binding site has yet to be identified at either the *Drosophila* or *Sciara* amplification origins. Chromatin immunoprecipitation (ChIP) studies have demonstrated the association of SpORC with *S. pombe* origins (25) and human ORC with the EBV OriP (18, 26, 27).

The complex nature of the interaction of ORC with DNA has made identification of a discrete DNA binding domain difficult. ScORC requires the five largest subunits (Orc1p–Orc5p) to recognize DNA, four of which (Orc1p, 2p, 4p, and 5p) are in close contact with the origin (21). Although Orc6p is not required for DNA binding, it remains essential for DNA replication and cell viability (28). The situation in *D. melanogaster* is somewhat different, as both DNA binding and DNA replication requires DmOrc6p (23). Unlike ScORC and DmORC, which lack clear DNA binding motifs (however, see the discussion of Cdc6p structure below), ORC derived from *S. pombe* (SpORC) has a repeated AT-hook DNA binding motif at the N terminus of SpOrc4p (29). Although not found in other Orc4p analogs, this element is essential for *S. pombe* viability (T. Kelly, personal communication). Previous studies of this motif have found that AT-hook domains interact preferentially with AT-rich DNA. Biochemical studies of SpOrc4p and SpORC show that both recognize the same specific regions of multiple *S. pombe* origins of replication (29, 29a, 29b). The DNA sequences identified have a strong propensity to have stretches of poly-A, similar to the very A-rich nature of the *S. cerevisiae* ACS. Consistent with this, SpORC lacking this domain of SpOrc4p shows reduced affinity for chromatin in vivo (30).

Because all origins must unwind during replication initiation, it is intriguing that ORC also shows affinity for single-stranded DNA (ssDNA). Comparative studies indicate that ORC binds ssDNA with a $K_d$ that is within threefold of the affinity of RPA, the primary eukaryotic ssDNA binding factor (31). Interestingly, this interaction is not influenced by either the sequence of the ssDNA or the presence of ATP. Instead, the primary determinant for the affinity of ORC for ssDNA is length: ssDNAs shorter than 30 bases have little or no affinity and ssDNAs greater than 80–85 bases have the highest affinity. As with ScORC binding to double-stranded DNA (dsDNA) at origins, the domain of ScORC required for ssDNA binding is unknown but requires Orc1–5p. In addition, competition studies indicate that the dsDNA and ssDNA binding sites overlap or require mutually exclusive conformations (31).

ATP REGULATION OF ORC ORC also binds and hydrolyzes ATP, and these functions have a significant influence on ORC function. Studies of both ScORC and DmORC indicate the ATP binding by Orc1p is required for DNA binding (10, 22, 23, 32). ScORC and DmORC also hydrolyze ATP, and this activity also depends on the Orc1p subunit; however, ATP hydrolysis is not required for DNA binding (23, 32). Consistent with this observation, origin DNA inhibits ATP hydrolysis by ScORC in a sequence-dependent manner. Wild-type origin DNA inhibits the ScORC ATPase ~10-fold, whereas origin DNA with mutations that inhibit ScORC binding show little or no inhibition (32). Similarly, *ori-β* DNA from the chorion amplification loci induces a two- to threefold decrease in DmORC ATPase activity (23). These findings suggest that once bound to the origin, ORC is retained in an ATP-bound state and that ATP hydrolysis is reserved for a downstream step in initiation.

Unlike double-stranded origin DNA, ssDNA stimulates ATP hydrolysis by ORC in a length- (but not sequence-) dependent manner (31). This suggests the interesting possibility that DNA unwinding at the origin stimulates ATP hydrolysis by ORC. Mutants in ScOrc1p that allow ScORC to bind but not hydrolyze ATP are lethal when overexpressed in combination with the remaining wild-type ORC subunits (33). The same phenotype is not caused by over-expression of wild-type ScORC (which continues to hydrolyze ATP) or a mutant ScORC that cannot bind ATP at ScOrc1p, which suggests that the ATP-bound conformation of ScORC is important to induce lethality. Additional data suggests that this lethality is due to titration of Cdc6p away from the origin by the overexpressed ATP-bound ORC (33). Together these findings suggest that ORC needs to be in an ATP-bound state to interact with Cdc6p and that a downstream event that produces ssDNA (e.g., initiation) promotes ATP hydrolysis by ORC. If ATP hydrolysis results in the release of bound replication factors from ORC (like Cdc6p), such a chain of events could play a key role in the transition from initiation to elongation stages of replication. Further studies of reconstituted replication reactions will be required to test this hypothesis.

CONTROL OF ORC SUBUNIT ASSOCIATION Recent studies in mammalian cells suggest that not all ORC subunits remain tightly associated as part of the complex throughout the cell cycle. Unlike ORC from budding yeast, *Drosophila*, and *Xenopus*, the subunits of the SpORC and mammalian ORC are difficult to extract as a stable complex (30, 34, 35). For example, SpOrc4p is retained on chromatin under conditions that elute the remainder of SpORC (30). Similarly, whereas mammalian Orc2p is found constitutively on the chromatin, mammalian Orc1p is removed from the chromatin at the end of S phase and rebinds only as cells re-enter G1 (34, 36, 37). Recent studies suggest that HsOrc1 (in *Homo sapiens*) may be proteolyzed during S phase as a mechanism to prevent re-replication (36); however, other studies have found HsORC1p to be stable throughout the cell cycle (38, 39; T. Kelly, personal communication). Yet another study has observed that Hamster Orc1p is stable through the cell cycle but is regulated in its association with chromatin by cell cycle regulated ubiquitination (39a). These substantial differences are unlikely to be due to simple technical differences but instead might indicate variations in the regulation of this key factor in different cell lines.

Biochemical studies addressing the assembly of HsORC found that a core subcomplex of HsOrc2–5p was readily assembled in insect cells (40, 41). Interestingly, if the nuclei of the same cells are extracted in the presence of high salt, a six-subunit HsORC was isolated (41). Collectively, these findings suggest that the six-subunit HsORC may be assembled on the chromatin only during the G1 and S phase of the cell cycle. This hypothesis also suggests that a more limited complex of HsORC subunits is competent for chromatin binding, although the specificity of this interaction is unclear. It is also noteworthy that the different mammalian ORC subunits co-immunoprecipitate with each other (15, 42–45); however, the majority of these subunits are not stably associated in a single complex (35). Thus, it is possible that there is a small amount of intact ORC in mammalian cells that is sufficient for replication initiation and that this is difficult to detect in the cell extracts. The finding that the reduction of HsOrc2p to less than 10% of wild-type levels allows normal S-phase progression supports this hypothesis (18).

CONTROL OF ORC DNA BINDING DURING THE CELL CYCLE ORC binding to chromatin is regulated through the cell cycle in some but not all species. *S. cerevisiae* and *S. pombe* ORC appear to associate constitutively with origin sequences throughout the cell cycle (25, 46–48). Similarly, DmORC remains associated with chromosomes at all stages of the cell cycle, although differences in sample preparation have significant effects on the detection of DmORC epitopes during metaphase (49). In contrast, studies in *Xenopus* cells and egg extracts found that XlORC is cleared from the chromatin during metaphase as measured by immunofluorescence (11, 13). It is possible that epitope masking could be responsible for the observed metaphase clearance of XlORC from chromosomes; direct measurement of XlORC association with metaphase chro-

mosomes (e.g., using chromatin precipitation assays) is needed to address this issue. The regulation of mammalian ORC is consistent with the removal of at least part of the complex from the chromosome at metaphase, as several groups have detected that Orc1p chromatin association is diminished in mitosis (34, 36, 37). The removal of ORC could serve to eliminate pre-RC formation prior to the completion of M phase (see below). Alternatively, it is possible that ORC removal is a consequence of, or critical for chromatin condensation during mitosis.



OTHER FACTORS IN THE SELECTION OF ORC BINDING SITES Although ScORC is commonly thought to be directed to origins by the conserved 11-bp ACS, there is substantial degeneracy in this sequence, and biochemical studies of ORC DNA binding suggest that its ability to distinguish specific from nonspecific sequences is limited (R. Klemm & S. Bell, unpublished data). Nevertheless, whole-genome location analysis indicates that ScORC is localized to a limited number of sites in the genome (~460), typically separated by 20–30 kb and almost always within intergenic regions (49a). This frequency is significantly lower than the frequency observed for matches to the ACS, which suggests that other factors influence the localization of ORC in the genome. Recent studies of the interaction between ORC and Cdc6p suggest that Cdc6p may increase the stability of ORC DNA binding (50), and this interaction could increase the selectivity of ORC binding in vivo. ORC is also unable to interact with its binding site in the context of a nucleosome (J. R. Lipford & S. P. Bell, unpublished results), raising the possibility that nucleosomes can restrict the number of available binding sites for ScORC. Interestingly, origins that are active on episomes but weak or inactive in the chromosome frequently show reduced ORC association in the latter context (49a), which suggests that the chromosomal context alters ORC ability to associate with these origins. If so, factors affecting local chromatin structure such as promoter-associated transcriptional regulators could influence ORC DNA binding and origin function. In this context, it is also interesting to note that human ORC interacts with the histone acetyltransferase Hbo1p, which could provide a mechanism to alter the local nucleosome configuration (51, 52).

The difficulty in identifying well-defined ORC binding sites in species other than *S. cerevisiae* raises the possibility that other DNA binding factors facilitate ORC localization and origin selection. Strong support for this hypothesis is provided by studies of the localization of DmORC during chorion gene amplification. Mutations in the DNA binding domains of the *Drosophila* transcription factors dDp and dE2F1 reduce DmORC localization and chorion amplification (53). Besides facilitating the recruitment of DmORC to sites of amplification, dE2F1 recruits regulatory factors such as *Drosophila* Rb (Rbf), and there is genetic evidence that Rbf inhibits replication. dE2F1, dDp, Rbf, and DmORC are in a complex independent of DNA, which raises the possibility that they are corecruited to the origin of replication (54). Rbf is required for the interaction of

dE2F1 with DmORC; however, Rbf is not required to mediate the stimulation of ORC localization by dE2F1. This raises the interesting possibility that DNA binding by dE2F1 helps localize DmORC to its binding site(s) by indirect mechanisms such as changing local chromatin structure, rather than acting through direct recruitment. Similar to DmORC association with dE2F, HsORC has been shown to interact with the viral origin binding protein, EBNA1 (18, 27). It is not clear, however, whether this interaction serves to localize HsORC to OriP.

The mechanism controlling the selection of origins in *Xenopus* egg extracts (and presumably early *Xenopus* embryos) is significantly less specific; virtually any DNA sequence is replicated in this system (55). Despite the apparent lack of sequence specificity for origin selection, ORC appears to saturate the added DNA by binding on average once every 16 kb (12, 56). This finding suggests that even in early *Xenopus* embryos there are mechanisms that limit the sites of ORC association. Interestingly, a similar spacing is observed for the sites of initiation on templates added to *Xenopus* egg extracts (57, 58). The finding that ORC is limiting for the initiation of replication in *Xenopus* extracts and the apparently similar spacing of origins and ORC on the chromatin both support the hypothesis that ORC binding determines the location of initiation sites, as has been found in *S. cerevisiae*. The mechanism that limits the amount of ORC binding is unknown. It is possible that this reflects a periodicity in the chromatin that is assembled on the added DNA. Alternatively, the loading of other replication factors may selectively stabilize a subset of bound ORC molecules.

## The Cdc6 Protein

ScCDC6 was first identified in the original screen for *S. cerevisiae* mutants with changes in the cell division cycle (59). Cdc6p is a member of the large family of AAA+ ATPases and is highly related to Orc1p and, to a more limited extent, to Orc4p, Orc5p, and the Mcm2–7 proteins (60). Work in the last decade has demonstrated that Cdc6p plays a crucial role in the assembly of the pre-RC at a step after ORC and before the Mcm2–7 proteins. Regulation of Cdc6p activity plays a critical role in regulating the formation of the pre-RC during the cell cycle. Below we focus on how Cdc6p is regulated both through its interaction with nucleotide and by changes in its abundance and localization during the cell cycle.

ATP REGULATION OF Cdc6p Like ORC, Cdc6 is an ATP-binding protein that participates in the pre-RC. Cdc6p requires ORC to associate with chromatin and is in turn required for Mcm2–7p chromatin association (61). Direct association with the origin has been demonstrated by ChIP studies (48), and genetic studies in *S. cerevisiae* indicate Cdc6p is required for the Mcm2–7p, but not ORC, association with the origin (46). Mutations in the ATP binding motif of ScCdc6p strongly suggest that ATP binding and hydrolysis play a key role in ScCdc6p

function. Mutations in the Walker A-motif of ScCdc6p result in alleles that exhibit either reduced or no function in vivo (62–64). Analysis of a mutation in the Walker B-motif shows that it either has no effect (64) or results in a protein (Cdc6-d1p) that has reduced function when expressed in single copy but that is dominant lethal when overexpressed (63, 64). The mechanism of action of this dominant negative allele is unclear. Analysis by in vivo DNAse I footprinting suggests that the overexpression of the mutant Cdc6p allows limited pre-RC formation; however, this allele is defective in promoting the association of Mcm5p with chromatin (64). This apparent contradiction may be explained by the different sensitivities of the two assays for pre-RC formation or that the other Mcm2–7 proteins are still loaded in the presence of Cdc6-d1p. It is also not clear whether Cdc6-d1p acts dominantly by loading at the origin and preventing subsequent steps in pre-RC formation or by titrating one or more replication factors away from the origin. The Cdc6 analog in fission yeast, Cdc18 (hereafter referred to as SpCdc6), also requires intact nucleotide binding and hydrolysis motifs for DNA replication in vivo. Mutations expected to affect hydrolysis enter but cannot complete S phase, whereas mutations expected to inhibit ATP binding have defects in entry into S phase and the S-phase checkpoint (65, 66).

Studies of HsCdc6p support the idea that the function of this replication factor is dependent on its ability to bind and hydrolyze nucleotides. Recombinant HsCdc6p is the only Cdc6p analog that has been shown to bind and hydrolyze ATP in vitro (67). As is often the case for ATP binding proteins, association of HsCdc6p with nucleotide results in a conformational change as detected by limited proteolysis. Consistent with previous mutations of other proteins containing Walker A- or B-motifs, mutations in the Walker A-motif eliminate ATP binding, whereas mutations in the Walker B-motif allow ATP binding but not ATP hydrolysis. Interestingly, microinjection of these mutant alleles resulted in the inhibition of DNA replication but only when injected during the G1 phase of the cell cycle.

**CELL CYCLE–REGULATED PROTEOLYSIS AND LOCALIZATION OF Cdc6p** The abundance of yeast Cdc6p is regulated during cell cycle progression. As cells pass into S phase, ScCdc6p is targeted for degradation by $SCF^{CDC4}$-dependent ubiquitinylation and is subsequently degraded by the proteosome (68–71). Biochemical experiments support a ubiquitin-dependent degradation of ScCdc6p following cyclin-dependent kinase (CDK) phosphorylation. A reconstituted ubiquitinylation assay demonstrated that Cdc6p was ubiquitinylated in a Cdc4p-, Cdc34p-, and Clb5p/Cdc28p-dependent manner (69). Similarly, SpCdc6 is also targeted for degradation by the $SCF^{CDC4}$ analogs SpPop1 and SpPop2/SpSud1p, which act together as a complex (72, 73). Mutations in several consensus sites for CDK phosphorylation at the N terminus of both SpCdc6p and ScCdc6p inhibit the degradation, which suggests that CDK activity triggers this proteolysis (69, 70, 72, 74, 75). Indeed, the phosphorylation of the SpCdc6p at the N terminus is required for in vivo binding to SpPop2/SpSud1p (72).

In mammalian cells, Cdc6p activity is controlled by a different mechanism. As cells enter S phase, phosphorylation of the HsCdc6p by CDK promotes the export of the protein from the nucleus (39, 76–78). Mechanistically, the phosphorylation at the N terminus of HsCdc6p appears to expose a nuclear export sequence at the C terminus but has no effect on the nuclear localization sequence (79). In mitosis, mammalian Cdc6p is degraded following ubiquitinylation by the anaphase promoting complex (APC) (78, 80). This is mediated by N-terminal destruction motifs (both a KEN-box and a D-box) that are recognized by the $APC^{CDH1}$ E3 ubiquitin ligase (78).

Despite the likely involvment of Cdc6p regulation in preventing the reformation of pre-RCs outside of G1, mutations that override this regulation are not sufficient to induce re-replication. For example, overexpression of either wild-type or mutant Cdc6p from human or *S. cerevisiae* cells is not sufficient to induce re-replication in their cognate cell types. In contrast, overexpression of SpCdc6p is known to result in re-replication (67). It is likely, however, that this re-replication is the result of CDK inhibition that resets the cell cycle clock to G1 (73). Although misregulation of Cdc6p is not sufficient to induce re-replication, recent studies strongly indicate that regulation of Cdc6p is one of several redundant mechanisms that prevent re-replication in eukaryotic cells (see below).

STRUCTURE OF A Cdc6p-RELATED PROTEIN The recent determination of the crystallographic structure of a Cdc6p/Orc1p-related protein from the archaebacteria *Pyrobaculum aerophilum* has provided the first look at the molecular details of these proteins (66). Three structural domains were identified and similar to other AAA+ ATPases domains I and II of the protein form the ATP binding/hydrolysis domain. More intriguing is the observation that domain III is structurally related to a winged-helix domain (81). This domain is responsible for DNA recognition in a number of transcription factors (e.g., E2F4/Dp2), which suggests the possibility that this domain of Cdc6p interacts with origin DNA. Mutational studies of *SpCDC6* provide support for the importance of this domain. Four different mutations in the winged-helix domain result in either cell cycle arrest with mostly replicated DNA or a complete loss of DNA replication (66). Despite the interesting possibility, there is no direct evidence for DNA binding by this motif of Cdc6p, although it remains possible that this interaction occurs only in the context of higher-order replication complexes. One previous report has detected DNA binding of a glutathione-S-transferase-ScCdc6p; however, this was mediated by the poorly conserved and nonessential N terminus of Cdc6p (82). An alternative possibility, raised by the recent determination of the structure of the *E. coli* $\gamma$ clamp loading complex, is that the C-terminal winged-helix domain of Cdc6p mediates protein-protein interactions with other AAA+ ATPase in the pre-RC (e.g., Orc1p, Orc4p) to form a "rotator cuff" involved in MCM loading (83, 84).

Domains I and II of Cdc6p form a cashew-shaped molecule that binds ATP in the cleft between the two domains. In addition to the Walker A- and B-motifs,

these domains contain the sensor I and II motifs, which interact with the phosphate groups of bound ATP. These sensor motifs are conserved in a variety of other ATP binding proteins and are thought to play a key role in distinguishing between the ATP- and ADP-bound states and to mediate subsequent conformational changes. Additional alleles of ScCDC6 support a function for the conserved Sensor I and Sensor II regions. The cdc6–3 allele that has mutations in two positions, one of which is in Sensor I, results in constitutive MCM association with chromatin and partial overreplication of genomic DNA (85). A more systematic analysis of mutations in other conserved domains of ScCDC6 suggests that mutations in the Sensor II domain destabilize the protein whereas mutations in several other conserved motifs, including Sensor I, produce inactive but stable proteins (86). Similarly, Sensor I and II mutations in the SpCdc6p result in nonfunctional proteins (66). Thus, although these domains are critical for Cdc6p function, their exact role remains unclear.

## The Cdt1 Protein

*SpCDT1* was originally identified as a gene in *S. pombe* that was regulated by the SpCdc10p transcription factor (87) and subsequently has been implicated as a key factor in pre-RC assembly. Mutations in *SpCDT1* resulted in a block to DNA replication and defects in the S-phase checkpoint (87). More recently, SpCdt1p has been shown to associate with the C terminus of SpCdc6p to cooperatively promote the association of MCM proteins with chromatin (88). Consistent with its regulation by SpCdc10, SpCdt1 protein levels peak at G1 and decay as cells pass through S phase. Like other members of the pre-RC complex, Cdt1p is conserved in other eukaryotes including *Xenopus* (XCDT1/XlCdt1p), humans (HsCdt1p), and *D. melanogaster* (Double parked/Dup/DmCdt1p) (89–91). Experiments in *Xenopus* egg extract reveal that XlORC, but not XlCdc6p, is required for chromatin association of XlCdt1p (89). As in *S. pombe*, XlCdc6p and XlCdt1p are found to be independently required for MCM chromatin association. Mutations in the *Drosophila* Cdt1 analog (Dup) show a failure to undergo S phase during the postblastoderm divisions and defects in DNA replication and amplification in the adult ovary. Consistent with a role in the pre-RC, DmCdt1p colocalizes with DmORC at sites of DNA amplification and requires DmORC for this localization (90). Although initially it was thought that there was no *S. cerevisiae* analog, recent studies have identified a functional homolog, ScCdt1p (originally identified as Tah11p) (S. Tanaka & J. Diffley, personal communication). As with Cdc6p, the function of Cdt1p in the recruitment of the MCMs to the chromatin remains unclear, although it is clear that Cdt1p is the target of the replication inhibitor geminin (see below).

## The Mcm2–7 Protein Complex

The genes that encode the Mcm2–7p complex were originally identified in genetic screens for proteins involved in plasmid maintenance, cell cycle progres-

sion, and chromosome missegregation and were originally grouped together based on their sequence similarity (reviewed in 93). Although they originally had a number of different names, standard names for Mcm2–7p were subsequently agreed upon (92). Interestingly, each MCM is highly related to all others, but unique sequences distinguishing each of the subunit types are conserved across eukaryotes. All eukaryotes appear to have exactly six MCM protein analogs that each fall into one of the existing classes (Mcm2–7p), which argues that each MCM protein has a unique and important function. This hypothesis is supported by the lethality that results from deleting any individual MCM gene in *S. cerevisiae* and *S. pombe* (93, 94). Subsequent biochemical studies identified a multiprotein complex consisting of XlMcm2–7p as a critical component of the replication licensing system in *Xenopus* egg extracts (95–97). Both biochemical and genetic studies strongly indicate that these proteins function together as a complex in the cell. The assembly of the MCM proteins onto chromatin requires the coordinated function of ORC, Cdc6p, and Cdt1p. Interestingly, once the MCM proteins have been loaded on chromatin, ORC and Cdc6p can be removed from the chromatin without preventing subsequent DNA replication, which suggests the primary role of the pre-RC is MCM loading (56, 98).

**IS THE Mcm2–7p COMPLEX THE DNA HELICASE AT THE REPLICATION FORK?** Unlike other members of the pre-RC, current data strongly support a role for the MCM proteins in both the initiation and elongation steps of DNA synthesis. ChIP assays demonstrated that several of the MCM proteins associated with the origin and origin-proximal DNA sequences in a temporal manner that was very similar to the pattern of association of the replicative DNA polymerase $\epsilon$ (46, 99, 100). Further in vivo support for a role for the MCM proteins at the replication fork was provided by the analysis of "degron" alleles of the ScMCM proteins that result in their complete degradation at the nonpermissive temperature (101). These mutant alleles resulted in no replication if inactivated prior to S phase and rapidly arrested the progression of the replication fork if inactivated during S phase. Interestingly, in contrast to biochemical studies of mammalian MCM proteins that suggest only a subset of the MCM proteins are required for DNA helicase activity (see below), inactivation of any of the six *S. cerevisiae* MCM proteins prevents further progression of the replication fork. Consistent with the requirement of ORC, Cdc6p, and Cdt1p function to assemble MCM proteins at the origin, reexpression of the degraded Mcm protein in S phase failed to rescue replication arrest (101). It is also possible that the failure to reestablish replication could be due to an inability of the newly expressed protein to be incorporated into a Mcm2–7p complex.

Mutation of the B2 element of ARS1, which has been implicated as the site of DNA unwinding at the origin, interferes with MCM association with the origin (102, 103). This could result from an inability to unwind origin DNA; however, there is no evidence that DNA unwinding occurs during MCM origin association. Alternatively, the B2 element could be a binding site for Mcm2–7p or for a

protein required for Mcm2–7p loading (e.g., ScCdt1p). Consistent with the latter possibility, recent studies have identified a requirement for specific DNA sequences related to the ACS at the *ARS1* B2 element (103a). Indeed, the ability of these sequences to facilitate Mcm2–7p loading is unrelated to their helical instabilty. These studies have not identified the protein that recognizes these sequences, although overexpression of ScCdc6p can partially rescue the mutant phenotype.

Although it is clear that the MCM proteins act at both the replication fork and origin, their biochemical function at these sites remains undetermined. Biochemical data support the hypothesis that MCM proteins act as a DNA helicase. A complex containing the mouse Mcm4, 6, 7 proteins has a weak, nonprocessive, intrinsic $3' \rightarrow 5'$ DNA helicase activity (104, 105). Similarly, studies of the *S. pombe* MCM proteins have identified a DNA helicase activity dependent on SpMcm4, 6, 7p (106). Conventional helicase substrates with only a single exposed end show low processivity with this enzyme (106). When a template that contained exposed $5'$ and $3'$ tails was used with excess SpMcm4/6/7p and *E. coli* ssDNA binding protein, more extensive regions of DNA were displaced ($>$500 bp) (107). This finding raises the possibility that the MCM complex requires a forklike structure for processive helicase activity. In neither the mammalian nor the *S. pombe* studies was helicase activity detected with hexameric complexes containing all six MCM subunits. Indeed, helicase activity of MCM4/6/7p was inhibited by the addition of Mcm2p or the Mcm3, 5p complex (105, 106, 108). Together these findings suggest a model in which the Mcm4, 6, 7 proteins act as the replicative DNA helicase and the Mcm2, 3, and 5 proteins act as inhibitors of this activity (but see below).

Further support for the hypothesis that MCM proteins act as a DNA helicase comes from the biochemical analysis of the single MCM analog isolated from the archaebacteria *Methanobacterium thermoautotrophicum*. Purified recombinant MtMCM protein was found to form a large multimeric complex (109–111). Both the native size of the purified protein and electron microscopic (EM) studies suggest that MtMCM proteins form a double hexamer, as is found for a number of other replicative DNA helicases. More importantly, this complex was found to act as a robust $3' \rightarrow 5'$ DNA helicase with a processivity of $>$500 bp. Although clearly related to eukaryotic MCM proteins, evidence that this protein acts during DNA replication in *M. thermoautotrophicum* is lacking.

**Mcm2–7p COMPLEX FORMATION** EM studies of the eukaryotic MCM complex support a hexameric structure for the MCM complex. EM of a purified *S. pombe* complex containing all six MCM proteins shows a doughnutlike structure with a central cavity (112). EM of the human Mcm4/6/7p complex shows a similar toroidal structure with six lobes surrounding a central cavity (108). Interestingly, addition of ssDNA results in the formation of a structure resembling the "beads on a string" seen for nucleosome-packaged dsDNA. Although not definitive, this structure is consistent with Mcm4/6/7p becoming associated with the ssDNA,

potentially by passing the ssDNA through the central cavity of the Mcm4/6/7p complex. If this is the case, it raises the question of how the MCM protein ring is loaded. In these experiments, a circular M13 ssDNA was used as substrate, which indicates that either the ssDNA or the MCM ring would have to be broken for MCM complex loading. This raises the interesting possibility that the helicase activity of the Mcm4/6/7p complex relative to the inactivity of the Mcm2–7p complex may be because the former complex can oscillate between an open and a closed ring to allow ssDNA loading.

Although typically found as a heterohexamer containing one of each of the six MCM proteins, the MCM proteins have also been purified in a variety of subassemblies that may shed light on the architecture of the MCM complex. In particular, complexes containing Mcm2/4/6/7p, Mcm4/6/7p, Mcm3/7p, and Mcm3/5p are commonly detected (106, 113–115). There is evidence that subcomplexes derived from *Xenopus* egg extracts can re-form into heterohexamers and function in DNA replication (114, 115). Interestingly, some of these subassemblies were able to associate with chromatin independently of the intact MCM complex. This association, however, is Cdt1-independent and cannot be complemented for replication by the addition of the remaining MCM subunits. Indeed, all six MCM proteins must be present simultaneously to associate with chromatin in a productive manner. Thus, although subcomplexes may be able to assemble in an ordered fashion (114, 115), this mechanism is nonproductive and unlikely to be relevant to the normal in vivo assembly of MCM proteins into the pre-RC.

NUCLEOTIDE REGULATION OF THE MCM COMPLEX Studies of the associated ATPase activity of the ScMCM complex have shed additional light on the coordinated function of the MCM subunits and the structure of the complex. Unlike the situation for the DNA helicase activity, the intact MCM hexamer has a robust associated ATPase activity (115a). Mutation of any one of the putative ATP binding sites conserved in all six MCM proteins reduces this activity dramatically, which indicates that ATP hydrolysis is a coordinated event involving all six subunits of the MCM complex. In contrast to the hypothesis that Mcm4/6/7p acts alone, only low levels of ATPase activity are detected for this complex. Although Mcm2/3/5p have no ATPase activity either individually or as a group, addition of these proteins to the weakly active ScMcm4/6/7p largely restores ATPase activity to that of the intact heterohexamer. Interestingly, despite the stimulatory nature of the Mcm2/3/5p subunits, mutation of multiple ATP binding motifs indicates that only the Mcm4/6/7p ATP binding motifs are required for ATPase activity, and the Mcm2/3/5p ATP binding motifs act only in a regulatory manner. Finally, studies of individual MCM subunits indicate that specific pairs of MCM proteins function together. For example, Mcm3p but not Mcm6p can activate ATP hydrolysis by Mcm7p. On the other hand, Mcm2p can activate Mcm6p activity. In addition to shedding light on the mechanisms of coordinated ATP hydrolysis, these studies suggest a specific architecture for the

MCM hexamer with Mcm3p adjacent to Mcm7p, Mcm2p adjacent to Mcm6p, and Mcm4p adjacent to Mcm5p. These findings also suggest that both members of the catalytic pairs contribute to achieving a conformation or structure that allows ATP binding and hydrolysis. This type of coordinated ATPase activity and the mixture of active and inactive subunits has precedence in studies of both the F1 ATPase and the T7 DNA helicase (116, 117).

CONTROL OF Mcm2–7p NUCLEAR LOCALIZATION The nuclear localization of MCM proteins is tightly regulated in budding yeast cells. The MCM proteins are present in the nucleus in G1 and S phase, but during G2 and M they are exported to the cytoplasm. Studies in both *S. pombe* and *S. cerevisiae* indicate that an intact six-subunit MCM complex is required for the complex to enter into the nucleus (118–120). This requirement is a result of the presence of consensus nuclear localization sequences (NLSs) on only a subset of the subunits (Mcm2p and Mcm3p). During S, G2, and M phases, MCM proteins that are not bound to chromatin are actively exported from the nucleus. In *S. pombe*, export of the MCM complex is mediated by the Crm1 nuclear export factor (120); however, it appears that the bulk of the SpMCM proteins are constitutively nuclear. In *S. cerevisiae*, nuclear export is promoted by cyclin-dependent kinase (CDK) activity, although it is unclear whether this is a result of direct modification of MCM proteins (118, 119). This process may initiate as early as late G1, as the accumulation of G1 cyclins is sufficient to promote the export of a subset of the MCMs (118). Export is likely to act only on a subset of free MCMs, however, as the export of chromatin-bound MCMs requires the completion of S phase. The protection of chromatin-bound MCMs from export may be due to a lack of accessibility to CDK or the export machinery or, alternatively, to the anchoring of the MCMs to the chromatin until the completion of replication. Unlike the situation in *S. cerevisiae*, metazoan MCM proteins are constitutively in the nucleus, although their association with chromatin is progressively weaker as cells proceed through S phase (121).

Although it remains most likely that the MCM proteins are the replicative helicase, several pieces of data do not fit easily into this hypothesis. The lack of helicase activity for the intact MCM hexamer is problematic. Moreover, MCM proteins do not strongly localize to the sites of DNA replication in *Xenopus* chromatin. Finally, the abundance of MCM proteins loaded onto chromatin is greatly in excess of the level expected if a single heterohexamer functions at each replication fork (122–126). If the MCM proteins do not act as the replicative helicase, what other function could they perform? One suggestion is that the MCM proteins are somehow involved in the regulation of chromatin structure in advance of replication fork passage (127). Biochemical studies have identified interactions between MCM proteins and histones, and modification of the local chromatin structure at origins interferes with MCM association (102, 113). Finally, it could also be that the MCM proteins are loaded on chromatin in excess to permit distributive helicase activity to replicate long stretches of DNA. If this

were the case, then each MCM hexamer would act for only a short distance before being replaced by an adjacent MCM complex. Certainly, such a complex molecular machine could also perform numerous other functions.

## ASSEMBLY AND FUNCTION OF THE PRE-REPLICATIVE COMPLEX

The individual factors described above work together to direct the formation of the pre-RC, a key intermediate in the replication initiation process (Figure 1). Association of ORC with the origin, which is required to recruit both Cdc6p and Cdt1p, initiates assembly of this structure. Cdc6p and Cdt1p associate with ORC-bound chromatin independently of one another (88, 89). ORC, Cdc6p, and Cdt1p together are required for the stable association of the MCM proteins with the origin during G1. Although studies in a number of different organisms support the basic outline of these events, the molecular details of the recruitment of these different factors to the origin remain unclear.

There is ample evidence to support an important role for nucleotide in controlling pre-RC assembly. Of the known members of the pre-RC, 10 of the 14 proteins have consensus motifs for nucleotide binding (Mcm2–7p, Orc1p, Orc4p, Orc5p, and Cdc6p). Mutations in eight of these ten nucleotide binding motifs result in nonfunctional proteins [only mutants in the Orc4p and Orc5p ATP binding sites are viable (27)]. Recent biochemical and genetic studies have begun to elucidate the role of nucleotide in pre-RC formation. Studies of an in vitro assay for pre-RC formation indicate at least two ATP-requiring steps in the formation of the pre-RC. The first step is the association of ORC with the origin, as previously demonstrated by DNA binding experiments. The recruitment of MCM and Cdc6p, however, requires a second ATP-dependent step (128). Further studies using this promising new assay will determine whether this step involves ATP binding to Cdc6p, the MCM proteins, or both. Mutations in the Orc1p site for nucleotide binding support a role for ORC nucleotide binding in the recruitment of Cdc6p (see above) (28). Similarly, there is evidence that ATP binding to Cdc6p is required for its interaction with ORC (33, 50). Taken together, these results make it clear that ATP-bound forms of ORC and Cdc6p are likely to be required for productive interactions and pre-RC formation. The role of ATP hydrolysis by these proteins remains unknown.

Despite the likely requirement of ATP binding and hydrolysis for pre-RC formation, the exact molecular changes brought about by nucleotide are unclear. It has been suggested that a combination of ORC subunits and Cdc6p may perform a function similar to clamp loaders that assemble sliding clamps [e.g., proliferating cell nuclear antigen (PCNA)] onto DNA (63). In support of this hypothesis, Cdc6p, Orc1p, Orc4p, and Orc5p are all related to one another as well as to subunits of the known clamp loaders (32, 43, 45, 63, 129), similar to the relatedness of the subunits of the replication factor C (RFC) and $\gamma$-complex

clamp loaders. If Cdc6p and ORC act as a clamp loader, then what is the ring that is loaded? A likely candidate would be the ring-shaped MCM heterohexamer that has been suggested to act as the replicative DNA helicase. If this is the case, then the MCMs are likely to be loaded around dsDNA rather than ssDNA or a primer-DNA junction because there is no evidence that the dsDNA is unwound during pre-RC formation (130).

Recent studies in *Xenopus* extracts indicate that once MCM proteins are loaded onto chromatin, XlORC and XlCdc6p can be removed without interfering with initiation of DNA replication (56, 98). This can be accomplished either by treatment with cyclin A/Cdk2, by addition of metaphase extract, or by high salt extraction of chromatin. Once treated, this chromatin can be added to interphase extracts that are depleted of either ORC or Cdc6p, and no defect in replication initiation is observed. Consistent with these findings, the association of XlCdc6p (but not XlORC) with chromatin appears to be transient in *Xenopus* extracts (98). Similarly, studies of Cdc6p association with origins in *S. cerevisiae* indicate that MCM proteins remain bound at late origins after Cdc6p is degraded (131). The requirement for ORC function during late-origin activation has not been assessed. These findings argue that the presence of ORC and Cdc6p is not required to recruit essential replication factors after the MCM proteins are assembled; however, they cannot eliminate the possibility that the artificial removal of ORC or Cdc6p mimics part of the normal initiation function of these proteins. This is a particularly important consideration for ORC, as ample evidence indicates that ORC remains bound to the chromatin/origins throughout S phase in many different organisms.

## REGULATION OF THE FORMATION OF PRE-REPLICATIVE COMPLEX

Pre-RC formation is regulated by *trans*-acting factors that signal the position in the cell cycle and by *cis*-acting chromosomal structural cues that assist in origin selection. The regulation by the cell cycle is critical to ensure that origins fire once and only once during each cell cycle. Regulation by chromatin is likely to play an important role in coordinating DNA replication with other DNA trans-actions (e.g., transcription).

### CDK-Mediated Prevention of Re-Replication

CDK activity has a bipartite role in the regulation of eukaryotic DNA replication. First, elevated CDK activity activates origins of replication as cells enter into S phase. Second, the same elevated levels of CDK activity are also required to prevent reinitiation from origins during S, G2, and M phases of the cell cycle. In addition, CDKs have been implicated in controlling the time of replication initiation at specific origins. Recent studies have suggested that CDKs prevent

re-replication by inhibiting pre-RC formation (discussed in this section). Discussions of how CDKs target the pre-RC, activate replication initiation, and influence the time of origin firing are presented in the section "Kinases controlling the transition to replication."

Eukaryotic cells have developed multiply redundant mechanisms to prevent reinitiation of replication from the same origin during a single cell cycle that relies on elevated CDK activity preventing new pre-RC formation. An important role for CDKs in preventing re-replication was originally suggested by genetic studies in *S. pombe* and *S. cerevisiae.* In each case, inactivation of CDK activity in G2/M cells resulted in full re-replication of the genome (71, 132, 133). That this effect is mediated by inhibition of pre-RC formation is supported by findings in several species indicating that expression of elevated levels of CDK activity in G1 prevents new pre-RC formation (71, 132, 134, 135). More recent studies provide clear evidence that at least three of the components of the pre-RC (ORC, Cdc6p, and the MCMs) are phosphorylated by CDKs to prevent re-replication and pre-RC assembly. We summarize what is known about the CDK-dependent phosphorylation of each of these components and the effect of this phosphorylation on their function below.

INHIBITION OF Cdc6p BY CDK The phosphorylation of Cdc6p by CDK activity is well documented both in vitro and in vivo (93). As discussed above, these modifications appear to control either the degradation or nuclear export of Cdc6p at the G1/S transition (Figure 1). Whether Cdc6p is exported (mammalian cells) or degraded (yeast), the primary phosphorylation sites that control these processes are located at the N terminus of the protein. In yeast Cdc6p, additional phosphorylation sites in the C terminus contribute to this regulation. Mutation of all consensus sites of CDK phosphorylation, however, does not promote re-replication (68, 70, 74, 79, 136, 137), although it results in either the stabilization (*S. cerevisiae*) or nuclear retention (mammals) of Cdc6p. In contrast, overexpression of *S. pombe* Cdc6p causes overreplication, and alteration of its putative sites of CDK phosphorylation exacerbates this ability (138, 139). This modification on its own, however, does not result in re-replication without overexpression, and it has been suggested that the overexpression is necessary to inhibit CDK activity in G2 (73; see above).

In the normal cell cycle, pre-RC destabilization and replication activation by CDKs are likely to be initiated by distinct G1- and S-phase cyclins. Although it is clear that B-type cyclin-dependent kinases (Clb/CDKs) can act to phosphorylate and destabilize Cdc6p, recent studies in *S. cerevisiae* suggest that G1 CDK complexes (Cln/CDK) are also competent to phosphorylate Cdc6p and target it for degradation (75). This result is consistent with biochemical data from *X. laevis* that suggest that the essential function of Cdc6p is to load MCM proteins and that after MCM loading has occurred, Cdc6p is dispensable for initiation (56, 98). The ability of G1 CDKs to activate Cdc6p degradation provides an appealing model to separate the prevention of new pre-RC formation from the activation of

existing pre-RCs (75). By degrading Cdc6p prior to activating the Clb/CDKs responsible for activating the initiation of DNA synthesis, even the earliest origins activated would lack the Cdc6p necessary to re-form a pre-RC in early S phase. The fact that Cdc6p is selectively removed during late G1 whereas ORC and MCM are retained throughout G1 is consistent with the latter two molecules having a direct role in replication initiation.

CDK MODIFICATION OF Mcm2–7p CDK modification also controls the function of the MCM proteins, although the mechanism of this control is unclear. In vitro, CDK can phosphorylate at least Mcm2p and Mcm4p, and the same subunits are phosphorylated in vivo (140–144). For the most part the exact sites of phosphorylation have not been mapped; however, the mutation of six consensus CDK sites on mouse Mcm4p prevented phosphorylation by cyclin A/Cdk2 or cyclin B/Cdk1, and at least one of these sites is phosphorylated in a cell-cycle-dependent manner in vivo (143). Although studies in *S. cerevisiae* suggest that CDK activity is involved in regulating MCM localization (see below), the subunits modified and the sites of modification have not been determined.

The functional consequences of CDK modification of the MCM complex are currently under scrutiny. One possible mechanism is the direct inhibition of MCM complex activity. The helicase activity of mouse Mcm4/6/7p is inhibited by CDK phosphorylation in vitro. In addition, Mcm4/6/7p purified from G1- and G1/S-phase cells, but not from M-phase cells, is active as a helicase (143). Another mode of regulation is likely to be through the regulation of the nuclear localization of the MCM proteins. In *S. cerevisiae*, MCM proteins are localized to the nucleus in G1 and S phases but are exported from the nucleus in G2 and M phases (118, 119, 145). Inactivation of CDK activity results in the stable maintenance of MCM proteins in the nucleus. A third proposed mechanism is that CDK modification induces release from chromatin and prevents MCM reassociation. In *X. laevis*, XlMcm4p has been shown to be a substrate for mitotic CDKs, and hyperphosphorylation of XlMcm4p is correlated with the release of MCMs from chromatin (140, 142, 144, 146). In addition, biochemical studies have shown that hyperphosphorylated MCM complexes cannot be assembled onto chromatin (144). These different mechanisms are not mutually exclusive, and future studies will be required to elucidate how CDK modification of the MCM complex alters its activity.

MODIFICATION OF ORC BY CDK Like Cdc6 and the MCMs, ORC is a substrate for CDK modification. In particular, ScORC2, ScOrc6p, and SpOrc2p are phosphorylated in a CDK-dependent manner in vivo (147, 148). In addition, recombinant SpOrc2p has also been shown to be a direct target of immunoprecipitated Cdc2 in vitro (148). The in vivo phosphorylation of each of these proteins can be reduced or eliminated by mutation of consensus sites for putative CDK targets. The same mutations promote re-replication of the genome under

specific conditions (see below), which suggests that phosphorylation of ORC by CDKs negatively regulates replication. Although these modifications are clearly regulated in the cell cycle with a peak in M phase, the molecular consequences of phosphorylation of yeast ORC remain unclear. Studies in *Xenopus* egg extracts suggest that addition of CDK activity results in the release of ORC and Cdc6 from the chromatin, although direct evidence for the phosphorylation of ORC by CDK activity in these reactions is lacking (56, 98). This finding is consistent with the observation that XlORC is released from the chromatin during M phase (11, 61) and that purified CDKs complexed with cyclin E, A, or B can phosphorylate recombinant XlOrc1p and XlOrc2p (140).

Unlike the situation for Cdc6p and the MCM proteins, the aspect of ORC function inhibited by CDK phosphorylation is unclear. Such a modification could prevent ORC from binding DNA or associating with one or more pre-RC components, or it could alter the stability of one or more ORC subunits. In *S. cerevisiae* it is unlikely that the mechanism works through the inhibition of DNA binding, as ORC is bound to the DNA both in vivo and in vitro in the presence of mitotic CDK activity (46, 48, 128). In mammalian cells, CDK-dependent degradation is an intriguing possibility given the reports that HsOrc1p is degraded in S, G2, and M phases (36).

## Regulation of Pre-Replicative Complex Formation by Geminin

The discovery of geminin in metazoans as an inhibitor of pre-RC formation indicated that not all mechanisms that prevent inappropriate pre-RC formation are mediated by elevated CDK activity. High levels of geminin appear as S phase proceeds and continue to accumulate until late M phase when geminin is degraded in an APC-dependent reaction (Figure 2). Geminin acts by preventing the loading of MCM proteins onto chromatin without interfering with ORC and Cdc6p association (149). Human geminin was found to be associated with HsCdt1p, and the inhibition of pre-RC assembly by geminin could be reversed by addition of excess HsCdt1p to *Xenopus* egg extracts (91). Purification of RLF-B, a component of the replication licensing system essential for replication in *Xenopus* egg extracts, revealed that RLF-B is XlCdt1p and that this activity is inhibited by *Xenopus* geminin (150). The appearance of geminin after cells enter S phase and its destruction in mitosis suggests that it contributes to the prevention of re-replication by inhibiting Cdt1p function after entry into S phase (91, 149, 150). Studies of DNA replication for Epstein-Barr virus (EBV) support a role for geminin in preventing pre-RC assembly in human cells. Replication from EBV OriP, which has been shown to require HsOrc2p (18), is inhibited by geminin. Moreover, excess HsCdt1p rescues this inhibition.

## Redundancy of Inhibition of Re-Replication

Clearly, inhibition of any component of the pre-RC during S, G2, and M phases would be sufficient to block re-replication; however, recent studies in *S. cerevisiae* indicate that at least three components of the pre-RC are targeted for inhibition (147). In these studies, mutations in the CDK target sites of ORC were combined with mutations in Cdc6p and the MCM proteins designed to eliminate the consequences of CDK phosphorylation; the result was induction of re-replication in G2/M cells. Importantly, this re-replication was not the consequence of unscheduled CDK inhibition, as the levels of CDK activity remained high before and during re-replication. These results strongly suggest that the CDKs use redundant mechanisms to inhibit re-replication, and the results are consistent with the finding that alteration of the regulation of individual pre-RC components failed to induce re-replication in mammalian cells (39, 77, 79, 136, 137, 147). Because the re-replication observed in these studies was incomplete, inhibition of ORC, Cdc6p, and MCM activity by CDKs may not be the only mechanisms to inhibit re-replication. Mechanisms such as inhibition of Cdt1 function by geminin (91, 150) have already been described in mammalian cell and *Xenopus* egg extracts, providing further protection against re-replication of the genome. Indeed, in *S. pombe* cells, cooverexpression of specific phosphorylation-deficient alleles of SpCdc6p and SpCdt1p results in re-replication (150a). Similarly, cooverexpression of wildtype SpCdc6p and SpCdt1p in a mutant background that limits CDK activation also results in re-replication (150b).

Why does the cell go to such apparent extremes to prevent re-replication? Certainly the consequences of re-replication are dire for the cell, leading to significant genomic instability. Perhaps the simplest reason for this redundancy is to eliminate the possibility that mutation of a single target could allow re-replication. The use of a geminin-dependent mechanism, which is apparently independent of CDK activity, provides an additional level of protection from re-replication under conditions in which CDKs are inhibited by checkpoint activation. In addition, because even a single origin firing out of turn could cause substantial re-replication (as much as 200 kb), the use of redundant mechanisms could be selected to minimize any chance of multiple rounds of initiation in a single cell cycle. Finally, although each individual mechanism may be sufficient to provide some level of re-replication protection, it is possible that they are not equally effective at all times of the cell cycle or at all origins.

## Regulation of Pre-Replicative Complex Formation by Chromatin

All the events of DNA replication must also contend with chromatin-assembled templates, yet the effects of chromatin on replication are relatively unclear at a molecular level. There is ample circumstantial evidence to suggest that the conformation of local chromatin domains has a profound effect on both the activity and timing of origins. First, late origins are associated with heterochro-

matic domains (151). Second, changes in origin usage in developing *Drosophila* and *Xenopus* embryos are correlated with changes in transcription profile and the associated changes in chromatin structure (152, 153). Finally, chromatin remodeling factors (e.g., CHRAC, FACT, and histone acetylases) are thought to influence DNA replication (51, 52, 154, 155). As the understanding of the events that control the initiation of replication has increased, it has become possible to more directly assess the influence of chromatin at a molecular level.

In addition to the interactions between DNA replication factors and the DNA, there is evidence of interactions between local nucleosomes and the replication machinery. Early studies of the chromatin structure of the *S. cerevisiae* TRP1-ARS1 plasmid indicated that the *ARS1* origin was located in a nucleosome-free region and the nucleosomes on either side were tightly positioned (156). Movement of a nucleosome over the A-element of *ARS1* by a strong nucleosome positioning sequence resulted in the inhibition of replication of the plasmid (157). Most likely this effect is due to the inhibition of ORC binding, as biochemical experiments indicate that ORC cannot bind to its binding site in the context of a nucleosome (J. R. Lipford & S. P. Bell, unpublished data). Similar inhibitory effects of chromatin have been documented for XlORC binding (158). In addition to these negative effects of chromatin, the positioned nucleosomes adjacent to *ARS1* play a positive role in pre-RC formation. Displacement of nucleosomes immediately adjacent to ORC at the *ARS1* origin of replication causes a significant defect in origin function by reducing pre-RC formation (as measured by MCM association with the origin) (102). This suggests the intriguing possibility that the MCM proteins interact directly with nucleosomes during pre-RC formation. Such a hypothesis is supported by the finding that MCM2 from mouse cells interacts with histone H3 in vitro (113); however, data supporting this interaction in vivo are lacking.

Studies of the replication of Chinese hamster ovary (CHO) cell nuclei in *Xenopus* egg extracts support a role of chromatin in origin selection. Metaphase CHO nuclei are dependent on *Xenopus* ORC for replication initiation. In contrast, G1 CHO nuclei appear to have already assembled hamster ORC on chromatin such that XlORC is dispensable. Despite this, the origin selection is random in nuclei derived from early G1 cells but becomes more defined as cells progress through a point in G1 referred to as the origin decision point (ODP) (159). Disruption of nuclear architecture by detergent or addition of naked DNA of the same sequence results in random initiation, which suggests a role for chromatin and/or nuclear structure in origin selection (160). Pre-RC formation appears to be necessary to progress through the ODP; however, there are mixed results as to whether pre-RC formation is sufficient for this event. In one study, the ODP appeared to correspond with the stable association of hamster Orc1, Orc2, and Mcm3 with chromatin (34). In contrast, other studies have demonstrated that ORC, Cdc6, and the MCM proteins associate with chromatin during telophase or very early G1 (38, 80), well before the ODP. The telophase nuclei can be added to geminin-supplemented *Xenopus* egg extracts and still replicate normally,

indicating that MCM association with origins (and thus pre-RC formation) had occurred in telophase before the ODP (38). If pre-RC formation is not sufficient for passage through the ODP, the event responsible for activating only a specific subset of pre-RCs formed in telophase remains unknown.

## THE TRANSITION TO REPLICATION

The formation of pre-RCs marks potential sites for the initiation of DNA replication; however, multiple other proteins or protein complexes must associate with the origin prior to successful initiation of DNA synthesis (Figure 2 and Figure 3). These proteins include regulatory factors as well as components of the DNA replication fork. In studies similar to those used to analyze the formation of the pre-RC, the order of association of these various factors with the origin is becoming clear. For the sake of clarity we have separated the events occurring during this transition into two stages: the first leading to unwound DNA at the origin (Figure 2) and the second leading to the loading of polymerases at the primer-template junctions (Figure 3). It is likely, however, that the two stages are not as cleanly separated in the cell. For example, some of the proteins described in Figure 3 [e.g., Dpb11 and Sld2 in the DNA polymerase (pol) $\epsilon$ complex] might have a role in the last stages of the events described in Figure 2. Below we describe the properties of these factors and the order in which they appear to assemble at the origin.

### The Mcm10/Dna43 Protein

The *DNA43/MCM10* gene was identified by two separate screens. *DNA43* was identified in a screen for temperature-sensitive mutant strains that had defects in mitotic DNA synthesis (161), whereas *MCM10* was identified in the same screen for strains with defects in plasmid stability that isolated *MCM2, MCM3,* and *MCM5* (162). Characterization of the *mcm10–1* allele indicates that this gene is required for efficient initiation of DNA replication (163). In addition, this allele results in pausing of replication forks at sites within or adjacent to origins of replication, which suggests a role in DNA elongation. ScMcm10p function is required for continued replication fork progression at times when intiation factors like ORC and Cdc6p are dispensable [(71, 165); O. Aparicio & S. P. Bell, unpublished data]. For example, ScMcm10p is required to complete DNA synthesis after cells are released from arrest by the DNA replication inhibitor hydroxyurea (HU) (165). Mutant alleles of *MCM10* exhibit genetic interactions with a number of genes encoding proteins involved in DNA elongation including subunits of DNA pol $\epsilon$ and $\delta$, *ScMCM7,* and *ScCDC45* (164, 165). Similarly, studies of Cdc23 (the *S. pombe* Mcm10p analog) show weak genetic interactions with Cdc24, a protein that is essential for S-phase completion and that interacts with PCNA and replication factor C (RFC) (two known components of the replication fork) (159).

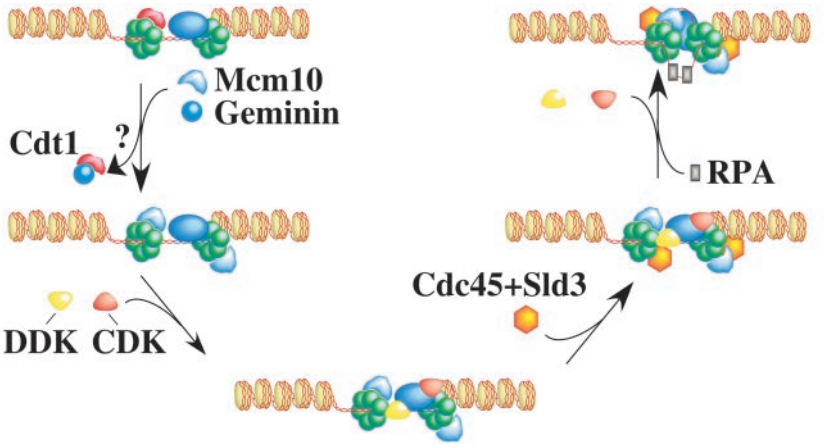


**Figure 2** The transition to replication: from the pre-RC to origin unwinding. Mcm10p, CDK, DDK, Cdc45, Sld3, and RPA are required for this transition. As with the model for pre-RC formation, the order rather than the relative stoichiometry is most clear from current data. We speculate that Mcm10p may displace Cdt1p from Mcm2–7p in the first step of this transition. It is also not clear whether, as indicated in the figure, CDK and DDK leave the initiation complex after they have accomplished the loading of Cdc45p. The additional Mcm2–7p illustrated in Figure 1 has been omitted for clarity. See text for further details.

Additional results suggest a role of Mcm10p in initiation of replication. ScMcm10p has been localized to origins using ChIP (164) and exhibits weak genetic and biochemical interactions with ORC (165). Despite this interaction, ORC and Mcm10p can associate with chromatin independently of one another (164). In addition, Mcm10p function is required for the continued association of Mcm2p with chromatin during G1. What remains unclear, however, is when in the cell cycle Mcm10p is associated with the origin and when it is associated with MCM proteins. Given the role of MCM protein in both initiation and elongation stages of DNA replication, it is possible that the genetic and physical interactions between Mcm10p and the Mcm2–7 proteins reflect a role of ScMcm10p in either or both of these stages. Further studies of the cell cycle regulation of ScMcm10p in its origin association and its interaction with Mcm2–7 proteins will be required to determine the function of this protein.

Studies of *Xenopus* Mcm10p function suggest a resolution to the multiple proposed functions of ScMcm10p. *Xenopus* Mcm10p chromatin loading requires chromatin-bound Mcm2–7p. Depletion of Mcm10p blocks the loading of Cdc45p and RPA and prevents the formation of topological intermediates that resemble an unwound origin (165a). This suggests that Mcm10p functions at the transition from the pre-RC to the elongating state. Although XlMcm10p is not required for DNA synthesis on M13 ssDNA, it remains an open question as to whether it

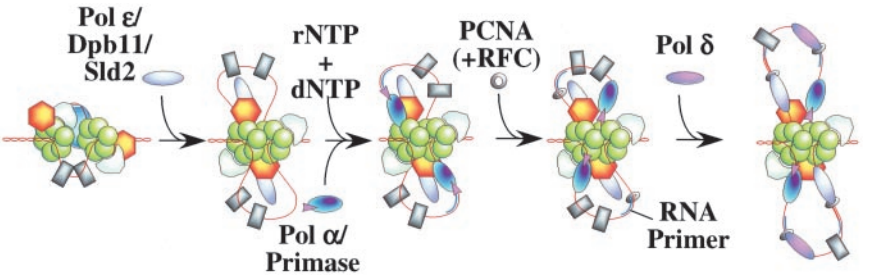


**Figure 3** The transition to replication: DNA polymerase loading. DNA polymerase (pol) $\epsilon$ and DNA pol $\alpha$/primase are loaded sequentially in a Cdc45p/Sld3p-dependent manner. The relative timing of DNA pol $\delta$ association with the origin is unknown. RNA/DNA primer synthesis allows the assembly of proliferating cell nuclear antigen (PCNA) using the replication factor C (RFC) clamp loader (not shown). Loading of PCNA allows pols $\epsilon$ and $\delta$ to bind the primer-template junction and continue DNA synthesis (polymerase switching). The juxtaposition of the Mcm2–7p helicase illustrated is based on the dodecameric nature of the archaeal MCM complex and studies of DnaB function at the *Escherichia coli* replication fork (228, 229). Similarly, we hypothesize that Cdc45p coordinates the action of leading and lagging strand polymerases on the same template strand at opposite replication forks rather than at opposite template strands at the same replication fork. The illustration has DNA pol $\epsilon$ involved in leading strand synthesis and DNA pol $\delta$ involved in lagging strand synthesis; however, these roles could be reversed or interchangeable (dNTP, deoxynucleoside triphosphate; rNTP, ribonucleoside triphosphate).



functions at the intact replication fork. The Mcm10p analog from human cells is found to associate with a nuclease-resistant structure in S- but not G2-phase cells (166). Consistent with HsMcm10p associating with human pre-RCs, coimmunoprecipitation and two-hybrid experiments suggest that HsMcm10p interacts with both HsOrc2p and two MCM subunits (166).

## Cdc45p (Sld4p)

Originally identified in a screen for cold-sensitive mutants in cell cycle progression (167), Cdc45p has emerged as a pivotal factor in the transition to replication. In vitro studies in *Xenopus* egg extracts indicate that Cdc45p interacts with DNA pol $\alpha$ and is required for the loading of this polymerase on chromatin (168). Similarly, studies in *S. cerevisiae* indicate that mutants in Cdc45p prevent the assembly of DNA polymerases $\alpha$ and $\epsilon$ at origins of replication (100, 169). In vitro studies using the human analogs of these proteins have detected similar interactions (170). Consistent with this hypothesis, studies of the early stages of replication initiation in *Xenopus* egg extracts show that Cdc45p loads prior to the

DNA unwinding step and the association of DNA polymerase (171, 172). In addition to the association with DNA polymerase $\alpha$, Cdc45p has been found to associate with ORC, DNA pol $\epsilon$, RPA, and the MCM proteins (100, 170, 173–175). The association of Cdc45p with both DNA polymerases and MCMs suggests that it may coordinate the function of these components of the replication fork and its association with RPA could help to tether it at the replication fork.

Once assembled at the origin, Cdc45p is incorporated into the replication fork. Studies of the association of Cdc45p with chromatin in yeast indicate that, like MCM proteins, Cdc45p colocalizes with polymerases at replication forks (100, 169). Genetic studies of a strong inactivating allele of *ScCDC45* support a role at the replication fork. This allele allows the rapid degradation of all ScCdc45p and prevents new initiation and continued progression of the replication fork at the nonpermissive temperature (176). Unlike the MCM proteins, reintroduction of ScCdc45p after its degradation during S phase allows replication to continue. This indicates that ScCdc45p can be reincorporated into the replication fork machinery in the absence of new pre-RC formation. Cdc45p is released from chromatin as cells progress through S phase in a manner similar to release of the MCM proteins, raising the possibility that release of Cdc45p could also prevent reinitiation after the completion of S phase (100, 103, 169, 174).

Loading of Cdc45p at the origin correlates with the time of initiation and requires CDK activity for tight binding. Studies of the association of Cdc45p with chromatin showed a strong dependence on CDK activity (103, 172). Interestingly, Cdc45p origin association correlates with the time of replication initiation (100, 169), which indicates that Cdc45p loading has requirements beyond simply CDK activation. Although there is controversy about the time of Cdc45p loading at early origins (100, 169, 173), it is clear that the level of association is increased substantially by elevated CDK activity. In *Xenopus* egg extracts, chromatin association of Cdc45p does not require RPA function (171, 172); however, studies of the *S. cerevisiae rfa2–2* allele suggest that Cdc45p origin association is RPA dependent (100). This distinction may be due to any number of differences including the different stringencies and specificities of the assays used to detect association with chromatin, residual RPA activity in RPA-depleted egg extracts, or dominant negative activity of the *rfa2–2* allele. Alternatively, the order of association may be different in *S. cerevisiae* and *X. laevis*.

## Dpb11 Protein, The Sld Proteins, and DNA Polymerases

The *ScDPB11* gene was identified as a multicopy suppressor of mutations in the catalytic and second largest subunits of DNA pol $\epsilon$ (177) and is related to the *S. pombe Cut5/Rad4* genes. The *dpb11–1* allele is synthetically lethal with mutations in other subunits of DNA pol $\epsilon$. Evidence in support of a physical interaction between Dpb11p and DNA pol $\epsilon$ was provided by coimmunoprecipitation assays (178). Detection of this association requires prior crosslinking,

which suggests that the interaction is relatively labile. Dpb11p associates with origins of replication and this association is codependent with DNA pol $\epsilon$ (Figure 3). Together, these findings suggest that Dpb11p is part of a DNA pol $\epsilon$ complex involved in origin recruitment. Consistent with a role in DNA pol $\epsilon$ recruitment, ScDpb11p and SpDpb11p are required for S phase (177, 179). *DPB11* mutants also show checkpoint defects. *S. cerevisiae* cells with defects in the *DPB11* gene are hypersensitive to treatment with DNA-damaging agents and are disrupted in the Rad53p-dependent checkpoint pathway (177, 180). In the absence of SpDpb11p, *S. pombe* cells enter mitosis without DNA replication (179), a phenotype similar to that seen with the absence of other proteins involved in replication initiation (e.g., SpCdc6p and SpOrc1p) (181, 182).

The *ScSLD* genes were identified in a screen for genes that were synthetically lethal with *dpb11–1* (183). *SLD4* was found to encode Cdc45p and *SLD1* encodes Dpb3p, the third largest subunit of DNA pol $\epsilon$. Two other *SLD* genes encode participants in the S-phase checkpoint: *SLD6* is the same as *RAD53* and *SLD2* is the same as *DRC1* and interacts with ScDpb11p (173, 180). Like a number of other proteins involved in the S-phase checkpoint, Sld2p is required for the initiation of DNA replication. Sld3p is also required for the initiation of replication and forms a complex with Cdc45p and associates with origins of replication in vivo (173). Like Cdc45p, Sld3p association with origins correlates with the time of replication initiation, consistent with a role in DNA polymerase recruitment. Consistent with the physical interaction between Sld3p and Cdc45p, the assembly of these two proteins at origins is mutually dependent. Moreover, a mutation in *SLD3* that reduces the Sld3p-Cdc45p interaction also results in a reduced interaction between Cdc45p and Mcm2p.

ASSEMBLY OF THE DNA POLYMERASES The ultimate goal of the DNA replication initiation machinery is the assembly of the DNA polymerases at the origins. The order of assembly of many of the components of the DNA polymerases and their accessory proteins has been investigated. Cdc45p is required for the assembly of many components of the DNA synthetic machinery at the replication fork, including RPA, PCNA, and DNA polymerases $\alpha$ and $\epsilon$ (100, 169, 171, 172). Although the exact order of assembly of these remaining components is unclear, studies in *S. cerevisiae* and *Xenopus* egg extracts agree that RPA is required for the loading of DNA pol $\alpha$ (100, 131, 171, 172). Intriguingly, neither RPA nor DNA pol $\alpha$ was required for the association of DNA pol $\epsilon$ with chromatin in *Xenopus* egg extracts (Figure 3) (166). Similarly, in *S. cerevisiae* ChIP studies indicate that DNA pol $\epsilon$ association with origin DNA is independent of DNA pol $\alpha$ (178). In contrast, the loading of DNA pol $\alpha$ requires DNA pol $\epsilon$ and Dpb11p. This order of polymerase loading may be to ensure that all of the necessary DNA polymerases are present at the origin prior to the synthesis of the first RNA primer by DNA pol $\alpha$ primase. Related experiments suggest that PCNA loading requires DNA pol $\alpha$, consistent with the requirement of a primer template junction for PCNA loading (171). Therefore, although DNA pol $\epsilon$ is recruited to

the origin prior to DNA pol $\alpha$, the processivity factor for DNA pol $\epsilon$, PCNA, is not present until after DNA pol $\alpha$ synthesizes bona fide replication primers.

## KINASES CONTROLLING THE TRANSITION TO REPLICATION

The assembly and activity of the DNA synthetic machinery at the origin is carefully controlled by the action of at least two kinases: Cdc7p/Dbf4p and CDKs. These kinases act during S phase to trigger specific steps in the transition to replication, particularly those leading to the association of Cdc45p with chromatin and the unwinding of the origin of DNA replication (Figure 2). Although the requirement for these kinases is clear, whether they simply regulate the assembly of proteins at the origin or also alter the enzymatic activity of these proteins remains to be determined. Here, we describe what we know about the role of these kinases in controlling the transition to replication.

### Cdc7p/Dbf4p (DDK)

The genes encoding Cdc7p and Dbf4p were originally identified in screens for genes that when mutated have characteristic phenotypes for cell-cycle arrest (reviewed in 179). The activity of Cdc7p/Dbf4p is dependent on the cyclical appearance of the regulatory Dbf4p subunit in a manner similar to the dependence of CDKs on the appearance of cyclins. The levels of Dbf4p peak in S phase and are controlled at the level of gene expression and regulated proteolysis. The similarity in the mechanisms of regulation between CDK and Cdc7p/Dbf4p has led to the latter being referred to as DDK (Dbf4p-dependent kinase).

DDK function requires localization to chromatin, but the mechanism of recruitment appears to vary between yeast and metazoans. In *S. cerevisiae*, Dbf4p chromatin association requires ORC, but is independent of either Cdc6p or MCM assembly (185). In contrast, studies of *Xenopus* egg extract reveal a strong requirement for MCM but not ORC or Cdc6 association with chromatin prior to Cdc7p recruitment (186, 187). It is possible that this difference stems from monitoring different partners in the DDK complex. Despite the different protein requirements for DDK recruitment, the kinetics of DDK chromatin association are consistent with loading occurring upon entry into S phase in both organisms (186–188). In addition, Cdc7p has been found to be associated with MCM proteins in both *S. cerevisiae* and *X. laevis*, although it is possible that this could represent tight substrate binding by DDK (see below and 185, 188a, 190).

The MCM proteins have emerged as prime candidates for the target of the DDK that triggers initiation of DNA replication. Genetic studies in *S. cerevisiae* identified an allele of the *MCM5* gene that bypasses the requirement for *CDC7* and *DBF4* for passage through S phase (188b). In addition, there is genetic suppression between alleles of *MCM2* and *DBF4* (189). In vivo phosphorylation

studies in *S. cerevisiae* indicate that five of the six MCM proteins are phosphorylated, but the cell cycle regulation of this phosphorylation is not well characterized (188, 189). Interestingly, despite the intriguing genetic suppression of *CDC7* and *DBF4* deletions by a mutation in *MCM5*, this is the only MCM subunit that has not been observed to be a phosphoprotein. In vivo evidence that MCM phosphorylation is due to DDK is limited, although one study has shown that phosphorylation of Mcm2p is reduced in a *dbf4–1* mutant strain at the nonpermissive temperature (188, 189).

Biochemical evidence supports the hypothesis that the MCM proteins are phosphorylated by DDK. Several different studies have demonstrated that purified DDK phosphorylates multiple MCM subunits when they are used individually as substrates (188, 189, 192, 193). There is substantial variability in the relative phosphorylation of the different subunits, although all groups agree that Mcm5p is not phosphorylated. The phosphorylation observed with individual subunits may be misleading. Studies of the *S. pombe* DDK analog (Hsk1p/Dfp1p) using the intact MCM complex as the substrate demonstrated that SpMcm2p was the preferred substrate (194). A similar preference for Mcm2p was observed when mammalian DDK complexes were used to phosphorylate intact MCM complexes (195, 196). Phosphopeptide mapping reveals that a subset of the Mcm2p peptides that are phosphorylated in vivo are modified by DDK in vitro (195). In addition to the MCM complex, in vitro studies have also demonstrated that the catalytic subunit of DNA pol $\alpha$/primase and ScCdc45p are phosphorylated by DDK (188, 197). Interestingly, mutations in *CDC7* and *DBF4* inhibit the association of Cdc45p with Mcm2–7p in vivo. Possible mechanisms could include reduced phsophorylation of Cdc45p or Mcm2p or an indirect effect of DDK loss.

Studies of the temporal requirement of DDK suggest that it acts individually at origins of replication to recruit Cdc45p at the time of initiation. Cell cycle studies in *S. cerevisiae* demonstrate that the inactivation of Cdc7 during S phase prevents firing of late origins but does not inhibit progression of the replication fork (198, 199). Studies of the assembly of replication factors on chromatin indicate that the *Xenopus* DDK has no effect on pre-RC formation but is required for assembly of XlCdc45p on the chromatin (172, 186). Although initial studies in *S. cerevisiae* indicated that ScCdc45p associates with chromatin in the absence of ScCdc7 function (103), further studies of the association of Cdc45p with origins (by ChIP) showed a strong dependence on ScCdc7p-ScDbf4p (100). A DDK requirement for Cdc45p association with origins is also consistent with genetic studies that indicate that these genes exert their functions in a codependent manner (200).

## Cyclin-Dependent Kinases (CDKs)

TARGETING CDKS TO ORIGINS Consistent with the role of CDKs in origin activation, CDKs have been reported to physically associate with multiple

components of the pre-RC. Interactions between ORC, Cdc6p, and CDKs are likely to be involved in recruiting CDKs to act at origins. In *X. laevis*, it is clear that CDKs are associated with chromatin in a reaction that depends on ORC and Cdc6p, making it likely that CDKs are present at origins (201). Because this association is not dependent on MCMs, it is unlikely to be a consequence of the formation of a replication fork or some other non-origin-associated protein complex. Corresponding data for the localization of yeast CDKs at the known origins of replication is lacking. Given that substrates of CDKs appear to have lasting physical interaction with the enzyme, it is also possible that this association is a consequence of CDK targeting the members of the pre-RC (see below).

There is substantial evidence for a direct interaction between CDKs and Cdc6p. Human Cdc6p has been coimmunoprecipitated with CDKs and the interaction has been recapitulated with purified recombinant proteins in vitro (39, 77, 136, 137). This interaction is dependent on a cyclin-binding motif (Cy-motif) originally identified on CDK inhibitors (202, 203). Mutation of this motif eliminates CDK-dependent phosphorylation of HsCdc6p and results in the constitutive nuclear localization of the latter protein throughout the cell cycle. The Cy-motif is a degenerate sequence of hydrophobic amino acids that often, but not always, contains RXL (204). It is present on many substrates, interacts with the hydrophobic patch on the cyclin, and accounts for a 100-fold increase in affinity of a substrate for CDK (205, 206). Indeed, it appears that together with the phospho-acceptor (S/TPXK/R) sequence, the Cy-motif forms a bipartite substrate binding site for CDKs. In *Xenopus* egg extracts, the XlCdc6p Cy-motif is essential for the interaction with and recruitment of cyclin E-Cdk2 to the chromatin (201). Mutation of the XlCdc6p Cy-motif renders the protein partially defective for the initiation of replication. The ability of this mutant Cdc6p to recruit MCMs and cyclin E-Cdk2 to chromatin in these assays was not tested, however, leaving open the possibility that the defect induced by this mutation is involved in other aspects of Cdc6p function besides cyclin E-Cdk2 recruitment (e.g., XlMCM loading). In this regard, it is noteworthy that a Cy-motif mutation in HsCdc6p reduces its in vitro ATPase activity, presumably because of conformational changes induced by the mutation (67). ScCdc6p also interacts with CDKs both in vitro and in vivo (62, 74); however, as described in more detail above, the domain responsible for this interaction [the N-terminal 47 amino acids (69)] is not essential for Cdc6 function. Instead, recent studies suggest that this domain, in addition to its role in regulating the stability of ScCdc6p, may also act as a CDK inhibitor in late mitosis (207, 208).

CDKs have also been reported to physically associate with subunits of ORC. SpOrc2p interacts with *S. pombe* Cdc2 in two hybrid and coimmunoprecipitation assays (148). Similarly, in *Xenopus* egg extracts, ORC copurifies and coimmunoprecipitates with cyclin A-Cdk1 (209). Human Orc1p physically associates with cyclin E-Cdk2 and cyclin A-Cdk1 in vitro (A. Dutta, unpublished data). The region of ORC required for these interactions in unknown; however, there is a Cy-motif in the human and *Xenopus* Orc1 proteins that could be responsible for

the interaction. Although intriguing, these interactions have yet to be tested for relevance in any organism.

CDK ACTIVATION OF REPLICATION There is an absolute requirement for CDK activity for DNA replication, yet the proteins that must be phosphorylated to promote replication remain unknown. The step at which CDKs function has been narrowed to a time after MCM loading but before Cdc45p association with chromatin/origins (172, 210, and see below). Although ORC, Cdc6, and MCM proteins are all targeted by CDKs, mutant forms of ScORC, SpOrc2p, and ScCdc6p that cannot be modified by CDKs continue to function during the initiation process (69, 75, 138, 139, 147, 148, 211). Thus, if these proteins are activated to initiate replication by phosphorylation, it is unlikely that this is a CDK-dependent event. Indeed, studies of SpOrc2p and XlMcm4p phosphorylation both suggest the existence of CDK-independent phosphorylation as cells enter S phase (146, 148). The Cdc7p kinase also is a substrate for CDK phosphorylation, and mutation of one of the sites modified by CDKs results in a substantial reduction in kinase activity; however, it has not been demonstrated that phosphorylation by CDK activates the kinase (196).

There is contradictory evidence regarding the role of phosphorylation of human Cdc6p in the promotion of replication initiation. Overexpression of human Cdc6p lacking phosphorylation sites has been reported to cause a block prior to replication initiation when released from cell cycle arrest (77, 136). This finding suggests that the nonphosphorylated form of HsCdc6p interacts with and titrates away an essential replication factor from the origin. In contrast, when asynchronous cultures are transfected with plasmids overexpressing the mutant HsCdc6p, there is no block to the cell cycle (79, 137). The reason for this discrepancy could be that the factor titrated by the mutant Cdc6p is limiting when cells are arrested in mitosis or G0 but is sufficiently abundant in cycling cells.

## Which Acts First, DDK or CDK?

Although it is clear that both CDKs and DDKs are absolutely required for the initiation of replication and more specifically for the tight association of Cdc45p with chromatin/origins, the order of function for these kinases is controversial. Reciprocal shift studies in *S. cerevisiae* cells support a model in which CDK function is required for subsequent DDK function (197). In contrast, by either depleting DDK activity or inhibiting CDK activity in *Xenopus* egg extracts, the opposite conclusion was reached: In this system DDK function is required for CDK function (186, 187). Although these different results could be due to fundamental differences in the action of these kinases in the different species, it is noteworthy that the activity of the kinases was not assessed in any of the above experiments, leaving open the possibility that residual activity of either kinase is providing misleading data. Additional methods of kinase inhibition may help to resolve these differences. It is also possible that these findings could reflect

differences between the time of origin association of CDKs and DDKs and their time of action. For example, these findings could be explained if DDK origin association is required for the execution of the CDK function, whereas DDK activity is required only after CDK activity. In this scenario, DDK activity would not be required for CDK function but DDK protein would be. Finally, these findings could also reflect different stability of the interaction of DDKs and CDKs with the origin. In *S. cerevisiae*, there is evidence that DDKs are associated with the origin but little or no evidence that CDKs are (185, 212). In contrast, both kinases are associated with chromatin in an ORC-dependent manner in *X. laevis*, which suggests that both are associated with origins.



## CONTROL OF REPLICATION TIMING

The time at which a eukaryotic origin initiates replication within S phase is a characteristic of each origin. Although the mechanisms that control this regulation are largely a mystery, there are a number of interesting clues. First, there is a strong correlation between the level of transcription of the surrounding genes and timing of replication. Early replicating origins tend to be associated with transcriptionally active regions, and late replicating origins are generally associated with transcriptionally repressed genes (151). Further evidence for the impact of surrounding chromatin on replication timing is provided by experiments in *S. cerevisiae* in which movement of origins from early to late or late to early loci indicates that the surrounding chromatin and not the origin itself is the major determinant of replication timing (213). Interestingly, the mechanisms that control the timing of replication appear to be established in late mitosis or G1 (214, 215).

Although how chromatin structure and local transcription affect the assembly of proteins at the origin is unclear, the differences between replication factor assembly at these sites are well documented. ORC, Cdc6, and MCM proteins associate with early and late origins equally during G1 (100, 131, 169, 216). In contrast, Cdc45p, RPA, and DNA pols $\alpha$ and $\epsilon$ all associate with origins in a manner correlated with their time of initiation (100, 131, 169). Thus, although the mechanisms that control replication timing do not affect pre-RC formation, they strongly regulate the association of the factors involved in the transition to replication. The robust formation of the pre-RC in late G1 at both early and late origins rules out the simple hypothesis that chromatin accessibility restricts assembly of proteins at late replication origins. Instead, more specific regulatory mechanisms must control the time of origin firing. Mutations in the Rad53p and Mec1p protein kinases allow late replicating origins to replicate prematurely (216, 217), which suggests that the intra-S-phase checkpoint is involved in controlling the time of origin firing. CDK function may also play an important role in the control of replication timing. Studies in *S. cerevisiae* indicate that Clb5p-Cdk1p can activate both early and late origins, whereas Clb6p-Cdk1p is

only capable of activating early origins of replication (218). It is possible that the differential effect of Clb5 and Clb6 is due to different overall levels of CDK activity; however, Clb2p cannot substitute for Clb5 to stimulate late origin activity, which suggests a qualitative difference in the ability of the two kinases to activate replication origins (219).

## DDK FUNCTION AND THE INTRA-S-PHASE CHECKPOINT

Studies of replication timing control in *S. cerevisiae* described above linked the Rad53/Cds1/Chk2 protein kinase to regulation of time of origin firing, particularly in the presence of DNA-damaging agents. Consistent with this, evidence has appeared suggesting that Rad53p regulates the activity of DDK, which is required to act at the time of initiation of each origin. Rad53p physically associates with Dbf4p, and the hydroxyurea-induced hyperphosphorylation of Dbf4p is dependent on Rad53p and Mec1p in *S. cerevisiae* and SpRad53p (Cds1p) in *S. pombe* (188, 220–222). This hyperphosphorylation partially inhibits the protein kinase activity of DDK, which raises the possibility that this is one mechanism by which Rad53 might inhibit late origin firing (188). In *S. pombe*, mutation of SpDbf4p (Dfp1 or Him1) results in growth suppression of cells in the presence of DNA-damaging agents and loss of the checkpoint pathway that prevents mitosis in the presence of incomplete DNA replication (222). In *S. cerevisiae*, the absence of Cdc7p (in the *bob1* strain) also suppresses growth in the presence of hydroxyurea, although in this case the pathway for checkpoint control in the S-M phases is not disrupted (188). Curiously, Rad53p appears to promote the expression of Dbf4p mRNA and protein through unknown mechanisms (223), although the significance of this regulation for checkpoint control is unclear. Taken together, the results suggest that the Rad53p-dependent intra-S-phase checkpoint pathway inhibits DDK activation at late origins. (For a more detailed discussion of the role of DDK in the intra-S-phase checkpoint, see the review in Reference 224).

## CONCLUDING REMARKS

Although progress in understanding the mechanism and control of DNA replication initiation has been rapid in the past several years, numerous critical questions remain. For example, how origins are selected in species other than *S. cerevisiae* remains largely undetermined. Although some degree of specificity is likely to be contributed by the DNA binding specificity of ORC, it is clear in all species (including *S. cerevisiae*) that other factors must also contribute to the origin selection process. Such factors are likely to include association with other pre-RC components, stimulation by transcription factors and local chromatin

structure, and potentially other components of chromosomal architecture. Although currently unknown, the determinants of the origin decision point (ODP) are likely to be key contributors to the origin selection process. It is intriguing that the Mcm2–7p complex is loaded before the ODP has occurred and that the loaded MCM proteins are in vast excess relative to the amounts of other replication fork proteins. It is possible that the ODP is related to loading of subsequent components of the initiation/elongation machinery (e.g., Mcm10p or Cdc45p) at a subset of the Mcm2–7p complexes.

Major progress has been made recently in identifying the proteins that are recruited to the origin and the order in which they are recruited; however, our knowledge of their biochemical activities once they arrive at the origin is still very limited. Given that 10 of the 14 proteins of the pre-RC are or have the potential to bind and hydrolyze ATP, it is likely that nucleotide control has a key role in these events. On the basis of studies of other nucleotide binding proteins, it is most likely that ATP binding is coupled to the formation of protein assemblies and that ATP hydrolysis is coupled to the disassembly of the same complex. The recent reconstitution of pre-RC assembly in vitro at *S. cerevisiae* origins (128) provides powerful tools to investigate the role of nucleotide and the mechanism of replication factor function within the pre-RC. Such studies also have the potential to identify additional components of the pre-RC. It is intriguing to note the similarity of these events to those that occur at the *E. coli* origin of replication. There are clear functional similarities between DnaA and ORC, DnaC and Cdc6p/Cdt1p, and DnaB and Mcm2–7p (225). In contrast, the transition to replication in eukaryotes appears more complicated than in prokaryotes both in terms of the number of proteins involved and their regulation by the cell cycle apparatus. It is worth noting that studies in a number of different eukaryotic species have provided key connections between different complexes that would have been missed had the studies been restricted to only one species. For example, the identification of Cdt1p in *S. pombe* and *X. laevis* as an accessory factor required for Mcm2–7p loading provided a strong impetus to identify a similar factor in *S. cerevisiae.*

In contrast to pre-RC formation, the transition to replication is only now being defined at the level of factor recruitment. Clearly, the biochemical changes induced by these factors will be critical to understand the earliest steps of the replication initiation process. For example, although Cdc45p and/or Mcm10p are required for the opening of the DNA duplex and recruitment of the polymerases, it remains unclear what the biochemical roles of these factors are in these events. Do they stimulate other proteins to unwind the DNA or do they directly participate in this event? Are they in direct contact with all of the polymerases and/or the ssDNA? Equally unclear is the mechanism by which these events are activated by the CDKs and DDKs. Again, the requirement of these kinases to trigger the key events of the transition to replication is undisputed, yet the substrates that are modified and the consequences of these modifications remain largely unknown.

New vistas for the study of eukaryotic DNA replication are also emerging. The sequencing of whole eukaryotic genomes will provide powerful new tools to investigate the replication process at a genome-wide level. Such studies will provide important new avenues to identify origins of replication and the connections between DNA replication and chromosome ultrastructure. Novel techniques, such as DNA combing (226), combine the advantages of observing replication forks on single molecules [as was done years ago by EM-autoradiography (227)] with the ability to investigate specific sequences in the genome. The discovery that viruses such as Epstein-Barr utilize the host replication apparatus will provide new models for experimental study of mammalian DNA replication. In addition, such findings suggest that eukaryotic DNA replication factors or the viral factors that recruit them could be targets of antiviral therapies.



## ACKNOWLEDGMENT

The authors would like to thank A. Schwacha for critical comments; B. Stillman, M. Botchan, S. Tanaka, J. Diffley, and T. Kelly for communicating results prior to publication; and M. DeVries for assistance with figures. Work in the laboratory of SPB is supported by grants from the National Institutes of General Medical Sciences (GM52339 and GM58701) and the Howard Hughes Medical Institute. Work in the laboratory of AD is supported by grants CA60499 and CA89406 from the National Cancer Institute and by RSG01–055-01-CCG from the American Cancer Society.

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

## LITERATURE CITED

1. Bell SP. 1995. *Curr. Opin. Genet. Dev.* 5:162–67
2. Clyne RK, Kelly TJ. 1995. *EMBO J.* 14:6348–57
3. Dubey DD, Zhu J, Carlson DL, Sharma K, Huberman JA. 1994. *EMBO J.* 13:3638–47
4. Dubey DD, Kim SM, Todorov IT, Huberman JA. 1996. *Curr. Biol.* 6:467–73
5. Bielinsky AK, Gerbi SA. 2001. *J. Cell Sci.* 114:643–51
6. Blow JJ. 2001. *EMBO J.* 20:3293–97
7. Waga S, Stillman B. 1998. *Annu. Rev. Biochem.* 67:721–51
8. Weinert T. 1998. *Cell* 94:555–58
9. Zhou BB, Elledge SJ. 2000. *Nature* 408:433–39
10. Bell SP, Stillman B. 1992. *Nature* 357: 128–34
11. Romanowski P, Madine MA, Rowles A, Blow JJ, Laskey RA. 1996. *Curr. Biol.* 6:1416–25
12. Rowles A, Chong JPJ, Brown L, Howell M, Evan GI, Blow JJ. 1996. *Cell* 87:287–96
13. Carpenter PB, Mueller PR, Dunphy WG. 1996. *Nature* 379:357–60
14. Pflumm MF, Botchan MR. 2001. *Development* 128:1697–707
15. Pinto S, Quintana DG, Smith P, Mihalek RM, Hou ZH, et al. 1999. *Neuron* 23:45–54
16. Landis G, Kelley R, Spradling AC, Tower J. 1997. *Proc. Natl. Acad. Sci. USA* 94:3888–92

17. Chesnokov I, Gossen M, Remus D, Botchan M. 1999. *Genes Dev.* 13:1289–96
18. Dhar SK, Yoshida K, Machida Y, Khaira P, Chaudhuri B, et al. 2001. *Cell* 106:287–96
19. Rao H, Stillman B. 1995. *Proc. Natl. Acad. Sci. USA* 92:2224–28
20. Rowley A, Cocker JH, Harwood J, Diffley JF. 1995. *EMBO J.* 14:2631–41
21. Lee D, Bell SP. 1997. *Mol. Cell. Biol.* 17:7159–68
22. Austin RJ, Orr-Weaver TL, Bell SP. 1999. *Genes Dev.* 13:2639–49
23. Chesnokov I, Remus D, Botchan M. 2001. *Proc. Natl. Acad. Sci. USA* 98:11997–2002
24. Bielinsky A, Blitzblau H, Beall EL, Ezrokhi M, Smith HS, et al. 2001. *Curr. Biol.* 11:1427–31
25. Ogawa Y, Takahashi T, Masukata H. 1999. *Mol. Cell. Biol.* 19:7228–36
26. Chaudhuri B, Xu HZ, Todorov I, Dutta A, Yates JL. 2001. *Proc. Natl. Acad. Sci. USA* 98:10085–89
27. Schepers A, Ritzi M, Bousset K, Kremmer E, Yates JL, et al. 2001. *EMBO J.* 20:4588–602
28. Li JJ, Herskowitz I. 1993. *Science* 262: 1870–74
29. Chuang RY, Kelly TJ. 1999. *Proc. Natl. Acad. Sci. USA* 96:2656–61
29a. Kong D, DePamphilis ML. 2001. *Mol. Cell. Biol.* 21:8095–103
29b. Lee JK, Moon KY, Jiang Y, Hurwitz J. 2001. *Proc. Natl. Acad. Sci. USA* 98:13589–94
30. Moon KY, Kong D, Lee JK, Raychaudhuri S, Hurwitz J. 1999. *Proc. Natl. Acad. Sci. USA* 96:12367–72
31. Lee DG, Makhov AM, Klemm RD, Griffith JD, Bell SP. 2000. *EMBO J.* 19:4774–82
32. Klemm RD, Austin RJ, Bell SP. 1997. *Cell* 88:493–502
33. Klemm RD, Bell SP. 2001. *Proc. Natl. Acad. Sci. USA* 98:8361–67
34. Natale DA, Li CJ, Sun WH, DePamphilis ML. 2000. *EMBO J.* 19:2728–38
35. Thome KC, Dhar SK, Quintana DG, Delmolino L, Shahsafaei A, Dutta A. 2000. *J. Biol. Chem.* 275:35233–41
36. Kreitz S, Ritzi M, Baack M, Knippers R. 2001. *J. Biol. Chem.* 276:6337–42
37. Tatsumi Y, Tsurimoto T, Shirahige K, Yoshikawa H, Obuse C. 2000. *J. Biol. Chem.* 275:5904–10
38. Okuno Y, McNairn AJ, den Elzen N, Pines J, Gilbert DM. 2001. *EMBO J.* 20:4263–77
39. Saha P, Chen J, Thome KC, Lawlis SJ, Hou ZH, et al. 1998. *Mol. Cell. Biol.* 18:2758–67
39a. Li CJ, DePamphilis ML. 2002. *Mol. Cell. Biol.* 22:105–16
40. Dhar SK, Delmolino L, Dutta A. 2001. *J. Biol. Chem.* 276:29067–71
41. Vashee S, Simancek P, Challberg MD, Kelly TJ. 2001. *J. Biol. Chem.* 276: 26666–73
42. Gavin KA, Hidaka M, Stillman B. 1995. *Science* 270:1667–71
43. Quintana DG, Hou ZH, Thome KC, Hendricks M, Saha P, Dutta A. 1997. *J. Biol. Chem.* 272:28247–51
44. Quintana DG, Thome KC, Hou ZH, Ligon AH, Morton CC, Dutta A. 1998. *J. Biol. Chem.* 273:27137–45
45. Tugal T, Zou-Yang XH, Gavin K, Pappin D, Canas B, et al. 1998. *J. Biol. Chem.* 273:32421–29
46. Aparicio OM, Weinstein DM, Bell SP. 1997. *Cell* 91:59–69
47. Santocanale C, Diffley JF. 1996. *EMBO J.* 15:6671–79
48. Tanaka T, Knapp D, Nasmyth K. 1997. *Cell* 90:649–60
49. Pak DT, Pflumm M, Chesnokov I, Huang DW, Kellum R, et al. 1997. *Cell* 91:311–23
49a. Wyrick JJ, Aparicio JG, Chen T, Barnett JD, Jennings EG, et al. 2001. *Science* 294:2357–60
50. Mizushima T, Takahashi N, Stillman B. 2000. *Genes Dev.* 14:1631–41

51. Burke TW, Cook JG, Asano M, Nevins JR. 2001. *J. Biol. Chem.* 276:15397–408
52. Iizuka M, Stillman B. 1999. *J. Biol. Chem.* 274:23027–34
53. Royzman I, Austin RJ, Bosco G, Bell SP, Orr-Weaver TL. 1999. *Genes Dev.* 13:827–40
54. Bosco G, Du W, Orr-Weaver TL. 2001. *Nat. Cell Biol.* 3:289–95
55. Mahbubani HM, Paull T, Elder JK, Blow JJ. 1992. *Nucleic Acids Res.* 20:1457–62
56. Rowles A, Tada S, Blow JJ. 1999. *J. Cell Sci.* 112:2011–18
57. Blow JJ, Gillespie PJ, Francis D, Jackson DA. 2001. *J. Cell Biol.* 152:15–25
58. Lucas I, Chevrier-Miller M, Sogo JM, Hyrien O. 2000. *J. Mol. Biol.* 296:769–86
59. Hartwell LH. 1973. *J. Bacteriol.* 115: 966–74
60. Neuwald AF, Aravind L, Spouge JL, Koonin EV. 1999. *Genome Res.* 9:27–43
61. Coleman TR, Carpenter PB, Dunphy WG. 1996. *Cell* 87:53–63
62. Elsasser S, Lou F, Wang B, Campbell JL, Jong A. 1996. *Mol. Biol. Cell* 7:1723–35
63. Perkins G, Diffley JF. 1998. *Mol. Cell* 2:23–32
64. Weinreich M, Liang C, Stillman B. 1999. *Proc. Natl. Acad. Sci. USA* 96:441–46
65. DeRyckere D, Smith CL, Martin GS. 1999. *Genetics* 151:1445–57
66. Liu J, Smith CL, DeRyckere D, DeAngelis K, Martin GS, Berger JM. 2000. *Mol. Cell* 6:637–48
67. Herbig U, Marlar CA, Fanning E. 1999. *Mol. Biol. Cell* 10:2631–45
68. Drury LS, Perkins G, Diffley JF. 1997. *EMBO J.* 16:5966–76
69. Elsasser S, Chi Y, Yang P, Campbell JL. 1999. *Mol. Biol. Cell* 10:3263–77
70. Perkins G, Drury LS, Diffley JF. 2001. *EMBO J.* 20:4836–45
71. Piatti S, Bohm T, Cocker JH, Diffley JF, Nasmyth K. 1996. *Genes Dev.* 10:1516–31
72. Jallepalli PV, Tien D, Kelly TJ. 1998. *Proc. Natl. Acad. Sci. USA* 95:8159–64
73. Wolf DA, McKeon F, Jackson PK. 1999. *Mol. Gen. Genet.* 262:473–80
74. Calzada A, Sanchez M, Sanchez E, Bueno A. 2000. *J. Biol. Chem.* 275: 9734–41
75. Drury LS, Perkins G, Diffley JF. 2000. *Curr. Biol.* 10:231–40
76. Fujita M, Yamada C, Goto H, Yokoyama N, Kuzushima K, et al. 1999. *J. Biol. Chem.* 274:25927–32
77. Jiang W, Wells NJ, Hunter T. 1999. *Proc. Natl. Acad. Sci. USA* 96:6193–98
78. Petersen BO, Wagener C, Marinoni F, Kramer ER, Melixetian M, et al. 2000. *Genes Dev.* 14:2330–43
79. Delmolino LM, Saha P, Dutta A. 2001. *J. Biol. Chem.* 276:26947–54
80. Mendez J, Stillman B. 2000. *Mol. Cell. Biol.* 20:8602–12
81. Gajiwala KS, Burley SK. 2000. *Curr. Opin. Struct. Biol.* 10:110–16
82. Feng L, Wang B, Driscoll B, Jong A. 2000. *Mol. Biol. Cell* 11:1673–85
83. Jeruzalmi D, O'Donnell M, Kuriyan J. 2001. *Cell* 106:429–41
84. Jeruzalmi D, Yurieva O, Zhao YX, Young M, Stewart J, et al. 2001. *Cell* 106:417–28
85. Liang C, Stillman B. 1997. *Genes Dev.* 11:3375–86
86. Schepers A, Diffley JF. 2001. *J. Mol. Biol.* 308:597–608
87. Hofmann JF, Beach D. 1994. *EMBO J.* 13:425–34
88. Nishitani H, Lygerou Z, Nishimoto T, Nurse P. 2000. *Nature* 404:625–28
89. Maiorano D, Moreau J, Mechali M. 2000. *Nature* 404:622–25
90. Whittaker AJ, Royzman I, Orr-Weaver TL. 2000. *Genes Dev.* 14:1765–76
91. Wohlschlegel JA, Dwyer BT, Dhar SK, Cvetic C, Walter JC, Dutta A. 2000. *Science* 290:2309–12

92. Chong JP, Thommes P, Blow JJ. 1996. *Trends Biochem. Sci.* 21:102–6
93. Dutta A, Bell SP. 1997. *Annu. Rev. Cell Dev. Biol.* 13:293–332
94. Kelly TJ, Brown GW. 2000. *Annu. Rev. Biochem.* 69:829–80
95. Chong JP, Mahbubani HM, Khoo CY, Blow JJ. 1995. *Nature* 375:418–21
96. Kubota Y, Mimura S, Nishimoto S, Masuda T, Nojima H, Takisawa H. 1997. *EMBO J.* 16:3320–31
97. Madine MA, Khoo CY, Mills AD, Laskey RA. 1995. *Nature* 375:421–24
98. Hua XH, Newport J. 1998. *J. Cell Biol.* 140:271–81
99. Tanaka H, Tanaka K, Murakami H, Okayama H. 1999. *Mol. Cell. Biol.* 19:1038–48
100. Zou L, Stillman B. 2000. *Mol. Cell. Biol.* 20:3086–96
101. Labib K, Tercero JA, Diffley JF. 2000. *Science* 288:1643–47
102. Lipford JR, Bell SP. 2001. *Mol. Cell* 7:21–30
103. Zou L, Stillman B. 1998. *Science* 280: 593–96
103a. Wilmes GM, Bell SP. 2002. *Proc. Natl. Acad. Sci. USA* 99:101–6
104. You ZY, Komamura Y, Ishimi Y. 1999. *Mol. Cell. Biol.* 19:8003–15
105. Ishimi Y. 1997. *J. Biol. Chem.* 272: 24508–13
106. Lee JK, Hurwitz J. 2000. *J. Biol. Chem.* 275:18871–78
107. Lee JK, Hurwitz J. 2001. *Proc. Natl. Acad. Sci. USA* 98:54–59
108. Sato M, Gotow T, You ZY, Komamura-Kohno Y, Uchiyama Y, et al. 2000. *J. Mol. Biol.* 300:421–31
109. Chong JP, Hayashi MK, Simon MN, Xu RM, Stillman B. 2000. *Proc. Natl. Acad. Sci. USA* 97:1530–35
110. Kelman Z, Lee JK, Hurwitz J. 1999. *Proc. Natl. Acad. Sci. USA* 96:14783–88
111. Shechter DF, Ying CY, Gautier J. 2000. *J. Biol. Chem.* 275:15049–59
112. Adachi Y, Usukura J, Yanagida M. 1997. *Genes Cells* 2:467–79
113. Ishimi Y, Ichinose S, Omori A, Sato K, Kimura H. 1996. *J. Biol. Chem.* 271: 24115–22
114. Maiorano D, Lemaitre JM, Mechali M. 2000. *J. Biol. Chem.* 275:8426–31
115. Prokhorova TA, Blow JJ. 2000. *J. Biol. Chem.* 275:2491–98
115a. Schwacha A, Bell SP. 2001. *Mol. Cell* 8:1093–104
116. Boyer PD. 1993. *Biochim. Biophys. Acta* 1140:215–50
117. Hingorani MM, Washington MT, Moore KC, Patel SS. 1997. *Proc. Natl. Acad. Sci. USA* 94:5012–17
118. Labib K, Diffley JF, Kearsey SE. 1999. *Nat. Cell Biol.* 1:415–22
119. Nguyen VQ, Co C, Irie K, Li JJ. 2000. *Curr. Biol.* 10:195–205
120. Pasion SG, Forsburg SL. 1999. *Mol. Biol. Cell* 10:4043–57
121. Lei M, Tye BK. 2001. *J. Cell Sci.* 114:1447–54
122. Burkhart R, Schulte D, Hu D, Musahl C, Gohring F, Knippers R. 1995. *Eur. J. Biochem.* 228:431–38
123. Donovan S, Diffley JF. 1996. *Curr. Opin. Genet. Dev.* 6:203–7
124. Lei M, Kawasaki Y, Tye BK. 1996. *Mol. Cell. Biol.* 16:5081–90
125. Mahbubani HM, Chong JP, Chevalier S, Thommes P, Blow JJ. 1997. *J. Cell Biol.* 136:125–35
126. Walter J, Newport JW. 1997. *Science* 275:993–95
127. Labib K, Diffley JF. 2001. *Curr. Opin. Genet. Dev.* 11:64–70
128. Seki T, Diffley JF. 2000. *Proc. Natl. Acad. Sci. USA* 97:14115–20
129. Bell SP, Mitchell J, Leber J, Kobayashi R, Stillman B. 1995. *Cell* 83:563–68
130. Geraghty DS, Ding M, Heintz NH, Pederson DS. 2000. *J. Biol. Chem.* 275: 18011–21
131. Tanaka T, Nasmyth K. 1998. *EMBO J.* 17:5182–91

132. Dahmann C, Diffley JF, Nasmyth KA. 1995. *Curr. Biol.* 5:1257–69
133. Nishitani H, Nurse P. 1995. *Cell* 83:397–405
134. Detweiler CS, Li JJ. 1998. *Proc. Natl. Acad. Sci. USA* 95:2384–89
135. Hua XH, Yan H, Newport J. 1997. *J. Cell Biol.* 137:183–92
136. Herbig U, Griffith JW, Fanning E. 2000. *Mol. Biol. Cell* 11:4117–30
137. Petersen BO, Lukas J, Sorensen CS, Bartek J, Helin K. 1999. *EMBO J.* 18:396–410
138. Jallepalli PV, Brown GW, Muzi-Falconi M, Tien D, Kelly TJ. 1997. *Genes Dev.* 11:2767–79
139. Lopez-Girona A, Mondesert O, Leatherwood J, Russell P. 1998. *Mol. Biol. Cell* 9:63–73
140. Findeisen M, El-Denary M, Kapitza T, Graf R, Strausfeld U. 1999. *Eur. J. Biochem.* 264:415–26
141. Fujita M, Yamada C, Tsurumi T, Hanaoka F, Matsuzawa K, Inagaki M. 1998. *J. Biol. Chem.* 273:17095–101
142. Hendrickson M, Madine M, Dalton S, Gautier J. 1996. *Proc. Natl. Acad. Sci. USA* 93:12223–28
143. Ishimi Y, Komamura-Kohno Y. 2001. *J. Biol. Chem.* 276:34428–33
144. Pereverzeva I, Whitmire E, Khan B, Coue M. 2000. *Mol. Cell. Biol.* 20:3667–76
145. Hennessy KM, Clark CD, Botstein D. 1990. *Genes Dev.* 4:2252–63
146. Coue M, Kearsey SE, Mechali M. 1996. *EMBO J.* 15:1085–97
147. Nguyen VQ, Co C, Li JJ. 2001. *Nature* 411:1068–73
148. Vas A, Mok W, Leatherwood J. 2001. *Mol. Cell. Biol.* 21:5767–77
149. McGarry TJ, Kirschner MW. 1998. *Cell* 93:1043–53
150. Tada S, Li A, Maiorano D, Mechali M, Blow JJ. 2001. *Nat. Cell Biol.* 3:107–13
150a. Gopalakrishnan V, Simancek P, Houchens C, Snaith HA, Frattini MG, et al. 2001. *Proc. Natl. Acad. Sci. USA* 98:13114–19
150b. Yanow SK, Lygerou Z, Nurse P. 2001. *EMBO J.* 20:4648–56
151. Diller JD, Raghuraman MK. 1994. *Trends Biochem. Sci.* 19:320–25
152. Blumenthal AB, Kriegstein HJ, Hogness DS. 1974. *Cold Spring Harbor Symp. Quant. Biol.* 38:205–23
153. Hyrien O, Maric C, Mechali M. 1995. *Science* 270:994–97
154. Alexiadis V, Varga-Weisz PD, Bonte E, Becker PB, Gruss C. 1998. *EMBO J.* 17:3428–38
155. Wittmeyer J, Joss L, Formosa T. 1999. *Biochemistry* 38:8961–71
156. Thoma F, Bergman LW, Simpson RT. 1984. *J. Mol. Biol.* 177:715–33
157. Simpson RT. 1990. *Nature* 343:387–89
158. Gillespie PJ, Blow JJ. 2000. *Nucleic Acids Res.* 28:472–80
159. Wu JR, Gilbert DM. 1996. *Science* 271: 1270–72
160. Gilbert DM, Miyazawa H, DePamphilis ML. 1995. *Mol. Cell. Biol.* 15:2942–54
161. Solomon NA, Wright MB, Chang S, Buckley AM, Dumas LB, Gaber RF. 1992. *Yeast* 8:273–89
162. Maine GT, Sinha P, Tye BK. 1984. *Genetics* 106:365–85
163. Merchant AM, Kawasaki Y, Chen YR, Lei M, Tye BK. 1997. *Mol. Cell. Biol.* 17:3261–71
164. Homesley L, Lei M, Kawasaki Y, Sawyer S, Christensen T, Tye BK. 2000. *Genes Dev.* 14:913–26
165. Kawasaki Y, Hiraga S, Sugino A. 2000. *Genes Cells* 5:975–89
165a. Wohlschlegel JA, Dhar SK, Prokhorova TA, Dutta A, Walter JC, 2002. *Mol. Cell.* 9:233–40
166. Izumi M, Yanagi K, Mizuno T, Yokoi M, Kawasaki Y, et al. 2000. *Nucleic Acids Res.* 28:4769–77
167. Moir D, Stewart SE, Osmond BC, Botstein D. 1982. *Genetics* 100:547–63
168. Mimura S, Takisawa H. 1998. *EMBO J.* 17:5699–707

169. Aparicio OM, Stout AM, Bell SP. 1999. *Proc. Natl. Acad. Sci. USA* 96:9130–35
170. Kukimoto I, Igaki H, Kanda T. 1999. *Eur. J. Biochem.* 265:936–43
171. Mimura S, Masuda T, Matsui T, Takisawa H. 2000. *Genes Cells* 5:439–52
172. Walter J, Newport J. 2000. *Mol. Cell* 5:617–27
173. Kamimura Y, Tak YS, Sugino A, Araki H. 2001. *EMBO J.* 20:2097–107
174. Saha P, Thome KC, Yamaguchi R, Hou ZH, Weremowicz S, Dutta A. 1998. *J. Biol. Chem.* 273:18205–9
175. Uchiyama M, Griffiths D, Arai K, Masai H. 2001. *J. Biol. Chem.* 276: 26189–96
176. Tercero JA, Labib K, Diffley JF. 2000. *EMBO J.* 19:2082–93
177. Araki H, Leem SH, Phongdara A, Sugino A. 1995. *Proc. Natl. Acad. Sci. USA* 92:11791–95
178. Masumoto H, Sugino A, Araki H. 2000. *Mol. Cell. Biol.* 20:2809–17
179. McFarlane RJ, Carr AM, Price C. 1997. *Mol. Gen. Genet.* 255:332–40
180. Wang H, Elledge SJ. 1999. *Proc. Natl. Acad. Sci. USA* 96:3824–29
181. Kelly TJ, Martin GS, Forsburg SL, Stephen RJ, Russo A, Nurse P. 1993. *Cell* 74:371–82
182. Grallert B, Nurse P. 1996. *Genes Dev.* 10:2644–54
183. Kamimura Y, Masumoto H, Sugino A, Araki H. 1998. *Mol. Cell. Biol.* 18: 6102–9
184. Deleted in proof
185. Pasero P, Duncker BP, Schwob E, Gasser SM. 1999. *Genes Dev.* 13:2159–76
186. Jares P, Blow JJ. 2000. *Genes Dev.* 14:1528–40
187. Walter JC. 2000. *J. Biol. Chem.* 275: 39773–78
188. Weinreich M, Stillman B. 1999. *EMBO J.* 18:5334–46
188a. Sclafani RA. 2000. *J. Cell Sci.* 113: 2111–17
188b. Hardy CFJ, Dryga O, Seematter S, Pahl PMB, Sclafani RA. 1997. *Proc. Natl. Acad. Sci. USA* 94:3151–55
189. Lei M, Kawasaki Y, Young MR, Kihara M, Sugino A, Tye BK. 1997. *Genes Dev.* 11:3365–74
190. Roberts BT, Ying CY, Gautier J, Maller JL. 1999. *Proc. Natl. Acad. Sci. USA* 96:2800–4
191. Deleted in proof
192. Kihara M, Nakai W, Asano S, Suzuki A, Kitada K, et al. 2000. *J. Biol. Chem.* 275:35051–62
193. Oshiro G, Owens JC, Shellman Y, Sclafani RA, Li JJ. 1999. *Mol. Cell. Biol.* 19:4888–96
194. Brown GW, Kelly TJ. 1998. *J. Biol. Chem.* 273:22083–90
195. Jiang W, Hunter T. 1997. *Proc. Natl. Acad. Sci. USA* 94:14320–25
196. Masai H, Matsui E, You Z, Ishimi Y, Tamai K, Arai K. 2000. *J. Biol. Chem.* 275:29042–52
197. Nougarede R, Della Seta F, Zarzov P, Schwob E. 2000. *Mol. Cell. Biol.* 20:3795–806
198. Bousset K, Diffley JF. 1998. *Genes Dev.* 12:480–90
199. Donaldson AD, Fangman WL, Brewer BJ. 1998. *Genes Dev.* 12:491–501
200. Owens JC, Detweiler CS, Li JJ. 1997. *Proc. Natl. Acad. Sci. USA* 94:12521–26
201. Furstenthal L, Kaiser BK, Swanson C, Jackson PK. 2001. *J. Cell Biol.* 152: 1267–78
202. Adams PD, Sellers WR, Sharma SK, Wu AD, Nalin CM, Kaelin WG Jr. 1996. *Mol. Cell. Biol.* 16:6623–33
203. Chen JJ, Saha P, Kornbluth S, Dynlacht BD, Dutta A. 1996. *Mol. Cell. Biol.* 16:4673–82
204. Wohlschlegel JA, Dwyer BT, Takeda DY, Dutta A. 2001. *Mol. Cell. Biol.* 21:4868–74
205. Schulman BA, Lindstrom DL, Harlow E. 1998. *Proc. Natl. Acad. Sci. USA* 95:10453–58

206. Takeda DY, Wohlschlegel JA, Dutta A. 2001. *J. Biol. Chem.* 276:1993–97
207. Calzada A, Sacristan M, Sanchez E, Bueno A. 2001. *Nature* 412:355–58
208. Weinreich M, Liang C, Chen HH, Stillman B. 2001. *Proc. Natl. Acad. Sci. USA* 98:11211–17
209. Romanowski P, Marr J, Madine MA, Rowles A, Blow JJ, et al. 2000. *J. Biol. Chem.* 275:4239–43
210. Zou L, Mitchell J, Stillman B. 1997. *Mol. Cell. Biol.* 17:553–63
211. Pelizon C, Madine MA, Romanowski P, Laskey RA. 2000. *Genes Dev.* 14: 2526–33
212. Dowell SJ, Romanowski P, Diffley JF. 1994. *Science* 265:1243–46
213. Friedman KL, Diller JD, Ferguson BM, Nyland SV, Brewer BJ, Fangman WL. 1996. *Genes Dev.* 10:1595–607
214. Raghuraman MK, Brewer BJ, Fangman WL. 1997. *Science* 276:806–9
215. Dimitrova DS, Gilbert DM. 1999. *Mol. Cell* 4:983–93
216. Santocanale C, Diffley JF. 1998. *Nature* 395:615–18
217. Shirahige K, Hori Y, Shiraishi K, Yamashita M, Takahashi K, et al. 1998. *Nature* 395:618–21
218. Donaldson AD, Raghuraman MK, Friedman KL, Cross FR, Brewer BJ, Fangman WL. 1998. *Mol. Cell* 2:173–82
219. Donaldson AD. 2000. *EMBO Rep.* 1:507–12
220. Brown GW, Kelly TJ. 1999. *Proc. Natl. Acad. Sci. USA* 96:8443–48
221. Cheng LA, Collyer T, Hardy CFJ. 1999. *Mol. Cell. Biol.* 19:4270–78
222. Takeda T, Ogino K, Matsui E, Cho MK, Kumagai H, et al. 1999. *Mol. Cell. Biol.* 19:5535–47
223. Dohrmann PR, Oshiro G, Tecklenburg M, Sclafani RA. 1999. *Genetics* 151: 965–77
224. Jares P, Donaldson A, Blow JJ. 2000. *EMBO Rep.* 1:319–22
225. Lee DG, Bell SP. 2000. *Curr. Opin. Cell Biol.* 12:280–85
226. Herrick J, Bensimon A. 1999. *Biochimie* 81:859–71
227. Huberman JA. 1968. *Cold Spring Harbor Symp. Quant. Biol.* 33: 509–24
228. Fang L, Davey MJ, O'Donnell M. 1999. *Mol. Cell* 4:541–53
229. Lemon KP, Grossman AD. 2001. *Genes Dev.* 15:2031–41

*Annual Review of Biochemistry*
*Volume 71, 2002*

# CONTENTS

ERRATA

An online log of corrections to *Annual Review of Biochemistry* chapters may be found at http://biochem.annualreviews.org/errata.shtml